

Conformational Dynamics and Stability of Structured Peptides and Small Proteins

INAUGURALDISSERTATION

zur

Erlangung der Würde eines Doktors der Philosophie

vorgelegt der

Philosophisch-Naturwissenschaftlichen Fakultät

der Universität Basel

von

Andreas Reiner

aus Deutschland

Basel, 2009

Genehmigt von der Philosophisch-Naturwissenschaftlichen Fakultät
auf Antrag von

Prof. Dr. Thomas Kiefhaber

Prof. Dr. Joachim Seelig

Basel, den 11. Dezember 2007

Prof. Dr. Hans-Peter Hauri

Dekan

Table of Contents

1	Introduction	1
1.1	Proteins and Protein Folding.....	1
1.1.1	The Native State.....	2
1.1.2	The Free Energy Surface of Proteins.....	3
1.1.3	Kinetics of Protein Folding.....	4
1.1.4	The Unfolded State Ensemble.....	8
1.2	Elementary Steps in Protein Folding.....	9
1.2.1	Formation of Contacts.....	10
1.2.1.1	Triplet-Triplet Energy Transfer (TTET).....	11
1.2.1.2	Chain Dynamics in Unstructured Polypeptides.....	13
1.2.2	Formation of α -Helices	15
1.2.2.1	Thermodynamic Aspects of α -Helix Formation.....	16
1.2.2.2	Dynamics of α -Helices.....	17
1.2.3	Formation of Hairpins and β -Sheets.....	18
1.3	The Fibrin Foldon Domain as a Model System for Oligomerisation.....	19
1.4	The Villin Headpiece Subdomain as a Small Model Protein.....	21
2	Aims of Research	23
3	Summary of Work Ready for Submission.....	26
3.1	Dynamics of α -Helical Peptides Probed by TTET.....	26
3.2	Thioamide Backbone Substitutions in α -Helices	29
3.3	Structure and Dynamics of the Foldon Monomer.....	31
3.4	Contact Formation Dynamics in the Villin Headpiece Subdomain.....	33
4	Summary of Unpublished Results.....	36
4.1	Length Dependence of Dynamics in α -Helical Peptides.....	36
4.1.1	Design and Helix Content of Peptides.....	36
4.1.2	Local Dynamics in the Helix Centre Measured by TTET.....	38
4.1.3	Comparison to Predictions from a Kinetic Ising Model.....	43
4.2	Peptide Model of the Foldon Hairpin Motif.....	48
4.2.1	Structure and Stability.....	48
4.2.2	TTET Measurements to Study Hairpin Dynamics.....	58
4.2.3	Material and Methods.....	63
4.3	Incorporation of Triplet Donor and Acceptor Labels via Cysteines.....	65
4.3.1	Incorporation of Naphthalene.....	65
4.3.2	Incorporation of Xanthone.....	67
4.3.3	Material and Methods.....	69
5	Summary.....	72

6	References	75
7	Appendix	90
8	Curriculum Vitae.....	95
9	Acknowledgements.....	96
10	Manuscripts.....	97
10.A	Local Conformational Dynamics in α -Helices Measured by Fast Triplet Transfer	A
10.B	Temperature Dependence of Local Equilibrium Folding and Unfolding of an α -Helix	B
10.C	Effect of Thiopeptide Bonds on α -Helix Structure and Stability.....	C
10.D	NMR Structure of a Monomeric Intermediate on the Evolutionarily Optimized Assembly Pathway of the Trimeric Foldon Domain.....	D
10.E	Conformational Dynamics in the Native and Denatured State of the Villin Headpiece Subdomain Measured by Triplet-Triplet Energy Transfer.....	E

1 Introduction

1.1 Proteins and Protein Folding

In the course of evolution, proteins, which are linear polymers of α -amino acids, have become the most versatile class of biomolecules, playing a central role in nearly all biological processes.¹ Unique to proteins is the formation of complex structures that are highly adaptable to specific needs, for example encountered in catalysis, transport, signal transduction, structural organisation or regulation.

The sequences in which the twenty different α -L-amino acids are combined to proteins are genetically encoded, ensuring both, stable inheritance and amenability to evolutionary processes. Based on this information proteins are synthesised at the ribosomes, which they leave through a narrow exit tunnel in a mainly unfolded, i.e. unstructured form. The structure of proteins is fully encoded in their sequence, as folding does not depend on the ribosome or any additional folding machinery, but occurs spontaneously.² The folded state of a protein usually remains in a dynamic equilibrium with unfolded conformations and hence most proteins fold and unfold many times during their lifetime. *In vivo*, accessory proteins like chaperones,³ disulfide isomerases⁴ and peptidyl-prolyl isomerases⁵ are present to ensure efficient folding. This is of particular importance for larger proteins, which tend to aggregate as long as they are unfolded. Aggregation of proteins and misfolding into fibrillar structures are related to a number of diseases.

Although the building principle of proteins is rather simple, it gives rise to highly diverse structures, much more complex than systems that have been realized in synthetic chemistry.⁶ The polypeptide chain is typically arranged in local secondary structure elements, which are combined to domains, the tertiary structure. Different domains might be further linked to multi-domain proteins or assembled to oligomers.

The folded state of proteins is in all cases a delicate balance of innumerable weak, non-covalent and competing interactions. The biological function often relies on subtle changes in this network of interactions, for example during a catalytic cycle or as response to ligand binding. It is puzzling how this balance, starting from an unfolded polypeptide chain, can be formed so efficiently.⁷⁻⁹ To achieve a comprehensive understanding of these phenomena, which are peculiar to proteins, remains a major challenge in molecular biology, chemistry and physics.

1.1.1 The Native State

The native state denotes the preferred state under physiological conditions, which is related to the function of the protein and is usually well structured. An exception are so-called natively unfolded proteins, in which structure formation is mostly coupled with the binding of interaction partners.¹⁰

The structures of proteins belonging to many different classes have been determined in atomic detail¹¹ using X-ray crystallography^{12,13} or nuclear magnetic resonance (NMR) spectroscopy.¹⁴ The interactions responsible for the formation of structures had been proposed from chemical considerations¹⁵⁻¹⁷ and were confirmed by the structural data and mutational studies.^{1,18} The majority of the backbone amide groups is involved in hydrogen bonds (H-bonds),¹⁹ which often define regular secondary structures (see 1.2.2 and 1.2.3). Charged amino acid side chains and the dipolar character of the amide bond give rise to long-range electrostatic interactions. Charged residues are typically distributed all over the protein surface to form networks of electrostatic interactions,²⁰ but defined salt-bridges can be observed as well.²¹ Another significant driving force for folding arises from the hydrophobic effect.^{16,22} The entropy loss of water molecules getting ordered around hydrophobic groups causes a minimization of the hydrophobic surface - a densely packed hydrophobic core is common to all globular proteins. The dense packing of proteins adds extensive van der Waals interactions.

More difficult to capture are the dynamics of the native state, which is not a rigid structure but subject to thermal fluctuations. Motions occur on many timescales, in correlated or uncorrelated fashion. Single bonds vibrate within some femtoseconds, whereas the rotation of unhindered side chains and the exchange of loosely associated water molecules take place in picoseconds.²³ Loop motions typically require nanoseconds, large amplitude motions of domains and local unfolding events microseconds to milliseconds. Dynamics in folded proteins can extend to hours²⁴ and global unfolding is in some cases so slow that it will be hardly observed. Flexibility is clearly a prerequisite for the function of proteins. Classical examples are the induced fit of enzymes²⁵ and the various experiments on motions in myoglobin,²⁶⁻²⁸ which led to the concept of conformational substates. The dynamic picture of proteins is now complemented by single molecule spectroscopy²⁹ and new NMR techniques,³⁰ providing insight into the action of molecular machines and the concerted motions that enzymes undergo in the course of catalytic cycles.^{31,32}

1.1.2 The Free Energy Surface of Proteins

The thermodynamic stability of the native state turns out to be rather low in many cases. Proteins can be readily denatured, applying perturbations such as high temperature, high pressure, extreme pH or denaturing agents like urea or guanidinium chloride (GdmCl). However, a key observation of these *in vitro* experiments is that denaturation of proteins is a reversible process,³³⁻³⁶ i.e. proteins can refold spontaneously into the native conformation, when the conditions are changed back again. This demonstrates that the formation of structure is fully encoded in the amino acid sequence and does not depend on any accessory factors. Anfinsen concluded that the folded state is attained, simply because it is the state, in which the free energy of the system is minimal.²

A full, statistical thermodynamic description of a protein chain and the surrounding solvent molecules would encompass all possible conformations, populated according to their free energies. However, most small single-domain proteins unfold in a single, cooperative transition, which can be fully described by a simple two-state model. This assumes that the protein molecule is either in the unfolded state U or in the folded state N and no intermediate states are populated at equilibrium (Eq. 1).



U and N are thermodynamic states, implying that they encompass many microscopic states of different energies, in this case ensembles of conformations, which equilibrate fast compared to the overall reaction. The relative population of U and N can be expressed by an equilibrium constant K (Eq. 2a), which is related to the difference in the Gibbs free energy ΔG° between the two macroscopic states (Eq. 2b).

$$K = \frac{[N]}{[U]} \quad K = e^{-\Delta G^\circ/RT} \quad (2a,b)$$

The thermodynamic stability of typical single-domain proteins, arising from a subtle balance of entropic and enthalpic contributions, is in the range of -15 to -60 kJ mol⁻¹.³⁷ These stabilities are low, especially taking into account, that a single salt-bridge or H-bond can attribute up to ~ 20 kJ mol⁻¹.^{17,21} Reasons for such low stabilities might well be that the native state has to be dynamic (see 1.1.1) and that proteins have to be unfolded in order to get degraded.

Since ΔG° is small, the stability of proteins is sensitive to all changes that affect the differences in enthalpy ΔH° , entropy ΔS° or chemical potential $\Delta\mu^\circ$ between the ground states. Furthermore, the thermodynamic equilibrium can be shifted by varying the

temperature T , the pressure p or by adding components n_i that affect the chemical potential (Eq. 3).

$$d\Delta G^\circ = \Delta V^\circ \cdot dp - \Delta S^\circ \cdot dT + \sum_i \Delta \mu_i^\circ \cdot dn_i \quad (3)$$

ΔV° denotes the volume change upon folding. It should be noted that ΔH° and ΔS° are temperature-dependent themselves, since the unfolded state has typically a larger heat capacity than the folded state ($\Delta C_p < 0$).³⁸ As a consequence, the stability of a protein is maximal at a certain temperature, but in most cases significant cold denaturation is only expected for temperatures below 0 °C. Common denaturants such as urea and GdmCl favour the exposure of surface and hence unfolding, by either improving the solvation of the polypeptide chain³⁶ or direct binding.^{15,39} The change in ΔG° is usually linearly related to the molar concentration of denaturant⁴⁰ and to the change in accessible surface area upon unfolding.⁴¹

The discussed two-state behaviour appears to be an oversimplification, which is certainly true for structures of low cooperativity (e.g. α -helices, see 1.2.2) or multi-domain proteins with several folding units. However, the observed equality of model derived enthalpies and calorimetrically measured enthalpies of many small single-domain proteins is a rigorous proof for the validity of this description.^{38,42}

1.1.3 Kinetics of Protein Folding

The conformational space, accessible to the backbone and the side chains of a polypeptide, is vast. It was pointed out early that protein folding cannot proceed via a random sampling of all conformations, since this would take an astronomic time.⁴³ Moreover, this view would neglect that different conformations are associated with different free energies. Efficient folding seems to be ensured by a small free energy bias towards productive conformations,⁴⁴ often viewed as a funnel on an energy landscape.⁴⁵ Nevertheless, it is puzzling how the native state, a delicate balance of many opposing interactions, can be reached without getting trapped in local minima of the free energy surface. Another question is, how the energy bias towards the native state is achieved, especially in the early stages of folding when only very few specific contacts exist at all.

The times that proteins require to fold from an unfolded state to the native state range from microseconds⁴⁶⁻⁴⁸ to several seconds.^{37,49} Additional slow processes might arise from the formation of disulfide bridges,^{50,51} steps limited by prolyl *cis-trans* isomerisation^{52,53} or the incorporation of cofactors. The folding and unfolding kinetics of

many small proteins are single-exponential and without any deviations from the behaviour expected for two-state reactions.^{37,54,55} In cases where folding is initiated by a fast change of solvent conditions a rapid chain collapse might be observed, which is not necessarily related to the folding reaction, but might reflect the adaptation of the unfolded state ensemble to the new solvent conditions. However, the collapse may also lead to specific structures or intermediates, separated from the unfolded state by a free energy barrier.⁵⁶ Multi-exponential folding kinetics^{57,58} or disagreement with interrupted refolding experiments⁵⁹ or equilibrium data point toward the presence of kinetic intermediates.⁵⁶ These are not necessarily obligatory on-pathway intermediates, but might belong to parallel pathways⁶⁰ or even might be dead-end intermediates.⁶¹ In addition to populated intermediates, which slow down folding, intermediates higher in energy as N and U have been observed as well, as indicated by non-linear rate equilibrium free energy relationships (see below).⁶² Intermediates are also commonly observed for protein unfolding reactions.⁶³

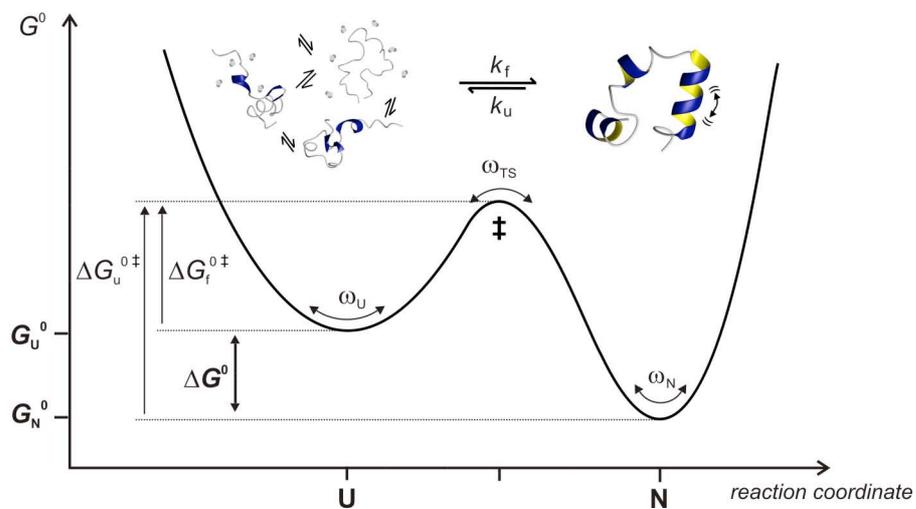


Figure 1.1. Free energy profile of a two-state folding reaction. The Gibbs free energy is given as function of a suitable reaction coordinate, describing the progress of the reaction. The macroscopic states, unfolded (U) and native (N) are separated by a significant free energy barrier (transition state \ddagger , TS). The thermodynamic stability of the protein is determined by the free energy difference ΔG° , whereas the folding and unfolding rate constants depend on the differences to the free energies of activation $\Delta G_{f,u}^{0\ddagger}$ and on the motional frequencies ω at the different positions of the surface.

The presence of a significant energy barrier between the unfolded state and the native state (Figure 1.1) has been shown for many proteins. In cases where the microscopic rate constants of folding k_f and unfolding k_u can be obtained reliably, the transition state theory of Eyring⁶⁴ might be applied (Eq. 4),

$$k = k_0 \cdot e^{-\Delta G^{\ddagger}/RT} \quad (4)$$

where the pre-exponential factor k_0 is the maximal rate constant in the absence of an energy barrier and ΔG^{\ddagger} denotes the free energy difference between the ground state and the extra-thermodynamic transition state (activation free energy).

The free energy barriers can be further characterised using rate equilibrium free energy relationships (REFERs).⁶⁵ The effects of perturbations on ΔG^{\ddagger} are often linearly related to the changes of ΔG° . Neglecting effects on the unfolded state, a change of ΔG° fully caused by a change of $\Delta G_f^{\circ\ddagger}$ ($\alpha = 1$) indicates that the transition state responds like the native state, i.e. is native-like. On the other hand, if the change in ΔG° can be fully attributed to $\Delta G_u^{\circ\ddagger}$ ($\alpha = 0$), the transition state remains unaffected by the perturbation, resembling the unfolded state. This can be used to place the transition state on the corresponding reaction coordinate. Variation of the denaturant concentration gives information about the accessible surface area of the transition state, variation of the pressure about its volume⁶⁶ and variation of the temperature about its entropy, enthalpy and heat capacity.^{55,67-69} By extending this concept to structural perturbations, achieved by amino acid substitutions, the role of individual side chain interactions can be assessed (ϕ -values).⁷⁰ Comprehensive analysis of several proteins revealed that both, diffuse transition states in which the most interactions are partly formed, as well as polarized transition states with separated regions of low and high ϕ -values, exist.⁷⁰ Non-linear REFERs are also commonly observed and can give further insight into the nature of the rate-limiting steps⁶² or the ground-state properties.⁷¹ In several proteins showing apparent two-state behaviour, folding actually seems to be a sequential process over consecutive energy barriers and through obligatory high-energy intermediates.⁶² Apparent movements of the transition state, as often observed in mutational studies, do not necessarily result from broad energy barriers. They can be mostly attributed to changes in the unfolded state (see 1.1.4), pointing to residual structure that is sensitive towards mutation.⁷¹

For chemical reactions involving the formation and breakage of covalent bonds in the gas phase, the pre-exponential factor k_0 in Eq. 4 is related to bond vibrations, the elementary reaction of these processes.⁶⁴ However, protein folding takes place in solution and is at

least partially controlled by diffusion. A more appropriate description should hence be the theory of Kramers^{72,73} which treats the escape over a barrier as thermally activated, diffusional process on a free energy surface (Figure 1.1) and considers different degrees of coupling with the medium. For reactions in solution the high friction limit is applicable (Eq. 5),

$$k = \frac{\omega_0 \omega_{\ddagger}}{2\pi \cdot \gamma} \cdot e^{-E_b/RT} \quad (5)$$

in which the pre-exponential factor is proportional to the frequencies of the motions in the starting well ω_0 and at the saddle point of the barrier ω_{\ddagger} , but inversely proportional to the friction γ . As a consequence, protein folding and unfolding rate constants should be inversely proportional to solvent viscosity and indeed strong viscosity dependencies of folding/unfolding have been observed.^{74,75}

In summary, conformational transitions of proteins are usually well described by two or very few thermodynamically distinguishable states separated by one or few major energy barriers. However, this does not explain the folding process in atomic detail. On the microscopic level, folding remains a highly statistical process. Molecular dynamic simulations can give ideas about possible routes to the native state, but one has to bear in mind that force-fields are of very limited accuracy. Experimental techniques, which circumvent averaging over large ensembles, do neither provide the necessary time nor spatial resolution. Several concepts describing the folding pathway have been discussed.⁷⁰ Folding might start from local (secondary) structures, which are only marginally stable but act as nuclei for folding (nucleation-growth mechanism).⁷⁶ With a certain probability also several marginally stable secondary structures might co-exist at the same time, diffuse together and coalesce (diffusion-collision model).^{77,78} One can also argue that nuclei might exist but are only stable in combination with nonlocal interactions, i.e. the rate-limiting step would comprise nucleation and condensation.⁷⁹ It is also conceivable that the polypeptide chain collapses around its hydrophobic side chains (tertiary contacts) into a more or less native-like topology, which then rearranges to the native state (hydrophobic collapse model).⁸⁰ The chain collapse may considerably speed up the search for native-like contacts as the internal diffusion is limited to shorter distances,⁸¹ but on the other hand the internal friction might be increased.

1.1.4 The Unfolded State Ensemble

Under physiological conditions proteins rarely populate the unfolded state. Low pH, high temperature, pressure or denaturants can be used to shift the equilibrium towards the unfolded state (Eq. 3) and to induce structural transitions. However, it has to be taken into account that the unfolded state ensemble is highly sensitive towards these perturbations itself. Linear REFERs (see 1.1.3) indicate that these changes are continuous and do not have a major effect on the folding process. Alternatively, the properties of the unfolded state might be investigated under more physiological conditions with protein fragments⁸²⁻⁸⁶ or proteins destabilised by mutation,⁸⁷ truncation or reduction of disulfide bonds. In all cases the experimental characterisation of the unfolded state ensemble is complicated by the heterogeneity of conformations and their fast interconversion (see also 1.2.1).

In the absence of nonlocal interactions, the backbone will adopt the ϕ, ψ -angles in the accessible regions on the Ramachandran plot according to their energies,⁸⁸ resulting in a random-flight chain.^{89,90} To calculate the conformational propensities, others than only steric effects between nearest neighbours might have to be taken into account. Some conformations favour clashes with more distant residues and hydration,⁹¹ as well as electrostatic interactions, side chain - backbone interactions and hyperconjugation might be considered.⁹² Most theoretical studies suggest a preference for the broad β -sheet / polyproline II region, i.e. for relatively extended conformations. Circular dichroism (CD) and NMR data indicate that the polyproline II region is indeed populated in model peptides and unfolded proteins, but it is difficult to obtain quantitative estimates from the large continuum of different conformations. Most importantly, no cooperativity between different segments was found for the population of the polyproline II region.⁹³ Hence the bias towards these conformations will not give rise to structure in the unfolded state and is not at all a contradiction to random coil behaviour.

The random coil view³⁶ has long been supported by the agreement of experimentally determined properties like hydrodynamic radii or radii of gyration⁹⁴ with scaling laws derived from polymer theory.^{90,95,96} The radius of gyration R_g is predicted to depend on R_0 , which includes the persistence length, the number of residues N and the exponential scaling factor ν (Eq. 6).

$$R_g = R_0 \cdot N^\nu \quad (6)$$

An ideal and infinitely thin chain in a good solvent, outweighing all nonlocal interactions by solvent interactions, is described by $\nu = 0.5$. If the volume excluded by other parts of the chain (a nonlocal effect) is considered, one expects $\nu = 0.588$. A recent survey of

literature data and new measurements found $\nu = 0.598$ for 26 proteins in high concentrations of urea or GdmCl.⁹⁷ Local effects will increase the persistence length and would thus enter in R_0 . The random coil view, however, is challenged by several NMR studies, which indicate the presence of defined, nonlocal hydrophobic interactions and H-bonds, even in the presence of high concentrations of denaturant.⁹⁸⁻¹⁰¹ The interactions in the denatured state do not necessarily have to be native-like and since proteins are hetero-polymers, denaturants cannot be optimal solvents for all groups.³⁶

A different situation is encountered for the unfolded state under physiological conditions, where the interactions with the solvent do not compensate all intramolecular forces. In that case compact states are expected and found. In aqueous solution, the artificial population of the unfolded state over longer periods of time most often leads to unspecific aggregation or misfolding into β -fibrillar oligomers. Other examples are molten globule states, formed after disruption of charge interactions at low pH and mildly denaturing conditions,^{80,102} or collapsed states observed in kinetic experiments after dilution from high concentrations of denaturant. These states are still dynamic ensembles of many different conformations and it is not clear, whether the collapse leads to specific structures or not and how they are related to the folding pathway.^{103,104} In several cases collapse leads to intermediates showing cooperative unfolding behaviour.⁵⁶ In other proteins electrostatic interactions seem to play an important role for the organisation of the unfolded state, inferred from significantly shifted pK_a -values.¹⁰⁵ Several kinetic experiments revealed a change of the ground-state properties upon mutation of single residues, as well pointing to the importance of specific interactions.⁷¹

Studies on protein fragments have proven that local structural motifs, such as α -helices (see 1.2.2) and β -hairpins, can form in isolation (see 1.2.3). Under favourable conditions they are present in the unfolded state of some proteins, albeit they are only marginally stable and in fast equilibrium with unfolded conformations. Their role in the folding process is not understood in general.

1.2 Elementary Steps in Protein Folding

On the atomic level, folding has to start with a rapid sampling of many unfolded conformations eventually leading into a productive folding channel. The formation of marginally stable secondary structure elements and transient tertiary interactions are certainly important steps on the route to the native state, although these events do not have to lead to folding *per se*. In the following such elementary steps, namely the

formation of single intrachain contacts and the formation of α -helices and β -hairpins will be discussed.

1.2.1 Formation of Contacts

The unfolded state ensemble is highly dynamic and many transient contacts are formed, governed by the diffusion of the chain segments. These interactions may lead to local structure elements or certain topologies, as shown in Figure 1.2.

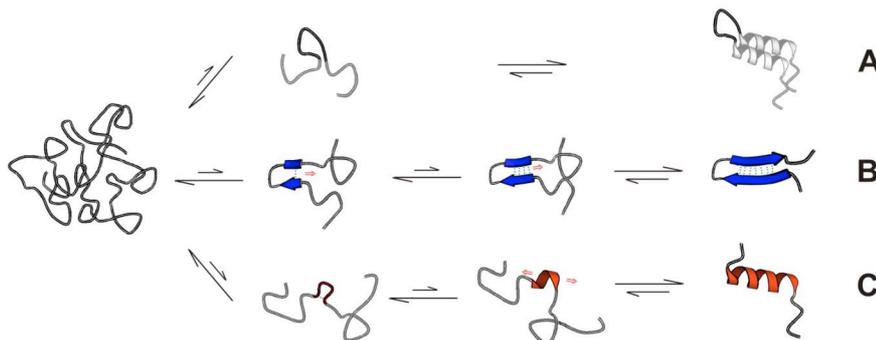


Figure 1.2. Contact formation as primary step to form structural elements such as (A) loops, (B) β -hairpins and (C) α -helices. Adapted from Krieger *et al.*¹⁰⁶

However, relative little is known about the dynamics of polypeptide chains, which determine how fast the complex free energy surface can be sampled. Furthermore, the chain dynamics may vary considerably in fully denatured proteins, collapsed states or folding intermediates. Using Förster resonance energy transfer (FRET), Haas and Steinberg obtained estimates of the intra-chain diffusion coefficients in short peptides by varying the solvent viscosity.¹⁰⁷ It is however difficult to separate the strongly correlated effects of the distance distribution between the labels and its change with time,¹⁰⁷ as also experienced by Lakowicz in a series of experiments on short polymers.¹⁰⁸ Two recent approaches try to avoid these complications using FRET donors with different lifetimes¹⁰⁹ or evaluating the autocorrelation of the donor intensity measured in single-molecule FRET measurements.¹¹⁰ The reported intrachain diffusion constants vary between $8 \cdot 10^{-8}$ and $1 \cdot 10^{-6} \text{ cm}^2 \text{ s}^{-1}$.

Another way to describe the dynamics of polypeptide chains is to measure the formation of contacts between specific groups. An early experiment addressed the dynamics in denatured cytochrome *c* by following the rebinding of methionine to the heme upon laser-induced dissociation of carbon monoxide.¹¹¹ The time constant of 35 - 40 μs obtained for the reaction between these groups separated by 50 - 60 peptide bonds was a significant overestimate of the time required to bring the two groups together.¹¹²

Nevertheless, this experiment triggered the development of more direct methods to measure contact formation,¹¹³ such as triplet-triplet energy transfer (see 1.2.1.1), triplet quenching or fluorescence quenching. Eaton and coworkers measured quenching of the tryptophan triplet state in a variety of peptides^{114,115-117} and proteins.¹¹⁸⁻¹²⁰ Since quenching is in many cases not completely diffusion controlled additional information is required to extract the absolute rate constants of contact formation. To employ fluorescence (singlet) quenching, either relatively long-lived chromophores have to be used¹²¹ or intensity fluctuations have to be measured by fluorescence correlation spectroscopy on a few or single molecules.^{122,123} The systematically lower rate constants of contact formation measured in such experiments indicate that neither fluorescence quenching does occur upon every encounter.

1.2.1.1 Triplet-Triplet Energy Transfer

A method particularly suited to obtain absolute rate constants of contact formation is triplet-triplet energy transfer (TTET).¹²⁴ In order to measure intrachain contact formation, triplet donor and acceptor groups are attached at specific sites of a polypeptide chain (Figure 1.3).

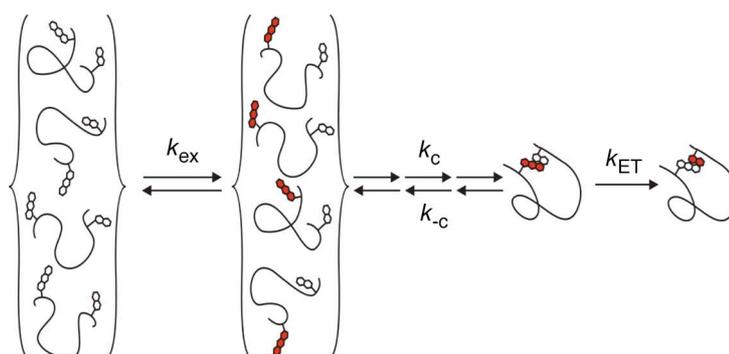


Figure 1.3. Scheme of triplet-triplet energy transfer (TTET) coupled to chain diffusion. Triplet donor and acceptor groups are attached at specific sites of a polypeptide chain. Excitation of the donor with a short laser pulse leads to a triplet state (red), which can be transferred to the acceptor upon van der Waals contact formation. Since k_{ET} is much faster than k_c and k_{c-} , TTET measurements directly give the absolute rate constant of contact formation k_c . Adapted from Krieger *et al.*¹⁰⁶

A triplet donor - acceptor pair well suited to measure TTET in aqueous solution is xanthone (Xan) and naphthalene (Nal). Xan is excited to its singlet state by a short laser pulse at 355 nm, leading to fast intersystem crossing to the triplet manifold (~ 2 ps)^{125,126} with a quantum yield of $\sim 99\%$ (Figure 1.4). Since the relaxation to the ground state is

spin forbidden, the Xan triplet state is relatively long-living in deoxygenated solutions ($\sim 30 \mu\text{s}$).

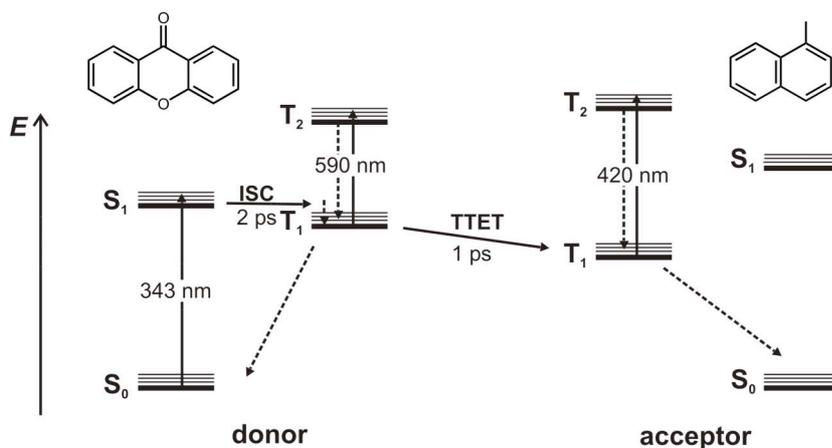


Figure 1.4. Jablonski diagram for TTET between xanthone (Xan) as donor and 1-methylnaphthalene (Nal) as acceptor. Xan is excited to its singlet state S₁ from where it undergoes fast intersystem crossing (ISC) to its triplet state T₁, which actually comprises two states in Xan ($^3n\pi^*$ and $^3\pi\pi^*$, not shown).^{125,126} Van der Waals contact with Nal leads to TTET. The triplet states can be monitored by the absorption bands arising from the T₁-T₂ transitions of Xan and Nal. Dashed arrows indicate internal conversion. The diagram reflects the relative energies.¹²⁷

If by chain diffusion the excited triplet donor forms contact with the triplet acceptor Nal (Figure 1.3) the excited state energy can be transferred by a simultaneous exchange of two electrons (Dexter mechanism).¹²⁸ Efficient energy transfer is ensured, since the Nal triplet state is lower in energy than the triplet state of Xan. The process can be directly followed by time-resolved absorption spectroscopy, since the triplet states of Xan and Nal give rise to characteristic absorption bands at 590 nm and 420 nm, respectively (Figure 1.4). The electron transfer mechanism requires orbital overlap between donor and acceptor, resulting in a strong, exponential distance dependence of the transfer rate constant k_{ET} (Eq. 7),

$$k_{\text{ET}} \propto e^{-2r/L} \quad (7)$$

where r is the distance between donor and acceptor and L the average of their van der Waals radii.¹²⁸ This means that TTET only occurs, when triplet donor and acceptor basically form van der Waals contact ($L \leq 5 \text{ \AA}$ for Xan and Nal). In comparison, FRET (singlet-singlet energy transfer) is mediated via dipolar interactions,¹²⁹ resulting in a weak distance dependence ($k_{\text{ET}} \propto r^{-6}$). Depending on the donor and acceptor pair, FRET is observed over distances up to $\sim 100 \text{ \AA}$.¹³⁰

The actual transfer (k_{ET}) between Xan and Nal in close contact occurs within 1 ps,¹²⁵ which is much faster than a typical encounter of the two labels lasts ($k_{ET} \gg k_c, k_{-c}$, Figure 1.3). Therefore the reaction is completely diffusion controlled and absolute rate constants of contact formation k_c can be obtained in a time-range from 10 ps to 30 μ s without any further assumptions.

1.2.1.2 Chain Dynamics in Unstructured Polypeptides

TTET measurements in poly(GlySer) and poly(Ser) peptides have provided detailed information on the chain dynamics of unstructured polypeptides.^{106,131-138} At ambient temperature contact formation can be described as single exponential process, in accordance with the theory of Szabo, Schulten and Schulten treating first passage times of diffusion controlled reactions.¹³⁹ Prerequisites for single exponential behaviour are that only a small fraction of molecules forms contact at a given time point and that the interconversion between all different conformations is fast compared to contact formation. In unstructured peptides contact formation occurs on the timescale of ten to hundreds of nanoseconds (Figure 1.5).

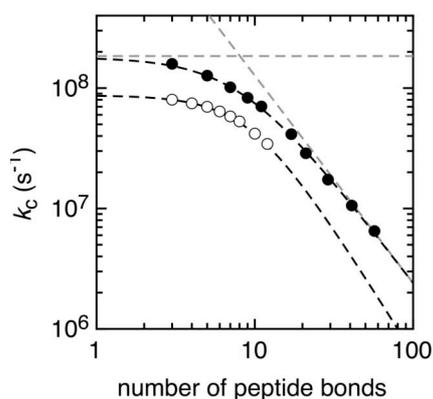


Figure 1.5. Rate constants of end-to-end contact formation in poly(GlySer) (●) and poly(Ser) (○) peptides of different length. The data were obtained at 22.5 °C using TTET between Xan and Nal (see 1.2.1.1). The lines show fits of Eq. 8. Data taken from Krieger *et al.*¹⁰⁶

The rate constant of end-to-end contact formation k_c was found to obey a scaling law (Eq. 8), as k_b scales with the number of peptide bonds N .

$$k_c = 1 / [1 / k_a + 1 / (k_b \cdot N^\gamma)] \quad (8)$$

The scaling factor $\gamma = -1.7$ determined for poly(GlySer)¹⁰⁶ agrees with polymer theory, which predicts that end-to-end contact formation scales with $N^{-1.5}$ in ideal Gaussian chains,^{139,140} with $N^{-1.8}$ if excluded volume effects are taken into account^{106,139} and with $N^{-1.7}$ to $N^{-2.0}$ in Rouse chains.¹⁴¹ In short peptides, which are strongly influenced by the chain stiffness and therefore do not longer behave like ideal chains, contact formation becomes length independent (Figure 1.5), expressed in the limiting rate constant k_a

(Eq. 8). Contact formation involving internal chain positions is additionally influenced by the size of the end-extensions, which can reduce k_c down to a limiting factor of 2.5 in the case of very large extensions.¹³⁸ Internal segments of a peptide chain are hence intrinsically less flexible than the chain ends.

Contact formation depends strongly on the inverse of the solvent viscosity (Eq. 9, $\beta \sim -1$),¹³⁵ as expected for diffusion-controlled reactions.

$$k_c = k_c^0 \cdot (\eta / \eta_0)^\beta \quad (9)$$

Fractional viscosities ($\beta > -1$) are observed for less flexible peptides or large viscogenic agents.¹³⁴ The effect of temperature on k_c in unstructured peptides is well described by Arrhenius behaviour with apparent activation energies of 20 - 35 kJ mol⁻¹.¹³⁴ Correction for the change of solvent viscosity with temperature (17 kJ mol⁻¹ for $\beta = -1$) reduces these activation energies to about 5 kJ mol⁻¹ for long poly(GlySer) peptides and to 15 - 20 kJ mol⁻¹ for less flexible peptides. Denaturants like urea and GdmCl affect $\ln k_c$ in an apparently linear fashion (Eq. 10).

$$\ln k_c = \ln k_c^0 - m \cdot [D] / RT \quad (10)$$

However, if k_c is corrected for the change of solvent viscosity (Eq. 9), the effect of denaturants is no longer linear, but can be ascribed to a binding isotherm.¹³⁵ Besides the dynamics of poly(GlySer) and poly(Ser) peptides the influence of different amino acid side chains has been addressed in host-guest studies.^{106,133,136} Glycine leads to significantly faster contact formation than all other amino acids, whereas proline gives double exponential contact formation kinetics. The faster phase could be assigned to peptides with a *cis* conformation of the amide bond preceding proline, causing a drastic restriction of the conformational space. Proline in the *trans* conformation leads to significantly slower contact formation.

Contact formation was also studied in some natural, but unstructured sequences, derived from carp parvalbumin or the GB1 hairpin.^{132,134} A recent study addressed fast contact formation events in short loops using femtosecond-laserflash spectroscopy and revealed a hierarchy of motions.¹³⁷ In addition to the diffusional processes on the 10 ns timescale, a subfraction of peptides showed complex contact formation kinetics on the timescale of 50 - 500 ps, probably representing motions within a local well of the free energy landscape. Only in a small fraction of molecules the labels had formed contact during the excitation.

1.2.2 Formation of α -Helices

The structure of α -helices was proposed by Pauling in 1950,¹⁴² who suggested them as structural elements in fibrous and globular proteins. Indeed, α -helices turned out to be the most abundant structural motif in proteins.¹⁴³ Since α -helices are stabilised by rather local interactions, they can also form independently of tertiary structure and isolated helices have been observed in peptides, protein fragments^{82,144,145} and several intermediates.^{102,103,146} The formation of α -helices can hence occur as one of the earliest steps in the folding of proteins.

Regular α -helices are defined by backbone H-bonds between the carbonyl groups of residues i and the amide protons of the residues $i + 4$ (Figure 1.6). H-bonds are the major stability determinant of α -helices,¹⁵ but the geometry is favoured for other reasons as well. The packing of the backbone is quite dense, thus adding stabilising van der Waals interactions. The dipolar peptide bonds are aligned in an energetically favourable parallel orientation, which results in a macro-dipole with a positive partial charge at the N-terminus and negative partial charge at the C-terminus. The side chains point outward in a tangential fashion, slightly directed towards the N-terminus. Their staggering of 100° avoids steric clashes, but allows interactions between side chains of residues i and $i + 4$.^{147,148}

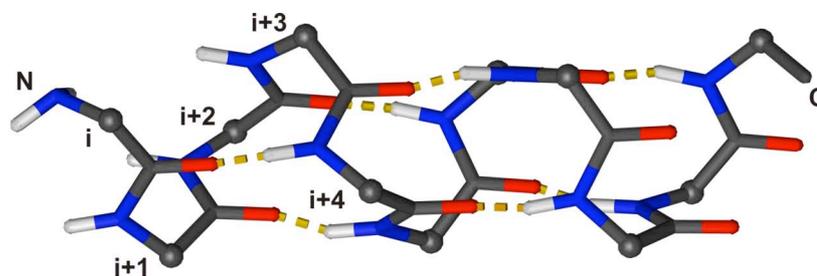


Figure 1.6. Model of a regular, right handed α -helix (3.6_{13} -helix). The helix has 3.6 residues per turn with a rise of 1.5 Å per residue. The backbone dihedral angles adopt ϕ, ψ -values of $(-57^\circ, -47^\circ)$. The $i, i + 4$ H-bonds (yellow) are 2.9 Å long. The C^α are symbolized as atoms, the side chains are omitted for clarity. The figure was prepared using MolMol.¹⁴⁹

All natural amino acids can be incorporated in regular α -helices, except proline that induces a kink ($\phi = -70^\circ$) and lacks the H-bond donor functionality. Less favourable are also glycine and β -branched amino acids.

A special situation is encountered at the N- and C-termini of helices, where four amide protons, respectively carbonyl groups, are not involved in backbone H-bonds. In proteins this is often compensated by capping motifs, including side chain - backbone

H-bonding.^{150,151} A structure related to the regular 3.6₁₃ α -helix is the 3₁₀-helix, which is characterised by a rise of three residues per turn and $i, i + 3$ H-bonds.¹ In general, this structure is energetically not as favourable as the α -helix, but less unsatisfied H-bond functionalities remain at the termini. As a consequence 3₁₀-helices are often found as N-terminal capping motifs in proteins and several studies suggest that they also prevail at the termini of isolated α -helices.¹⁵²

1.2.2.1 Thermodynamic Aspects of α -Helix Formation

The thermodynamic and kinetic properties of isolated α -helices are quite different from those of proteins. Due to the low cooperativity between the segments the helix-coil transition cannot be treated as two-state system but represents a multi-state system. Every single segment can be either in the helix or coil state, giving rise to 2^N different conformations for N segments. The restriction of the ϕ, ψ -angles to the helical space is associated with an entropy loss, which in the interior of helices is overcompensated by the enthalpy gain due to H-bonding. Helical segments close to helix boundaries remain energetically highly unfavourable, since they are only partly involved in $i, i + 4$ H-bonds (Figure 1.6). The state of a segment therefore depends to some extent on the conformation of the neighbouring residues, or more precisely, on the presence of H-bonding partners.

Statistical mechanical models resembling the cooperativity of the helix-coil transition have been devised early¹⁵³ and were applied to the helix-coil transition in polymers, such as poly- γ -benzyl-L-glutamate.¹⁵⁴ Zimm and Bragg attributed a statistical weight of $\sigma \cdot s$ to the first peptide bond of a helical structure with an H-bonded carbonyl group (nucleation) and a statistical weight of s (elongation) to every following peptide bond in a helix.¹⁵⁵ H-bonded carbonyls without preceding helical units were excluded and a statistical weight of *unity* assigned to all other peptide bonds. The formalism is usually simplified by considering only nearest neighbour interactions instead of $i, i + 4$ H-bonding partners for the assignment of the statistical weights (2 x 2 Zimm-Bragg matrix¹⁵⁵). A largely similar treatment¹⁵⁶ exclusively based on nearest neighbour interactions was given by Lifson and Roig.¹⁵⁷ Using matrix methods, the statistical weights assigned by Zimm and Bragg or Lifson and Roig can be used to calculate the partition function of any conformation, which means that all equilibrium properties of the system can be obtained.¹⁵⁸

While several polymer systems were known to form helices in different solvents,¹⁵⁴ α -helices in short peptides were considered for a long time as too unstable to exist in the

absence of tertiary interactions. The first contradicting example was the S-peptide of RNase S,⁸² which has a weak tendency to form helix in aqueous solution and was used as early model system to investigate the factors contributing to helix stability. Subsequently, the high helix forming potential of alanine was recognized,¹⁴⁷ which triggered the design and characterisation of many helix-forming peptides.^{148,151,159-167}

The overall helix content of peptides can be determined by CD spectroscopy and interpreted with the above-mentioned theories. By varying the peptide length the weights for nucleation ($\sigma \cdot s$) and elongation (s) in alanine-based helices were determined.¹⁶¹ The low value of $\sigma \approx 0.003$ mainly represents the entropy loss for fixing two consecutive pairs of ϕ, ψ -angles without formation of an H-bond, the value of $s \approx 1.35$ the enthalpy gain for adding another alanine (H-bond) to an existing helix. Thermal melting can be interpreted as a change of s and yields folding enthalpies of $\sim -4 \text{ kJ mol}^{-1}$ per residue, which is in agreement with enthalpies obtained model-free from calorimetric measurements.¹⁶⁸⁻¹⁷⁰ Other studies addressed the influence of denaturants^{148,165,167} or cosolvents¹⁴⁸ on helix stability. The influence of different side chains,¹⁶⁴ their interactions with each other¹⁴⁷ and their potential to act as capping motifs were investigated in a number of host-guest studies.^{151,166} The results from alanine-based guest-helices correlate with those from studies on copolymers¹⁷¹ and protein hosts.¹⁷² The data derived from short peptides can be used to obtain fairly good estimates for the tendency of a given sequence to form helices in the absence of tertiary interactions.^{151,173}

The helix-coil theories^{155,157} predict that these short peptides mainly contain one stretch of helical segments and that the ends of this helices are considerably frayed, i.e. the helix propensity is high in the helix centre where it reaches a plateau, but drops sharply towards the ends. Helix fraying was experimentally confirmed by measuring the effect of alanine to glycine substitutions at different positions,¹⁶² H/D-exchange studies,¹⁶³ as well as by electron spin resonance,¹⁷⁴ IR¹⁷⁵ and NMR spectroscopy.^{160,176,177}

1.2.2.2 Dynamics of α -Helices

Isolated α -helices are systems of low cooperativity, resulting in a highly dynamic equilibrium of many conformations. The addition and removal of single residues to an existing helix conformation is expected to be very fast, perhaps only limited by the conformational orientation. Successful nucleation events, in contrast, should occur orders of magnitudes slower, since nuclei below a critical size are unstable and will break down rapidly.

Data on helix dynamics are difficult to obtain. Opening and closing of single H-bonds can be probed by H/D-exchange, but the time resolution is much too slow to capture the fast dynamics and the expected equilibrium distribution is obtained.¹⁶³ Relaxation experiments¹⁷⁸ can provide the necessary time resolution, and dielectric¹⁷⁹ and ultrasonic¹⁸⁰ relaxation were applied early to study the kinetics of helical polymer systems. In principal, the 2^N possible conformations will give rise to $2^N - 1$ time constants, resulting in a continuous spectrum of relaxation times. The determined mean relaxation times are on the order of 20 ns up to 1 μ s.¹⁸⁰ The kinetic theory of Schwarz^{181,182} relates these maximum mean relaxation times to σ and to the rate constant k_+ for H-bond formation, yielding $k_+ \approx 10^{10} \text{ s}^{-1}$.¹⁸⁰

In the past decade relaxation studies were continued on small alanine-based helices using laser-induced nanosecond T -jumps, combined with fluorescence,^{183,184} UV resonance Raman¹⁸⁵ or IR^{121,186,187} spectroscopy. IR detection on ^{13}C -labelled carbonyl groups even allows site-specific measurements.^{121,188,189} The observed relaxation time constants are between 200 ns and 400 ns, with apparent activation energies of $\sim 34 - 50 \text{ kJ mol}^{-1}$ ^{185,187} and little variation for different positions.^{188,189} The molecular interpretation of these results remains however difficult, since the T -jumps of 10 to 15 K shift the multi-state equilibrium only slightly towards the unfolded state. In addition the apparent rate constants cannot be dissected into microscopic rate constants for folding or unfolding like in two-state systems. A NMR relaxation study detected dynamics on the microsecond timescale, but the microscopic origin of the exchange process remained unknown.¹⁹⁰ Initiation of a conformational transition by photo-induced *cis-trans* isomerisation of an azobenzene side chain cross-link led to complex dynamics extending to the microsecond timescale.¹⁹¹ An alternative way to probe the dynamics of helix-coil transitions is to couple a fast probing reaction to helix unfolding. Using triplet-quenching, Eaton and co-workers demonstrated that this approach can be used to monitor global helix dynamics.¹¹⁵

1.2.3 Formation of Hairpins and β -Sheets

β -Sheets often comprise distant chain regions or are even formed between different molecules, as in β -fibrils.¹⁹² A local motif, quite abundant in proteins, is the β -hairpin,¹⁹³ a pair of anti-parallel strand segments connected by a tight turn. Such hairpins appear to play an important role in the folding of several proteins, as they are often formed before the transition state is reached.^{68,194-196} Furthermore, β -hairpins have been observed in folding intermediates,¹⁹⁷⁻¹⁹⁹ unfolded proteins¹⁰¹ and protein fragments.^{200,201}

Several peptides forming monomeric β -hairpins in aqueous solution are known. Hairpin sequences derived from tendamistat^{83,202} and ferredoxin²⁰³ showed only a weak tendency for structure formation, whereas 16mers derived from the B1 domain of protein G²⁰⁰ and from ubiquitin adopt to some extent stable, native-like β -hairpin conformations.^{201,204} These studies stimulated the design of peptides with highly stabilised β -hairpins, like peptide I^{205,206} or the trpzip family.²⁰⁷ By combining two stable β -hairpin motifs small, three-stranded β -sheets can be built as well.

β -Hairpin peptides served as minimal model systems to investigate the stability determinants of hairpin and β -sheet formation,²⁰⁸ mainly using CD and NMR spectroscopy. In addition to intra-strand H-bonding three key factors for the stability of β -hairpins were identified. First, turn sequences that are compatible with the preferred right-handed twist of the strands are beneficial. Typical residues favouring such a geometry in two-residue turns (type I' and II' turns) are the non-natural D-proline²⁰⁹ or Xaa-Gly, where Xaa is often Asn or Asp.^{206,210,211} Second, the strands are predominantly composed of amino acids with bulky side chains like Val, Ile, Phe and Tyr, which have a high intrinsic propensity to populate the extended β -sheet region.²¹² Third, cross-strand interactions, often salt-bridges²¹³ or hydrophobic clusters,^{200,207,214} have a strongly stabilising effect. It is, however, difficult to obtain accurate data on the thermodynamic stability of hairpins, since the unfolding transitions are broad and the native baselines are unknown. Next to this, β -hairpins are not necessarily well approximated as two-state systems.

The folding and unfolding of β -hairpins occurs on the timescale of a few microseconds, as deduced from nanosecond *T*-jump measurements²¹⁵⁻²¹⁷ and NMR lineshape analysis,²¹⁸ i.e. their dynamics seem to be an order of magnitude slower than those of α -helices. In other studies hairpin folding was triggered by photo-induced cleavage of restraining cross-linkers²¹⁹ or photo-induced *trans-cis* isomerisation of an azobenzene turn,²²⁰ which gave folding time constants of 40 ns and 30 μ s, respectively.

1.3 The Fibrin Foldon Domain as a Model System for Oligomerisation

Many proteins exert their function as homo or hetero oligomers, especially structural proteins or systems that have to be well regulated. In principle, oligomerisation can occur after folding of the individual subunits has been completed or the two processes might occur in a concerted step.⁴⁹

An interesting system is the foldon domain of bacteriophage T4 fibritin.^{221,222} This C-terminally located domain of only 27 amino acid residues is essential for trimerisation of fibritin to the large segmental coiled-coil structure.²²³ Foldon can also be used to promote the trimerisation of other proteins, first demonstrated by fusion proteins with collagen²²⁴ and the HIV envelope protein gp140.²²⁵ Furthermore the foldon domain trimerises with high efficiency in isolation. The structure of the foldon trimer, determined by X-ray crystallography in the context of fibritin²²¹ and later by NMR spectroscopy in isolation,²²⁶ shows that each foldon subunit forms a twisted β -hairpin that are arranged to a trimeric β -propeller (Figure 1.7A). The trimer is stabilised by intermolecular backbone H-bonds (Figure 1.7B), a hydrophobic core around Trp20 and an intersubunit salt-bridge between Glu5 and Arg15'.

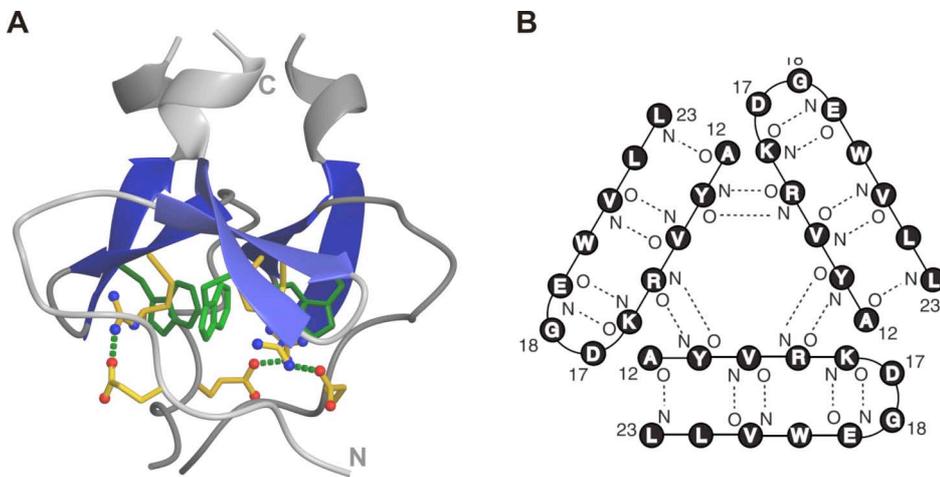


Figure 1.7: (A) Structure of the trimer formed by the foldon domain of bacteriophage T4 fibritin. The side chains of Trp20 and the intersubunit salt-bridge Glu5 - Arg15 are drawn. Each subunit consists of 27 amino acids with the sequence ¹GYIPEAPRDG¹¹QAYVRKDGEW²¹VLLSTFL (the hairpin strands are underlined). The model is based on PDB 1RFO²²⁶ and was prepared using MolMol.¹⁴⁹ (B) Scheme of the central β -hairpin regions showing the backbone H-bonds as dashed lines. Taken from Güthe *et al.*²²⁶

The folding/oligomerisation mechanism of foldon trimer was elucidated in a series of stopped-flow refolding experiments with different concentrations of foldon, in combination with interrupted refolding experiments.²²⁶ By this, the folding mechanism could be dissected into two bimolecular association steps, followed by a final rearrangement within the trimer. The association steps, with bimolecular rate constants of $2 \cdot 10^6 \text{ M}^{-1} \text{ s}^{-1}$ to the dimer and $5 \cdot 10^6 \text{ M}^{-1} \text{ s}^{-1}$ to the trimer, are close to the fastest known oligomerisation reaction of a natural protein, the dimerisation of wild type P22 Arc repressor.^{37,227}

A detailed amplitude analysis of the stopped-flow experiments revealed an intermediate on the monomer level of foldon, which forms within the millisecond dead-time of the experiment. This monomeric intermediate is likely of importance to ensure efficient oligomerisation. Its fluorescence spectrum differs from the spectra of the denatured and the native state and indicates that Trp20 is in a partially hydrophobic environment.²²⁶ The monomer is however not directly accessible, since in thermodynamic equilibrium it is only populated at very low concentrations (< 20 nM) or in the presence of denaturants that disrupt its structure.

The foldon trimer is significantly destabilised at pH 2, which makes it possible to investigate the acidic state (A-state) monomer by NMR methods.¹⁹⁹ Under these conditions the central β -hairpin is formed, remaining stable up to high temperatures, whereas the N- and C-terminal regions appear to be unstructured. Folding and unfolding of the hairpin motif seem to occur on the microsecond timescale.¹⁹⁹ The fluorescence spectra however indicate that the A-state is not identical to the monomeric folding intermediate, but significantly less structured.

1.4 The Villin Headpiece Subdomain as a Small Model Protein

The headpiece subdomain (HP35) of the villin from chicken consists of 35 amino acids.²²⁸ It is the smallest known protein unit folding in a cooperative manner in the absence of disulfide bonds or cofactors. High-resolution structures, determined by NMR²²⁹ and X-ray crystallography,²³⁰ show a three-helix bundle with a well defined hydrophobic core involving three phenylalanines (Figure 1.8).

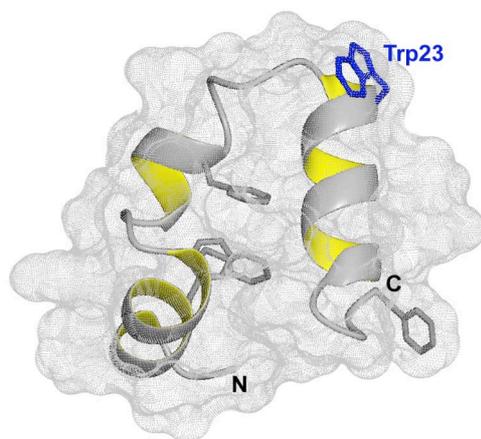


Figure 1.8. Ribbon and surface representation of the chicken villin headpiece subdomain (HP35). The complete sequence is ¹LSDEDFKAVF¹¹GMTRSAFANL
²¹PLWKQQNLKK³¹EKGLF, the helical regions are underlined. The figure is based on PDB 1YRF²³⁰ and was prepared using MolMol.¹⁴⁹

The thermodynamic stability of HP35 is low ($\Delta G^\circ = 8 - 12 \text{ kJ mol}^{-1}$), but no deviations from two-state behaviour have been found in equilibrium unfolding experiments.

The high protection of the most amide protons against H/D-exchange confirms the formation of a stable and well-defined structure.²²⁸ The two C-terminal amino acids, Leu34 and Phe35, seem to be flexible in HP35 and only ordered within the full headpiece domain. Despite the low overall stability and the high sequence homology to headpiece domains from other species, a number of reported amino acid substitutions had little effect on the structure and stability of HP35.^{47,231-233}

HP35 is one of the fastest folding proteins. The relaxation kinetics, investigated in several *T*-jump experiments utilizing different probes, were in all cases found to be double exponential.^{119,230,232} The slower, microsecond phase has been attributed to global unfolding/folding, the faster phase (~100 ns) to either local unfolding, helix-coil transitions in the unfolded state or changes in solvation. Assuming two-state behaviour to obtain folding and unfolding time constants suggests that folding occurs on the timescale of a few microseconds independent of temperature, whereas unfolding has to be much slower at room temperature (milliseconds). In *T*-jump experiments the shift of the equilibrium towards the unfolded state is however small and it cannot be proven that global folding and unfolding is observed in the relaxation experiments. NMR lineshape analysis confirmed that an exchange process occurs within microseconds²³⁴ and data from a tryptophan triplet quenching experiment¹¹⁹ agree with this observation as well. Several amino acid substitutions had only minor effects on the folding kinetics, but in the case of a designed variant the time constant for folding could be increased to 700 ns,⁴⁷ the fastest folding time observed for a protein so far.^{46,48}

A possible explanation for the unusual fast folding of HP35 would be that the unfolded state retains residual structure or a native-like topology.²³⁵ In fact, an isolated fragment consisting of helix 1 and helix 2 forms considerable amount of structure,^{86,236} whereas the isolated helical segments have low intrinsic helix propensities.

Its small size, simple topology and fast folding made HP35 to a major subject of computational studies and molecular dynamic simulations on protein folding, as it may allow closing the gap between experimental and computational approaches. Numerous molecular dynamic simulations, including the first 1 μ s all-atom simulation of Kollman²³⁷ and the distributed computing project of Pande,²³⁵ addressed the properties of the native and denatured state ensembles, as well as the course of folding. Using replica exchange molecular dynamics, Duan and colleagues recently achieved complete folding to the native state.²³⁸

2 Aims of Research

Folding of proteins involves the formation of local secondary structures and the formation of contacts between distant sites along a polypeptide chain. Dynamics of loop formation in unstructured model peptides and protein fragments have been characterised in detail using triplet-triplet energy transfer (TTET). The aim of this work was to get insight into the dynamic properties of secondary structures and into the dynamics of unfolded and partially folded proteins.

1. Dynamics of Secondary Structures

α -Helices and β -hairpins can form in the absence of tertiary interactions. The thermodynamics of these structural elements have been extensively characterised. Much less is known about their dynamics, since classical relaxation measurements, giving access to the kinetics on the nanosecond timescale, are difficult to interpret in the case of multi-state systems.

An alternative approach to study conformational transitions is to couple a fast probing reaction to one of the states, which allows to characterise the dynamics of the transition in equilibrium. Contact formation between a triplet donor and an acceptor group fulfils the requirements for such a probing reaction. If contact formation between the labels is prevented in the folded state but can occur in the unfolded state, this method should be well suited to characterise local folding and unfolding events in α -helices and β -hairpins.

We wanted to investigate the local dynamics in a series of alanine-based model helices by attaching the labels at residues i and $i + 6$, which places them on opposite sides of the helix. Basic experiments should establish, whether this approach can be used to obtain rate constants of local helix folding and unfolding. We further wanted to investigate the local dynamics at different helix positions and to study the temperature and length dependence of the folding and unfolding rate constants. The comparison of the experimental results with simulations based on kinetic models for the helix-coil transition should provide new insight into the dynamics of α -helical peptides and the underlying processes.

The same approach should be applicable to study the dynamics of β -hairpin formation. The β -hairpin region of the foldon domain from bacteriophage T4 fibritin was chosen as a model system for these studies. For this purpose, we planned to synthesise peptides comprising the hairpin sequence and to investigate their structure and stability. By

attaching triplet donor and acceptor groups at defined positions in the hairpin we intended to study its equilibrium dynamics in order to obtain rate constants for hairpin folding and unfolding.

Besides the equilibrium dynamics of secondary structures it would be desirable to follow structure formation starting from a fully unfolded state. Substantial helix folding or unfolding reactions might be induced using photoswitchable groups. Thioamide bonds, which can be incorporated as nearly isosteric backbone modifications, were shown to undergo a fast and reversible *cis-trans* isomerisation upon irradiation with light. We set out to investigate the effect of single amide-to-thioamide substitutions on the structure and stability of alanine-based model helices, since no conclusive data on this issue were available.

2. Structure and Dynamics of a Folding Intermediate

The dynamics and three-dimensional structure of folding intermediates are still poorly understood. Foldon, a small domain promoting the trimerisation of fibrin from bacteriophage T4, rapidly forms a monomeric intermediate during the folding process. This intermediate associates fast and with high efficiency to the dimer and subsequently to the trimer. The transient nature of the intermediate does not allow a detailed characterisation, but fluorescence and circular dichroism spectra point to a considerable amount of structure formation. The trimer dissociates at low pH and it was shown by NMR spectroscopy that the native β -hairpin is still formed in the acidic-state. This state does however not represent the folding intermediate, since its tryptophan fluorescence differs substantially from the spectrum of the folding intermediate. In order to obtain a better model of the folding intermediate for equilibrium studies, we wanted to disrupt an intersubunit salt-bridge between Glu5 and Arg15', which is essential for trimerisation. Investigation of a foldon variant with a single amino acid replacement (E5R) by NMR and optical spectroscopy should enable us to characterise the structure and dynamics of this key folding intermediate under physiological conditions. This should allow us to elucidate the structural basis for the rapid and efficient assembly reaction. Further experiments aim at the use of TTET to study the dynamics in the intermediate and to compare the results to the dynamics of the isolated hairpin.

3. Dynamics in the Unfolded and Native State of a Small Protein

The dynamics in the unfolded state of proteins remain largely unexplored, although they certainly affect the folding process and might be influenced by local structure or a collapse of the polypeptide chain. TTET measurements between different donor and acceptor positions should for the first time provide absolute rate constants of contact formation in a small protein, the villin headpiece subdomain (HP35). Besides loop formation in the denatured state, we wanted to investigate the dynamics in unfolded molecules under conditions strongly favouring the native state. This should contribute to understand the fast folding of HP35, which became a subject of many experimental and theoretical studies. We also planned to perform TTET measurements in the native state of HP35 in order to obtain information on structural fluctuations and local flexibility.

The extension of TTET measurements to probe the dynamics in larger proteins is highly desirable, but requires novel techniques to incorporate the triplet donor and acceptor labels. We therefore wanted to explore strategies to introduce the triplet labels into proteins using cysteine-reactive derivatives of the labels.

3 Summary of Work Ready for Submission

3.1 Dynamics of α -Helical Peptides Probed by Triplet-Triplet Energy Transfer

α -Helices can form in the absence of tertiary interactions and hence play likely an important role in the early stages of protein folding. Isolated α -helical structures are only marginally stable, weakly cooperative and highly dynamic. Peptides mainly consisting of alanine are well suited to study the formation of these fundamental structural elements, since they form relatively stable helices and are devoid of complex side chain interactions. Their thermodynamics have been intensively characterised and can be well described by the helix-coil theories originally developed for long homopolymers. The kinetics of helix formation in peptides have been investigated as well, mainly using T -jump techniques. In the case of α -helices, the interpretation of relaxation kinetics is however difficult, since their cooperativity is low, making them to multi-state rather than two-state systems. Another problem is associated with the measured probes, which often report on global changes and are therefore not directly related to microscopic events. For these reasons a different approach was used, studying the dynamics of α -helical peptides in thermodynamic equilibrium. We coupled a fast and irreversible probing reaction to the helix-coil transition in order to determine the rate constants of helix folding and unfolding (Figure 3.1).

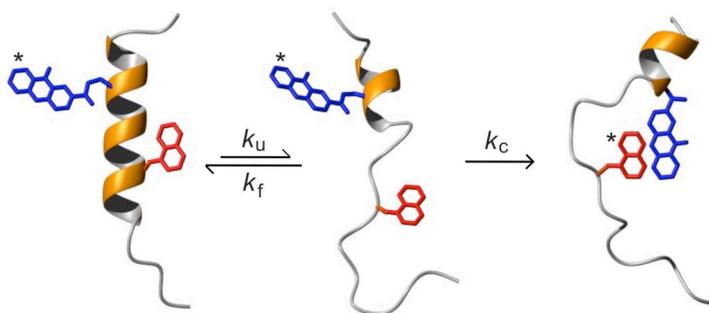


Figure 3.1. Probing local helix unfolding k_u and folding k_f by contact formation between $i, i + 6$ spaced triplet donor (Xan, blue) and acceptor (Nal, red) groups. A short laser pulse is used to generate a Xan triplet state (*), which can be monitored by its absorption at 590 nm. Upon local unfolding (k_u) the labels are free to form van der Waals contact with the rate constant k_c , as long the helix is not reformed (k_f). Contact formation leads to an instantaneous transfer of the triplet state from Xan to Nal.

As probing reaction contact formation between a triplet donor group (Xan) and an acceptor group (Nal) was chosen, which can be directly measured by TTET. If the labels are introduced at the side chains of residues i and $i + 6$ they are located on opposite sides of a folded helix. This is sufficient to prohibit contact formation between the labels. Upon local helix unfolding the chain becomes flexible and allows the labels to form contact, which leads to TTET. In the experiments with alanine-based model helices, we observed double exponential TTET kinetics as expected for the three-state system depicted in Figure 3.1. The apparent rate constants in combination with the amplitude information were used to extract the microscopic rate constants k_u , k_f and k_c in a global fitting procedure. The probing reaction was in all cases found to occur in the range of $1/k_c \approx 100$ ns.

After establishing the principle of these measurements we addressed a number of different aspects. First we investigated the dynamics at five different positions within a 21 residue peptide. Local helix folding was found to be position independent with time constants of $1/k_f \approx 400$ ns. In contrast, unfolding varied from $1/k_u \approx 1.4$ μ s in the helix centre to $1/k_u \approx 250$ ns at the termini (at 5 °C). This fraying of the helices towards the ends is in agreement with a number of equilibrium experiments and predictions from helix-coil theories. The experiments point to a slight asymmetry of our helices, since the C-terminus appears to be somewhat less stable than the N-terminus. Second, we systematically investigated the effect of urea on the local helix dynamics, which also allowed us to test our model over a wide range of conditions. We found that the addition of urea slows down folding at all helix positions in a similar way but has less pronounced effects on helix unfolding. In the helix centre urea slightly accelerates helix unfolding, at the termini it decelerates unfolding.

In order to gain mechanistic insight into the complex behaviour of this multi-state system we performed discrete time-step Monte-Carlo simulations on a kinetic Ising model, resembling the cooperativity of the α -helical peptides. The simulations demonstrate that the observed kinetics mainly result from addition and removal of helical segments to existing helix boundaries, whereas nucleation events seem to be of minor importance under these conditions. A quantitative comparison suggests that the addition of single segments to an existing helix (helix elongation) occurs on the timescale of about 50 ns.

In a second work we investigated the effect of temperature on the local dynamics in the helix centre. Local unfolding encounters a high enthalpic barrier, whereas folding is only weakly temperature dependent and shows a pronounced convex curvature in the Arrhenius plot. A comparison with predictions from our kinetic Ising model suggests

however, that this behaviour has to originate from high activation energies of both elementary processes, helix elongation and unwinding, of around 40 kJ mol^{-1} .

The experiments and results are described in detail in the following manuscripts (see 10.A and 10.B). Data on the dynamics in helices of different length are summarised as unpublished results (see 4.1, p. 36).

Fierz, B., # Reiner, A. # & Kiefhaber, T. (2007). Local Conformational Dynamics in α -Helices Measured by Fast Triplet Transfer. *To be submitted*.

Reiner, A., # Fierz, B. # & Kiefhaber, T. (2007). Temperature Dependence of Local Equilibrium Folding and Unfolding of an α -Helix. *To be submitted*.

equally contributing

3.2 Thioamide Backbone Substitutions in α -Helices

Thioamides ([CS-NH]) possess altered hydrogen-bonding and photophysical properties compared to amides, making them to promising backbone modifications for biophysical studies of peptides and proteins. A particularly interesting aspect is, that both, fast *trans*→*cis* and *cis*→*trans* isomerisation can be induced by irradiation with light of suitable wavelengths.^{239,240} On the other hand, the rotational barrier in the ground state is high and hence relaxation back to the thermodynamic equilibrium is slow. Thioamides are nearly isosteric to amide bonds and have therefore a high potential as non-invasive photoswitchable backbone modification that might allow to induce and to study fast conformational transitions.

The compatibility of thioamides with secondary structures has been addressed in several theoretical and experimental studies, but the effect of thioamide substitutions on α -helices remains unclear. For single amide-to-thioamide substitutions in an α -helical coiled-coil no destabilisation has been observed,²⁴¹ whereas theoretical studies suggest that thioamides should have a destabilising effect or even might be incompatible with α -helical structures. In this context, it was also proposed that thioamides might favour 3_{10} -helical conformations. Experimental data on thioamide substitutions in α -helical peptides have not been reported so far.

We investigated the effect of single amide-to-thioamide substitutions in alanine-based model helices, using CD and NMR spectroscopy. A single thioamide substitution in the centre of a 16 residue sequence leads to a significant decrease of the helix content, from around 68 % in the reference peptide to around 24 % in the thioamide peptide. This destabilisation is comparable to the effect of an alanine-to-glycine substitution at the same position. Using helix-coil theory, it can be estimated that the destabilisation caused by the thioamide corresponds to approximately 7 kJ mol⁻¹.

A thioamide substitution close to the N-terminus has a destabilising effect of similar strength as revealed by a second thioamide peptide. The overall helix content of this peptide is, however, less severely affected, due to the fraying of helical structures at their ends, rendering the termini less sensitive towards perturbations. The NMR data show that the thioamide bond between residues two and three is unequivocally involved as acceptor in an *i, i + 4* H-bond. This indicates that despite the strong destabilising effect helical conformations can be formed around thioamides and argues against the formation of 3_{10} -helical structures, which are defined by *i, i + 3* H-bonds. It had been suggested that 3_{10} -helices might generally prevail at the N-termini of α -helical peptides, as less

unsaturated H-bond donor functionalities remain. Thioamides should favour this conformation even more, since the somewhat widened conformation of 3_{10} -helices compared to canonical α -helices might better accommodate the thiocarbonyl group.

The data on the two thioamide-substituted peptides show that these substitutions have a strong destabilising effect on regular α -helices. This was anticipated because the longer C=S bond and the larger sulphur encounter a quite densely packed structure. Additional effects likely arise from the altered H-bonding properties of thioamides, which are stronger H-bond donors but weaker acceptors. In an earlier experimental study on thioamide substitutions in a coiled-coil, widened positions of the somewhat irregular structure were chosen to place the thioamide groups. This was apparently sufficient to avoid steric clashes and led to the general notion that helices tolerate these substitutions quite well. The destabilising effect of thioamides limits their use to induce conformational transitions in regular alanine-based α -helices. Nevertheless, thioamides can be incorporated into α -helical structures and might be used as perturbations in folding experiments for determining backbone ϕ -values.

The experiments and results are described and discussed in detail in the manuscript (see 10.C):

Reiner, A., Wildemann, D., Fischer, G. & Kiefhaber, T. (2007). Effect of Thiopeptide Bonds on α -Helix Structure and Stability. *To be submitted*.

3.3 Structure and Dynamics of the Foldon Monomer

The foldon domain, located at the C-terminus of bacteriophage T4 fibritin, promotes the trimerisation of this large rod-like protein. The foldon domain forms a stable trimer in isolation and its folding/oligomerisation mechanism has been elucidated in a number of kinetic experiments.

The monomeric state of foldon seems to be optimised for efficient association into the dimer and trimer, as high bimolecular rate constants of $2 - 5 \cdot 10^6 \text{ M}^{-1} \text{ s}^{-1}$ have been found for these steps. The fluorescence and CD spectra of the monomeric folding intermediate, obtained in stopped-flow experiments, point to a considerable amount of structure formation in the monomer.²²⁶ Furthermore, it was known that the foldon trimer dissociates at low pH and that the central β -hairpin is formed in the acidic-state (A-state) monomer, which had been characterised by NMR spectroscopy. However, the CD and fluorescence spectra indicate that the A-state is not identical, but less structured than the kinetic folding intermediate under physiological conditions.

In order to get a more adequate model of the monomeric folding intermediate for equilibrium studies, we disrupted the intersubunit salt-bridge between Glu and Arg15' by a single amino acid replacement (E5R). This exchange should have minor effects on the monomeric state, as Glu5 does not form any pronounced intrasubunit interactions. The foldon variant E5R was prepared by solid-phase peptide synthesis and later by expression as thioredoxin-fusion protein in *Escherichia coli* to obtain ^{15}N -labelled protein for NMR investigations.

As expected from previous pH-titration experiments,¹⁹⁹ the single amino acid replacement E5R destabilises the trimer sufficiently to study the monomeric state in concentrations up to $\sim 200 \mu\text{M}$. The blue-shifted fluorescence emission of Trp20 indicates that its side chain is embedded in a hydrophobic environment and the fluorescence quantum yield in foldon E5R is much higher than in the previously characterised A-state. Indeed, the fluorescence spectrum of foldon E5R closely resembles the spectrum of the kinetic folding intermediate.²²⁶ Further evidence for the structural similarity between foldon E5R and the monomeric intermediate is given by the corresponding CD spectra. The foldon variant E5R can hence be used as a model to study the monomeric state of foldon, which associates so efficiently.

Structure and dynamics of the foldon monomer E5R were investigated by NMR methods. Similar to the native trimer²²⁶ and the previously characterised A-state,¹⁹⁹ a central β -hairpin is formed in the region from Ala12 to Leu23. However, in foldon E5R

additional structure exists compared to the A-state, as a number of defined interactions between the hairpin and the N-terminal part of the protein were detected. The side chains of Pro4, Ala6 and Pro7 form a small hydrophobic cluster with Tyr13 and Trp20 from the hairpin region. The distance restraints obtained by NMR allowed to calculate the structure of the foldon monomer E5R. The ensemble of structures shows that the hydrophobic interactions between the N-terminal region and the central β -hairpin give rise to a compact topology, quite similar to the structure adopted by the individual subunits in the trimer. The N-terminal region, however, appears not to be fully structured as indicated by sparse contacts between the residues. Some tendency to form native-like conformations is observed for the four C-terminal residues, which form a short 3_{10} -helix in the native trimer.

The thermodynamic stability of the structure in the foldon monomer appears to be low. Large temperature coefficients of Ala6 and Pro7 side chain protons, which are strongly shifted due to ring current effects of Trp20 and Tyr13, point to a steady loss of structure upon heating. Large negative temperature coefficients of the amide protons of Val14 and Val21, involved in cross-strand H-bonds, indicate that also the β -hairpin gets significantly unfolded at higher temperatures.

The dynamics of the backbone were probed by the ^{15}N -relaxation behaviour. The longitudinal relaxation rate constants (R_1) show, besides end effects, uniform motions of the polypeptide chain in the picosecond to nanosecond time range. Several transverse relaxation rate constants (R_2) in the hairpin region were significantly increased, indicating exchange processes on the tens of microsecond timescale, which probably arise from hairpin folding and unfolding.

In summary, the foldon variant E5R is a model that represents the monomeric folding intermediate under physiological conditions. The structural preorganisation of the monomer, quite similar to the structure of the subunits in the trimer, explains the efficient association to the dimer and trimer. The stability of this structure appears to be low. A too stable monomer, on the other hand, would likely reduce the free energy gain upon trimerisation.

The experiments and results on the monomeric foldon variant E5R are described in detail in the following manuscript (see 10.D):

Habazettl, J., Reiner, A. & Kiefhaber, T. (2007). NMR Structure of a Monomeric Intermediate on the Evolutionarily Optimized Assembly Pathway of the Trimeric Foldon Domain. *To be submitted*.

3.4 Contact Formation Dynamics in the Villin Headpiece Subdomain

Despite continuous progress, mainly in NMR techniques, several timescales and types of protein motions remain unexplored. Especially the dynamics in the unfolded state ensemble, involving large amplitude motions of the polypeptide chain on the nanosecond timescale are difficult to capture. These dynamics, however, are closely related to the conformational search towards the native state and determine how fast interactions between different sites of a polypeptide chain can form.

The villin headpiece subdomain (HP35) was chosen as model protein to investigate for the first time absolute contact formation dynamics in the native and denatured state of a protein. Despite its small size of 35 amino acids it shows all properties characteristic for proteins. HP35 folds on the timescale of a few microseconds and thus became a major subject of experimental and computational studies. A fragment of HP35 has the tendency to form native-like structure, which led to the suggestion that the fast folding might result from residual structure in the unfolded state.²³⁶ From molecular dynamic (MD) simulations it was proposed that the unfolded state may have a native-like topology.²³⁵

In order to investigate the contact formation dynamics in HP35, we chose four different pairs of triplet donor and acceptor positions for TTET measurements. We prepared the double- and donor-only labelled HP35 variants by solid-phase peptide synthesis and introduced the labels at positions, where little effect on the structure of HP35 was expected. Characteristic one-dimensional NMR spectra indicated that the double-labelled variants folded into the native structure and the thermodynamic stability, assessed by CD measurements, was comparable to wild type HP35.

In two variants we had placed donor and acceptor at distant sites of the structure. In these variants no TETT was observed during the lifetime of the Xan triplet state of around $\sim 10 \mu\text{s}$ under the experimental conditions. This indicates that the labels are well separated, and furthermore, that global unfolding has to be significantly slower than $10 \mu\text{s}$ under these conditions. In another variant the labels were attached at the N- and C-termini, which are in close vicinity in the native state. As expected, this led to fast contact formation in the native state, but on two distinct timescales. In 50 % of the native molecules contact formation occurred in a sub-nanosecond process during the dead-time of the experiment, i.e. contact between the labels is formed or can be established by a few bond rotations. In the remaining molecules the labels formed contact with a time constant of $\sim 30 \text{ ns}$, indicating that the labels have a larger degree of freedom or that a small energy barrier opposes contact formation.

In the fourth variant the labels were placed along the C-terminal helix to probe the local dynamics of this region. Although the labels are separated by two and a half helical turns, efficient contact formation between the labels was observed. The contact formation kinetics are double exponential with time constants of 170 ns and 1 μ s, the faster being strongly sensitive towards denaturants. This points to local unfolding processes of the C-terminal helix to a much higher extent than the previously described flexibility of the last two amino acid residues.

The addition of denaturants leads to unfolding of the villin headpiece subdomain, with a midpoint of denaturation around 2.5 M GdmCl. Upon unfolding, we detected an additional TTET process in all variants, corresponding to contact formation in the denatured state. At high concentrations of denaturant (above 4 M GdmCl) HP35 is completely unfolded and only this contact formation process remains. In all cases, the unfolding transition probed by contact formation is in full agreement with the unfolding transition measured by CD, which is a hallmark of two-state behaviour and confirms the assignment of the observed phases.

Loop formation in the fraction of denatured molecules takes place on the timescale of 100 to 200 ns. These absolute rate constants of contact formation obtained by TTET are 20-fold faster than previous estimates from a triplet quenching experiment.¹¹⁹ This discrepancy is due to the fact that quenching of tryptophan by cysteine is not diffusion-controlled, i.e. not every encounter leads to quenching. The conformational equilibration of the polypeptide has to be much faster than 100 ns, since single exponential kinetics were observed in all cases. The loop formation dynamics, however, are already close to the estimated folding time constants of up to 700 ns.⁴⁷ Comparison with data from unstructured model peptides suggests that the chain dynamics in the denatured state of HP35 are not affected by residual structure, even under conditions strongly favouring the native state. Contact formation is neither slowed down by an unspecific collapse of the chain, nor accelerated between side chains forming contact in the native state.

In summary, TTET measurements between different donor and acceptor positions revealed heterogeneity in the native state of HP35 and gave an accurate description of the chain dynamics in the denatured state. The absolute rate constants of contact formation should be of great use for direct comparison with MD simulations. Even under conditions strongly favouring the native state no evidence for residual structure was found.

The experiments and results are described in detail in the manuscript (see 10.E):

Reiner, A., Henklein, P. & Kiefhaber, T. (2007). Conformational Dynamics in the Native and Denatured State of the Villin Headpiece Subdomain Measured by Triplet-Triplet Energy Transfer. *To be submitted.*

4 Summary of Unpublished Results

4.1 Dynamics in α -Helical Peptides of Different Length

The experiments on alanine-based model helices have shown that contact formation between $i, i + 6$ spaced triplet donor (Xan) and acceptor (Nal) groups is well suited to measure local helix folding and unfolding (see 3.1). Previous studies addressed the dynamics at different positions within the helices, as well as the influence of denaturants (see 10.A) and temperature (see 10.B). This section summarises the effect of helix length on local folding and unfolding dynamics.

The influence of the peptide length on the dynamics of α -helices is of general interest, since the size of the system is a fundamental parameter in all descriptions of the helix-coil transition.^{153,155-157} Length-dependent relaxation data were difficult to obtain with homopolymeric model systems and only one study on alanine-based model helices of different length has been reported.¹⁸⁷ Experimental details on the synthesis and characterisation of the peptides, on the TTET measurements and data evaluation and on the kinetic Ising model are given in the publication 10.A and the manuscript 10.B.

4.1.1 Design and Helix Content of the Peptides

Alanine-based model helices, similar to those used in the previous studies (see 10.A and 10.B) were chosen to investigate the influence of the peptide length on the dynamics and stability of α -helices. The synthesised peptides comprise eleven to 41 amino acid residues (Table 4.1), consist mainly of alanine¹⁴⁸ and carry an arginine in every fifth position in order to mediate solubility and to avoid oligomerisation.¹⁵¹ The termini were capped to prevent unfavourable interactions with the helix dipole. The shortest length of the peptides is given by the critical length necessary to form a significant amount of helix (see below), the upper length by practical limitations in Fmoc solid-phase peptide synthesis. The triplet donor group Xan was attached as carboxylic acid derivative via an amide bond at the side chain of α, β -diaminopropionic acid. The triplet acceptor naphthalene was incorporated as 1-naphthylalanine.

The labels were introduced at position i (Xan) and $i + 6$ (Nal), placing them on opposite sides of a helix, separated by one and a half helical turns. This is sufficient to prohibit contact formation between donor and acceptor as long as helical structure is formed in that region (see 10.A). In a series of peptides with 11, 13, 16, 21, 31 and 41 residues the

$i, i + 6$ spaced labels were placed in the centre (**11c**, **13c**, **16c**, **21c**, **31c**, and **41c**; Table 4.1). In three additional peptides with 16, 21, and 31 residues the labels were placed at the N-terminus (**16n**, **21n** and **31n**).

Table 4.1. Alanine-Based Model Helices of Different Length

Name	<i>N</i>	Label pos.	Sequence
N-terminal ($i, i + 6$) probing			
16n	16	X1-Z7	Ac- X AAAA AZARA AAARA A-NH ₂
21n	21	X1-Z7	Ac- X AAAA AZARA AAARA AAARA A-NH ₂
31n	31	X1-Z7	Ac- X AAAA AZARA AAARA AAARA AAARA AAARA A-NH ₂
Central ($i, i + 6$) probing			
11c	11	X3-Z9	Ac-AA X AA RAA Z A A-NH ₂
13c	13	X4-Z10	Ac-AA X A RAA Z RAA-NH ₂
16c	16	X5-Z11	Ac-AAAA X AAARA Z AARA A-NH ₂
21c	21	X7-Z13	Ac-AAAAA A XARA AA Z RA AAARA A-NH ₂
31c	31	X12-Z18	Ac-AAAAA AAARA A XARA AA Z RA AAARA AAARA A-NH ₂
41c	41	X17-Z23	Ac-AAAAA AAARA AAARA A XARA AA Z RA AAARA AAARA AAARA A-NH ₂
Donor-only reference			
21d	21	X7	Ac-AAAAA A XARA AAARA AAARA A-NH ₂

N gives the number of residues. **X** denotes the triplet donor Xan, attached to the β -amino group of selectively deprotected α, β -diaminopropionic acid; **Z** denotes the triplet acceptor Nal, incorporated as 1-naphthylalanine.

The shortest peptide **11c** was found to be prone to aggregation in the absence of denaturant. The Far-UV CD spectra of all other peptides are shown in Figure 4.1A. All but peptide **13c** form a significant amount of helical structure, indicated by maxima of the ellipticity at 190 nm and minima around 208 nm and 222 nm. The triplet labels Xan and Nal strongly absorb in the Far-UV and contribute significantly to the CD spectra, impeding a quantitative assessment of the helix content. The qualitative comparison of the spectra, however, clearly shows two trends, which are also expected from helix-coil theory.^{155,157} First, at a critical length, which is apparently between 13 and 16 residues, a sharp increase of the helix content is observed. Then, up to the longest peptide of 41 residues, the helix content increases more steadily. The incorporation of the $i, i + 6$ spaced triplet labels has a destabilising effect (see 10.A), especially in the helix centre, which is most sensitive towards perturbations.¹⁶² This explains the higher helix content observed for **21d** / **21n** compared to **21c** and of **31n** compared to **31c**. In contrast, **16n** appears to have a lower helix content than **16c**. This might be a false estimate due to different

contributions of the triplet labels in more helical regions (**16c**) and less helical regions (**16n**) that become more dominant in shorter peptides.

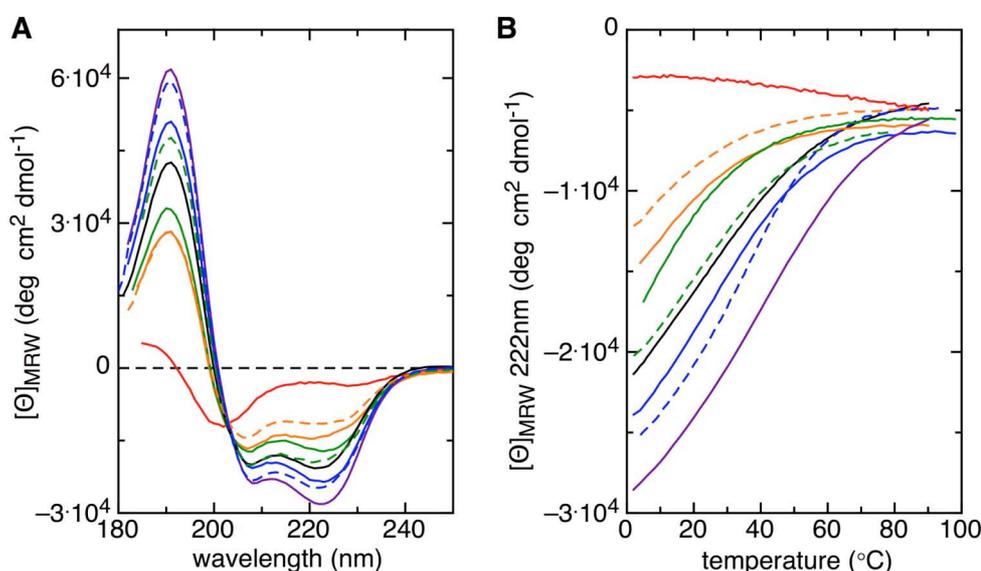


Figure 4.1. (A) Far-UV CD spectra of the peptides at 5 °C. The solid lines show spectra of **13c** (—), **16c** (—), **21c** (—), **31c** (—), **41c** (—) and **21d** (—), the broken lines of **16n** (---), **21n** (---) and **31n** (---). (B) Thermal melting of the peptides monitored at 222 nm, lines and colours as in (A). The peptide sequences are given in Table 4.1. The data were recorded in 10 mM potassium phosphate, pH 7.0.

The thermal unfolding transitions monitored by the ellipticity change at 222 nm confirm an increase of the helix stability with peptide length (Figure 4.1B). All peptides with an α -helical spectrum show signal changes that are indicative for the loss of helical structure. In contrast, the CD signal of **13c** at 222 nm decreases in an almost linear fashion, which is typical for unstructured alanine peptides.²⁴² Hence, no helical structure seems to be formed in **13c** at all, whereas the signal change observed for **16c** clearly confirms the presence of helix. A similar 12mer without destabilising labels and with lysine residues instead of arginine has been reported to have a helix content of $\sim 20\%$.²⁴² With increasing peptide length the cooperativity of the transitions increases as expected from helix-coil theories.^{154,155,157,161}

4.1.2 Local Dynamics in the Helix Centre Measured by TTET

The dynamics in the helical peptides were investigated using contact formation as probing reaction for local folding and unfolding of the segment between the i , $i + 6$ spaced labels (see 3.1). As long as helical structure between the two labels is formed, they point on opposite sides and cannot form contact. Upon local unfolding (k_u) the labels are free to

form contact with a time constant k_c , as long as the helical structure is not reformed (k_f) (Eq. 11).



Contact formation between Xan and Nal leads to TTET, which can be directly followed by the decay of the Xan triplet absorption band at 590 nm (Figure 4.2) and a concomitant increase of the Nal absorption band at 420 nm (data not shown). Under the experimental conditions the Xan triplet state has an intrinsic lifetime of $\sim 35 \mu\text{s}$, as observed in the donor-only helix **21d** (Figure 4.2). The intrinsic Xan triplet lifetime is independent of the α -helical environment, since it does not change if the helix is denatured by the addition of urea (data not shown). The Xan triplet decays in all double-labelled peptides are much faster than in the donor-only peptide, indicating that significant TTET occurs (Figure 4.2).

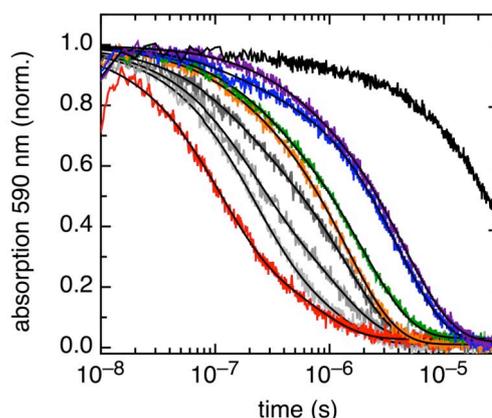


Figure 4.2. TTET measurements in peptides of different length (see Table 4.1). Shown are the Xan triplet absorption decays at 590 nm of **13c** (red), **16c** (orange), **21c** (green), **31c** (blue) and **41c** (violet). For **16c**, in addition, kinetics in the presence of 2 M, 4 M and 6 M urea are shown (dark to light grey). The solid lines represent double exponential functions fitted to data. The intrinsic Xan lifetime in the donor-only helix **21d** (black) is given for comparison. Measurements were performed in 10 mM potassium phosphate, pH 7.0 at 5 °C.

Figure 4.2 shows that the fastest triplet transfer is observed in the mainly unstructured peptide **13c**, the slowest transfer in the longest and most helical peptide **41c**. The addition of increasing amounts of urea, as shown for **16c**, leads to faster TTET kinetics due to unfolding of the helices.¹⁶⁵ The Xan triplet decays of all double-labelled peptides can be well described by the sum of two exponential functions (Figure 4.2), as expected for a three-state system as depicted in Eq. 11. The two apparent rate constants λ_1 and λ_2 and their amplitudes are a function of the microscopic rate constants k_f , k_u and k_c . If the

microscopic rate constants k_f , k_u and k_c are in a similar range, they can be obtained from a global fit of the apparent rate constants λ_1 and λ_2 in combination with the corresponding amplitude information (see 10.A and 10.B). In order to get reliable data on k_f , k_u and k_c the kinetics were measured at different urea concentrations, i.e. over a wide range of helix stabilities. With exception of **11c** and **13c**, in which the Xan triplet decays are basically single exponential in ≥ 2 M urea, the Xan triplet decays curves in all other peptides are double exponential over the whole range of denaturant concentrations. The apparent rate constants λ_1 and λ_2 and corresponding amplitudes of **16c**, **21c**, **31c** and **41c** measured as functions of the urea concentration are shown in Figure 4.3. The apparent rate constants are only slightly affected by the addition of urea, whereas the amplitudes change drastically. At low concentrations of urea the slow phase has the major amplitude, at high concentrations the fast phase prevails.

The microscopic rate constants k_f , k_u and k_c were determined by a global fit of the Xan triplet decays at different urea concentrations (see 10.B), assuming that the logarithms of the microscopic rate constants k_f , k_u and k_c are linear functions of the molar urea concentration (Eq. 12).

$$\ln k_i = \ln k_i^0 - \frac{m_i \cdot [\text{urea}]}{RT} \quad (12)$$

This model is able to describe the behaviour of **16c** (Figure 4.3A), **21c** (Figure 4.3B) and **31c** (Figure 4.3C) and of the N-terminally labelled helices **16n**, **21n** and **31n** (data not shown). The apparent rate constants and amplitudes predicted from the global fits are nearly identical to those obtained from the double exponential fits of the single decay curves. Deviations are limited to λ_1 in the absence of denaturant, for which large uncertainties are expected due to the low amplitude of the phase ($< 15\%$). The model becomes apparently less suitable to describe the kinetics in **41c** (Figure 4.3D). Here almost all apparent rate constants predicted from the global fit deviate from those obtained from fitting the decay curves individually. Applying the same model, as used for the global fit of the triplet decays, to fit λ_1 , λ_2 and the relative amplitudes individually, gives significantly different results (Figure 4.3D, broken lines). Therefore no reliable results can be obtained for helix folding and contact formation in **41c**. The rate constant for helix unfolding of **41c** obtained by the different fitting procedures appears to be more robust.

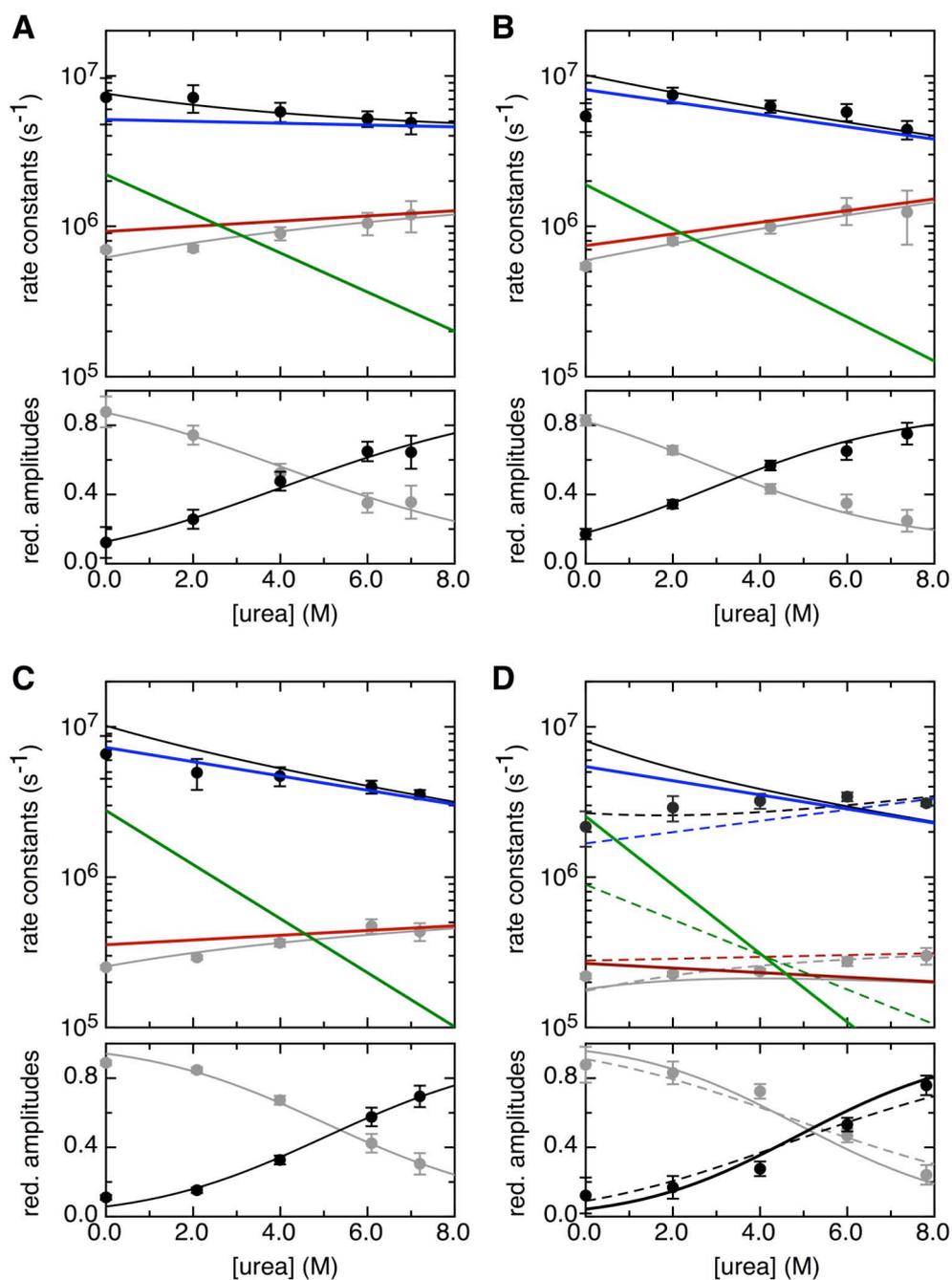


Figure 4.3. Urea-dependence of the TTET kinetics and results of the global fitting procedure for (A) **16c**, (B) **21c**, (C) **31c** and (D) **41c**. The upper panels give the apparent rate constants λ_1 (●) and λ_2 (●) of single Xan triplet decays as a function of urea, the lower panels the reduced amplitudes. The black and grey lines show the predicted rate constants and amplitudes from a global, simultaneous fit of the Xan decay curves. The obtained microscopic rate constants k_f (green), k_u (red) and k_c (blue) are shown as solid lines. For **41c**, in addition, results of an alternative fitting procedure are shown (dashed lines, see text). The measurements were performed in 10 mM potassium phosphate, pH 7.0 at 5 °C.

The results for all investigated peptides are summarised in Figure 4.4. For **41c** the different results obtained by the two fitting procedures, global or individual fits to the traces, are given. The time constants of contact formation $1/k_c$ are in the range of 80 ns to 180 ns (Figure 4.4A,B). The directly measured rate constants of contact formation in the unstructured peptides **11c** and **13c** are comparable to those obtained for **16n**, **21n** and **31n** using the three-state model (Eq. 11). In **16c**, **21c**, **31c**, and presumably also in **41c**, k_c is somewhat lower, which might be explained by the influence of stable helical structures or the influence of tails adjacent at both sides of the probed segments. In a study on the effect of end-extensions on contact formation in unstructured peptides tails reduced k_c up to a factor of 2.5.¹³⁸ Contact formation in peptides with extensions on both sides is additionally 1.7-fold slower than in peptides with one tail.¹³⁸

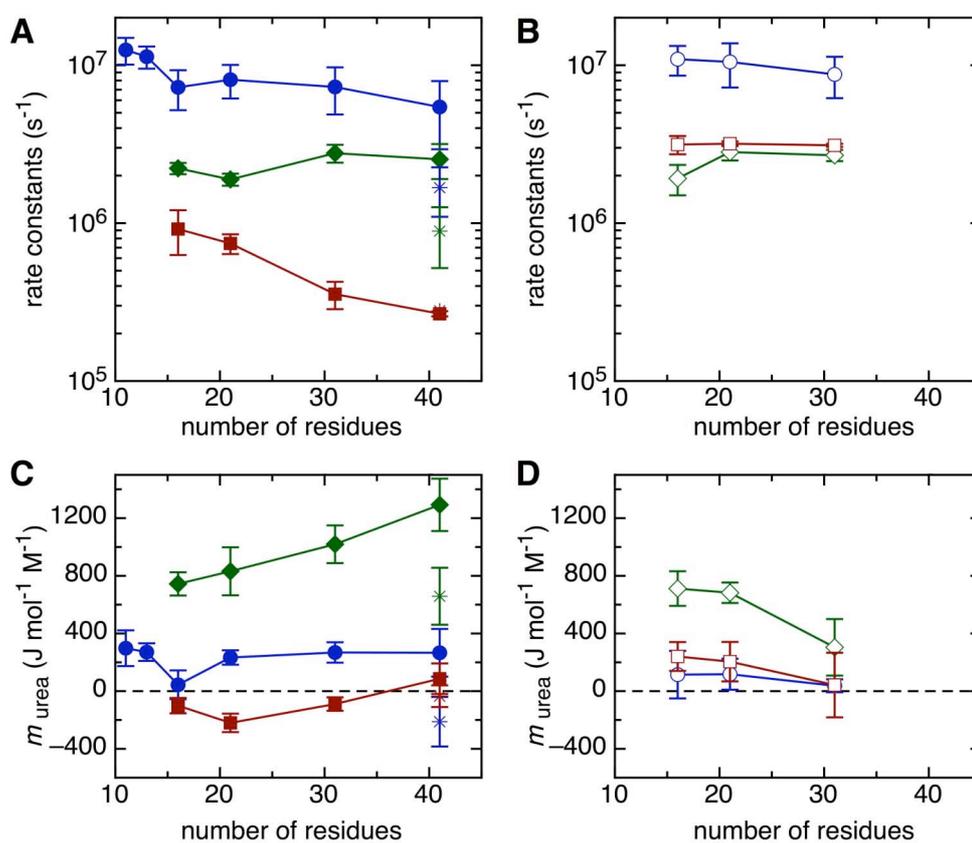


Figure 4.4. Influence of the peptide length on the local dynamics in the centre (A,C) and at the N-terminus (B,D) of α -helical peptides. (A,B) Microscopic rate constants of local helix folding k_f (\blacklozenge), unfolding k_u (\blacksquare) and contact formation k_c (\bullet). (C,D) Corresponding m -values (see Eq.12), colours and symbols as in (A). Results of the alternative fitting procedure for **41c** are given as well ($*$). The measurements were performed at 5 °C in 10 mM potassium phosphate, pH 7.0.

The m -value of contact formation m_c is around $250 \text{ J mol}^{-1} \text{ M}^{-1}$ in all centrally labelled helices (Figure 4.4C) and somewhat lower for contact formation at the N-terminus (Figure 4.4D). In all peptides with 16 or more residues the rate constants of helix folding k_f in the centre are higher than the rate constants of unfolding k_u , indicating that stable helices are formed in this region (Figure 4.4A). For **16c** and **21c** similar stabilities are found, as folding in both peptides is around 2.5-fold faster than unfolding. With increasing length the stability at the central positions becomes higher and in **31c** folding is almost ten times faster than unfolding. The stabilisation with peptide length appears to be mainly caused by a reduction of k_u , which is in **41c** almost four times lower than in **16c**, whereas only small changes were observed for k_f . In contrast, the helix dynamics at the N-terminus do not vary significantly with length (Figure 4.4B), with exception of helix **16n** in which folding was found to be somewhat slower. In **21n** and **31n** folding and unfolding are equally fast, i.e. at the N-terminus the equilibrium constant of the helix stability is one and does not depend on the peptide length. The sensitivities of k_f and k_u towards urea are summarised in Figure 4.4C,D. In the centre the m -value for helix folding m_f increases with chain length, whereas the trend is less pronounced for m_u . At the N-terminus a drastic reduction of m_f is observed for **31n**.

In summary, the dynamics at the N-terminus of helical peptides are almost independent of the peptide length. In the helix centre, unfolding is considerably slowed down with increasing peptide length, whereas we do not find a pronounced effect on the rate constants of helix folding. In order to discuss these observations it is not sufficient to remain in the local two-state description, but the state of the neighbouring segments has to be taken into account as well. In the next section the experimentally determined rate constants are compared to predictions from simulations based on a kinetic Ising model.

4.1.3 Comparison to Predictions from a Kinetic Ising Model

Isolated α -helices are systems of low cooperativity and do hence show multi-state behaviour. The measurement of local helix folding and unfolding provides rate constants of microscopically well-defined transitions, which can be directly simulated with a kinetic Ising model in a Monte-Carlo approach. The model and the simulations are described in detail elsewhere (see 10.A). Briefly the peptide is simplified to a chain of N identical segments (homopolymer), which can be either in the helical state h or the coil state c . A statistical weight is assigned to each segment based on its state and the states of the nearest neighbours, as it also done in the 2×2 approximation of Zimm and Bragg.^{155,156} Next, probabilities for transitions from c to h and *vice versa* are defined, which take as

well the states of the nearest neighbours into account and satisfy the equilibrium weights.¹⁸¹ Within every discrete time-step of the simulation, the probability for each segment to change is compared to randomly generated numbers in order to decide, whether the particular transition takes place or not.

Starting from an arbitrarily chosen helical conformation the behaviour of the system is followed over many time steps ($> 10^7$). To extract the local dynamics the times that a specified segment needs to change from *c* to *h* (folding) or to change from *h* to *c* (unfolding) are measured and stored in histograms. The obtained histograms of first passage times are then approximated by a single rate constant, giving reduced rate constants of local helix folding k_f/k_1^0 and unfolding k_u/k_1^0 (Figure 4.5). These are normalized to k_1^0 , the basic rate constant for adding one segment to an existing helix in the absence of denaturant, which is used to define the rate constants of all other transitions as well. The comparison of the experimental data with the predictions from simulations suggests that $k_1^0 \approx 2 \cdot 10^7 \text{ s}^{-1}$ (see 10.A). The model can be used to investigate the dynamics at different positions (Figure 4.5A) and can be adapted to include the effect of denaturants (Figure 4.5C,D). The following discussion focuses on the effect of the peptide length, whereas the position- and urea-dependence is discussed in detail elsewhere (see 10.A). In general the predicted m_f - and m_u -values were found to be systematically lower by around $300 \text{ J mol}^{-1} \text{ M}^{-1}$ than the measured m -values, which might be attributed to the influence of urea on the chain dynamics (see 10.A). The m_f - and m_u -values given in Figure 4.5D are corrected for this effect.

Figure 4.5A shows the reduced rate constants of local helix folding k_f/k_1^0 and unfolding k_u/k_1^0 in systems of different length as a function of the position. A general observation is that for all lengths local helix folding varies only slightly with the position, whereas unfolding is slow in the helix centre and becomes faster towards the termini. Both has been experimentally observed in the previously studied 21mer (see 10.A). The simulations suggest that the unfolding rate constants reach a plateau in the centre of long systems, which is not yet the case in the previously studied 21mer. For the local stability in the centre of a 17mer an equilibrium constant $K_{\text{eq}} = k_f/k_u$ of one is found. The local stability increases with peptide length, up to $K_{\text{eq}} = 35$ in the centre of a 41mer. At the termini the local dynamics are almost length independent and $K_{\text{eq}} \approx 1$.

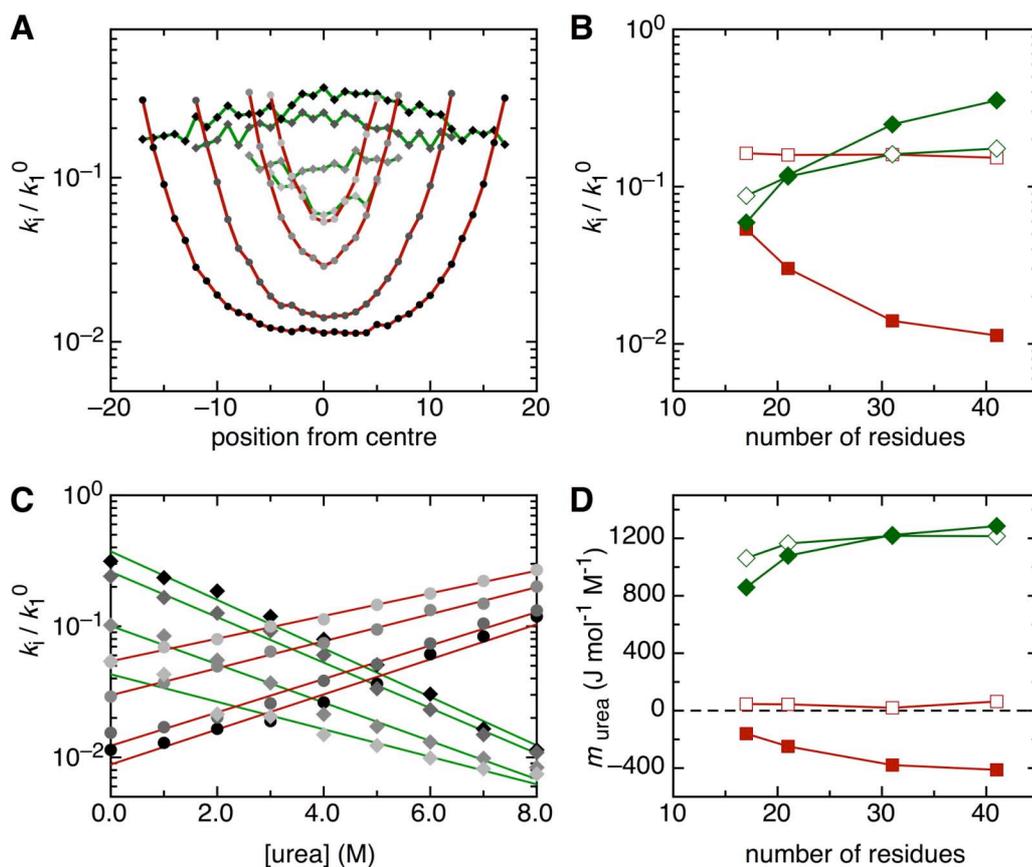


Figure 4.5. Results obtained from simulations on local folding and unfolding in α -helical peptides of different length. (A) Reduced rate constants of local helix folding k_f/k_1^0 (green) and unfolding k_u/k_1^0 (red) as function of the position in systems of 17, 21, 31 and 41 residues (grey to black). (B) Reduced rate constants of folding (\blacklozenge, \diamond) and unfolding (\blacksquare, \square) as function of chain length. Open symbols (\diamond, \square) denote the dynamics at the N-terminus (position 5), closed symbols ($\blacklozenge, \blacksquare$) the dynamics in the centre. (C) Reduced rate constants of folding and unfolding in the helix centre as function of the urea concentration, coloured as in (A). (D) Corresponding m_f - and m_u -values (shifted by $+300 \text{ J mol}^{-1} \text{ M}^{-1}$, see 10.A), colours and symbols as in (B). For the simulations $s^0(5^\circ\text{C}) = 1.3059$, $\sigma = 0.003$, $\gamma_h = \gamma_c = 1$, $m_{\text{eq}} = 96 \text{ J mol}^{-1} \text{ M}^{-1}$ and $\alpha_D = 0.5$ were used,^{161,165} for all other parameters refer to Manuscript 10.A. Closed segments were defined as at least four out of five residues in the helical state.

In the helix centre folding is predicted to become faster and unfolding slower with increasing peptide length (Figure 4.5B). In the 41mer k_f/k_1^0 is two times higher and k_u/k_1^0 fivefold lower than in the 16mer. In the shorter peptides the chain length is predicted to have a strong effect on k_f/k_1^0 and k_u/k_1^0 , whereas the dynamics in the 31mer and 41mer differ only slightly. In the experiments a decrease of k_u with peptide length was observed, with a much larger difference between **21c** and **31c** than between **31c** and **41c** (Figure 4.4A,B). However, the strong increase of k_f/k_1^0 with peptide length was not

observed experimentally. As a consequence, the measured local helix stabilities in **31c** and **41c** are much lower than predicted. One explanation for the discrepancy between the simulations and experimental data might be that the three-state model used to evaluate the data (Eq. 11) does not longer work in the centre of long, stable helices, which is also indicated by the complications to fit the kinetics of **41c** (see above). For the dynamics at the N-terminus a good agreement between the experiments and the simulation is found, as k_f is predicted to increase only slightly in short peptides and then becomes equal to k_u , which is length independent (Figure 4.4B and Figure 4.5B).

The response of the simulated system towards urea is shown in Figure 4.5C,D. In the studied size range helix folding in the centre is predicted to become slower and unfolding faster with increasing concentrations of urea (Figure 4.5C). The behaviour can be approximated by linear relations between $\ln k_i$ and the molar urea concentration, as it is also done in the evaluation of the experiments (Eq. 12). Only for long systems the simulations predict a deviation of $\ln k_u$ from linear behaviour. The deviations are however too small to explain the failure of the global fitting procedure in the case of **41c**. Figure 4.5D shows the resulting m_f - and m_u -values at the N-terminus and in the helix centre. The values for m_f are predicted to be high, whereas m_u is slightly negative in the centre and virtually zero at the N-terminus. The difference increases somewhat with peptide length and becomes constant for longer helices. In contrast, the experiments showed that both, m_f and m_u in the helix centre increase with chain length. The predicted m_u -values for unfolding are higher at the N-terminus than in the centre, which is in agreement with the experiments.

Conclusion. The dynamics of local helix folding and unfolding have been investigated in alanine-based peptides from 11 to 41 residues using TTET. In an 11mer and a 13mer no stable helices are formed, whereas from 16 residues on helical structure has been detected at the N-termini and in the centres of the peptides. In the studied length range, both the global helix content and the local stability in the helix centre increase with peptide length. Local helix unfolding gets about fourfold reduced from $(1.1 \pm 0.5) \mu\text{s}$ in the 16mer to $(3.7 \pm 0.4) \mu\text{s}$ in the 41mer, whereas folding appears to be independent of the peptide length, occurring with time constants between 360 ns and 500 ns. The dynamics at the N-terminus are similar in all peptides. The kinetic Ising simulations predicted similar behaviour with one major exception - helix folding is expected to become faster with increasing peptide length, leading to higher stabilities in longer systems as experimentally observed. The origin of this discrepancy is presently unclear.

Gai and coworkers studied similar peptides with varying length (19 to 39 residues) by T -jump measurements.¹⁸⁷ The relaxation time constants after T -jumps from 1 to 11 °C were virtually length independent. The authors compared these relaxation times to calculated first passage times of helix formation, which were predicted to increase with peptide length.²⁴³ The observed discrepancy was attributed to a variation of the elongation parameter s with peptide length.¹⁸⁷ However, relaxation kinetics obtained with global probes after small shifts of the helix-coil equilibrium towards the unfolded state do not directly represent the process of helix formation starting from pure random coil. A direct comparison of the relaxation data obtained by a global probe with the results of our local folding/unfolding experiments is not possible.

4.2 Peptide Model of the Foldon Hairpin

The studies on the monomeric states of the foldon domain from bacteriophage T4 fibrin indicate that the amino acid residues Ala12 to Leu23 form a highly stable β -hairpin. In the monomeric foldon variant E5R the hairpin region is associated with the N-terminal part of foldon (see 10.D), whereas in the A-state (pH 2.0) no tertiary interactions with other parts of the protein were detected.¹⁹⁹ In order to prove that foldon encompasses an autonomously forming β -hairpin under physiological conditions, model peptides were synthesised (**H-Fhp-OH** and **Ac-Fhp-NH₂**, Table 4.2) and characterised by CD and NMR spectroscopy.

Table 4.2. Synthesised Peptides Comprising the Hairpin Region of Foldon

Name	Sequence
H-Fhp-OH	H- ¹² AYVRKDGEWVLL ²³ -OH
Ac-Fhp-NH₂	Ac- ¹² AYVRKDGEWVLL ²³ -NH ₂
D/A-Fhp-1	Xan- ¹³ <u>F</u> VRKDGE <u>Z</u> VLL ²³ -OH
D/A-Fhp-2	Xan- ¹³ <u>V</u> VRKDGE <u>Z</u> VLL ²³ -OH

The numbers refer to the positions in the fibrin foldon domain.²²⁶ In **D/A-Fhp-1** and **D/A-Fhp-2** (see 4.2.2) the N-terminus is acylated with xanthonic acid (Xan) and Trp20 is replaced by 1-naphtylalanine (Z, Nal).

In addition, the foldon hairpin might be also suited to study hairpin folding and unfolding with a similar approach used to characterise the local dynamics of α -helix formation (see 3.1). For this purpose two foldon hairpin peptides with triplet donor and acceptor labels were synthesised and investigated (**D/A-Fhp-1** and **D/A-Fhp-2**, Table 4.2).

4.2.1 Structure and Stability

CD spectroscopy. Two peptides, **H-Fhp-OH** and **Ac-Fhp-NH₂**, with the natural foldon hairpin sequence but different termini were synthesised (Table 4.2). In **H-Fhp-OH** the termini are free, whereas in **Ac-Fhp-NH₂** the N-terminus is acetylated (Ac-) and the C-terminus amidated (-NH₂). The peptides are well soluble in aqueous solution (≥ 1 mM), both at pH 2.0 and pH 7.0.

Figure 4.6A shows Far-UV CD spectra of **H-Fhp-OH** and **Ac-Fhp-NH₂** at different pH values and temperatures. The CD signal has low intensity and the spectra are dominated by contributions from the aromatic side chains.²⁴⁴ At pH 2.0 and 5.0 °C the spectrum of **H-Fhp-OH** shows minima at 195 nm and 210 nm and a positive band at

228 nm. The latter band is likely arises from the 1L_a transition of Tyr13 (~ 230 nm) and/or the 1B_b transition of Trp20 (~ 225 nm),²⁴⁴ coupled with $\pi\pi^*$ and $\pi\pi^*$ transitions of amide groups. The intensity of the CD bands of up to $[\Theta] = 10^5$ deg cm² dmol⁻¹ can only be explained with the chromophores fixed in an asymmetric environment. Further evidence for structure formation involving Tyr13 and/or Trp20 is also given by a positive duplet centred at 277 nm ($[\Theta] = 2 \cdot 10^3$ deg cm² dmol⁻¹, data not shown).

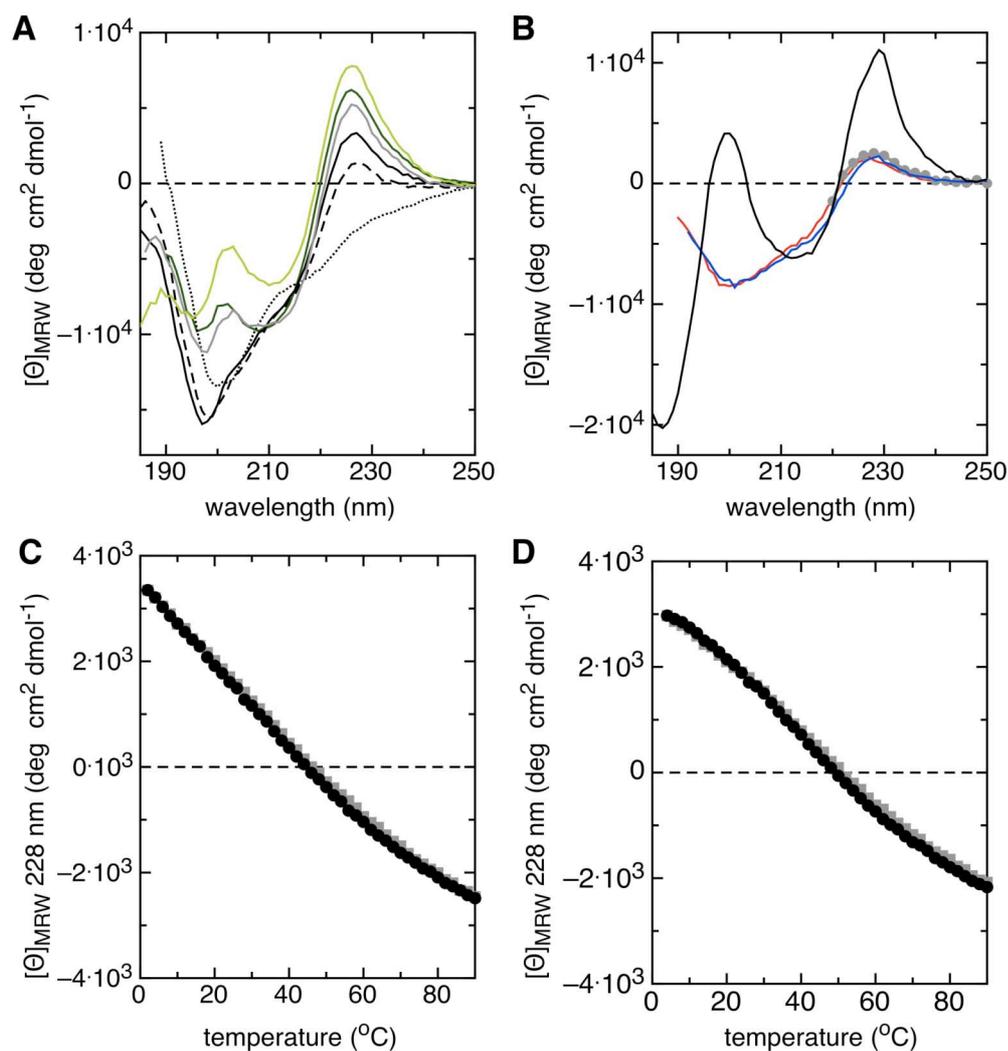


Figure 4.6. (A) Far-UV CD spectra of **H-Fhp-OH** (pH 7.0 dark green, pH 2.0 light green) and **Ac-Fhp-NH₂** (pH 7.0 black, pH 2.0 grey) at 5.0 °C (solid lines). In addition the spectra of **Ac-Fhp-NH₂** at 22.5 °C (dashed line) and 90.0 °C (dotted line) are shown at pH 7.0. (B) Far-UV CD spectra of the native foldon trimer (black), the monomeric folding intermediate (grey), the monomeric foldon variant E5R (blue) and the foldon A-state (red, pH 2.0) measured at 22.5 °C. Taken from Manuscript 10.D. (C) Temperature-induced equilibrium unfolding of **Ac-Fhp-NH₂** and (D) of foldon E5R monitored at 228 nm (heating: black, cooling: grey) in 10 mM potassium phosphate, pH 7.0.

Normalized CD spectra with at least two different peptide concentrations were in all cases super-imposable, indicating that no oligomerisation occurs in the concentration range of 5 - 200 μ M (see below). Strong deviations from classical β -sheet spectra due to aromatic side chain contributions have also been observed for other β -hairpin peptides.^{207,208} Similar CD spectra have been observed for the A-state of wild type foldon and the monomeric foldon variant E5R (Figure 4.6B). The characteristic band at 228 nm is present in all cases. A quantitative comparison of the spectra, however, is not possible, since the relative contributions of the aromatic side chains are larger in the 12mer hairpin peptides than in the 27 residue full-length foldon.

In the hairpin peptides the band at 228 nm is less intense at pH 7.0 than at pH 2.0 (Figure 4.6A). **Ac-Fhp-NH₂** at pH 2.0 has a spectrum comparable to that of **H-Fhp-OH** at pH 7.0. In the spectrum of **Ac-Fhp-NH₂** at pH 7.0 the ellipticity at 228 nm is further reduced and a pronounced minimum at 197 nm appears. With increasing temperature both peptides give spectra typical for random coil at high temperature. The thermal melting followed by the ellipticity change at 228 nm is consistent with a loss of structure (Figure 4.6C). The transition is however broad and no defined baselines are observed, impeding a quantitative thermodynamic analysis. The thermal melting of foldon E5R under identical conditions shows a more sigmoidal course (Figure 4.6D), indicative for a higher degree of structure formation (see 10.D).

NMR spectroscopy. In order to obtain more information on the structure of the peptides **H-Fhp-OH** and **Ac-Fhp-NH₂** one- (1D) and two-dimensional proton NMR spectra were recorded. As example the 1D ¹H-spectrum of **Ac-Fhp-NH₂** with well-dispersed signals in the amide region is shown (Figure 4.7). A complete assignment of the resonances was achieved using total correlation spectroscopy (TOCSY, Figure 4.7) and nuclear Overhauser enhancement spectroscopy (NOESY, see below). The complete assignment for **Ac-Fhp-NH₂** at 2 °C and pH 7.0 is listed in Table 4.3. Further assignments for **Ac-Fhp-NH₂** at 25 °C, pH 7.0, **Ac-Fhp-NH₂** at 2 °C, pH 2.0 and **H-Fhp-OH** at 2 °C, pH 7.0 are given in the Appendix (Table 7.1 - 7.3).

The amide proton (^NH) resonances of Val14, Lys16, Asp17 and Val21 are strongly downfield shifted, whereas the ^NH resonance of Glu19 experiences a strong upfield shift (Figure 4.7). In addition, the ^NH resonances of Val14, Asp17 and Val21 are extremely broadened, indicating exchange processes on the NMR timescale. The vicinal ³J_{H^NH^α coupling constants are generally between 7 – 8 Hz (Table 4.3), pointing to an extended conformation.¹⁴ The coupling constants are somewhat lower than expected for perfect}

anti-parallel β -sheets with ϕ -angles of 140° , which would give coupling constants of ~ 9 Hz.¹⁴ The geometry of isolated hairpins is however less perfect and the β -hairpin is likely in fast exchange with unfolded conformations, which leads to an averaging and reduction of the coupling constants.

next page →

Figure 4.7. ^1H -NMR characterisation of **Ac-Fhp-NH₂** at 2°C in 10 mM potassium phosphate, pH 7.0, 5 % (v/v) D₂O. The upper panels show regions of a 1D spectrum obtained with jump-and-return water suppression, the lower panels regions of a TOCSY spectrum with Watergate solvent suppression and MLEV-17 spin lock (110 ms mixing time). Left panels: H^N and aromatic protons with cross peaks to H ^{α} and side chain protons. Right panels: side chain protons with cross peaks to H ^{α} and side chain protons. The assignment is given in Table 4.3.

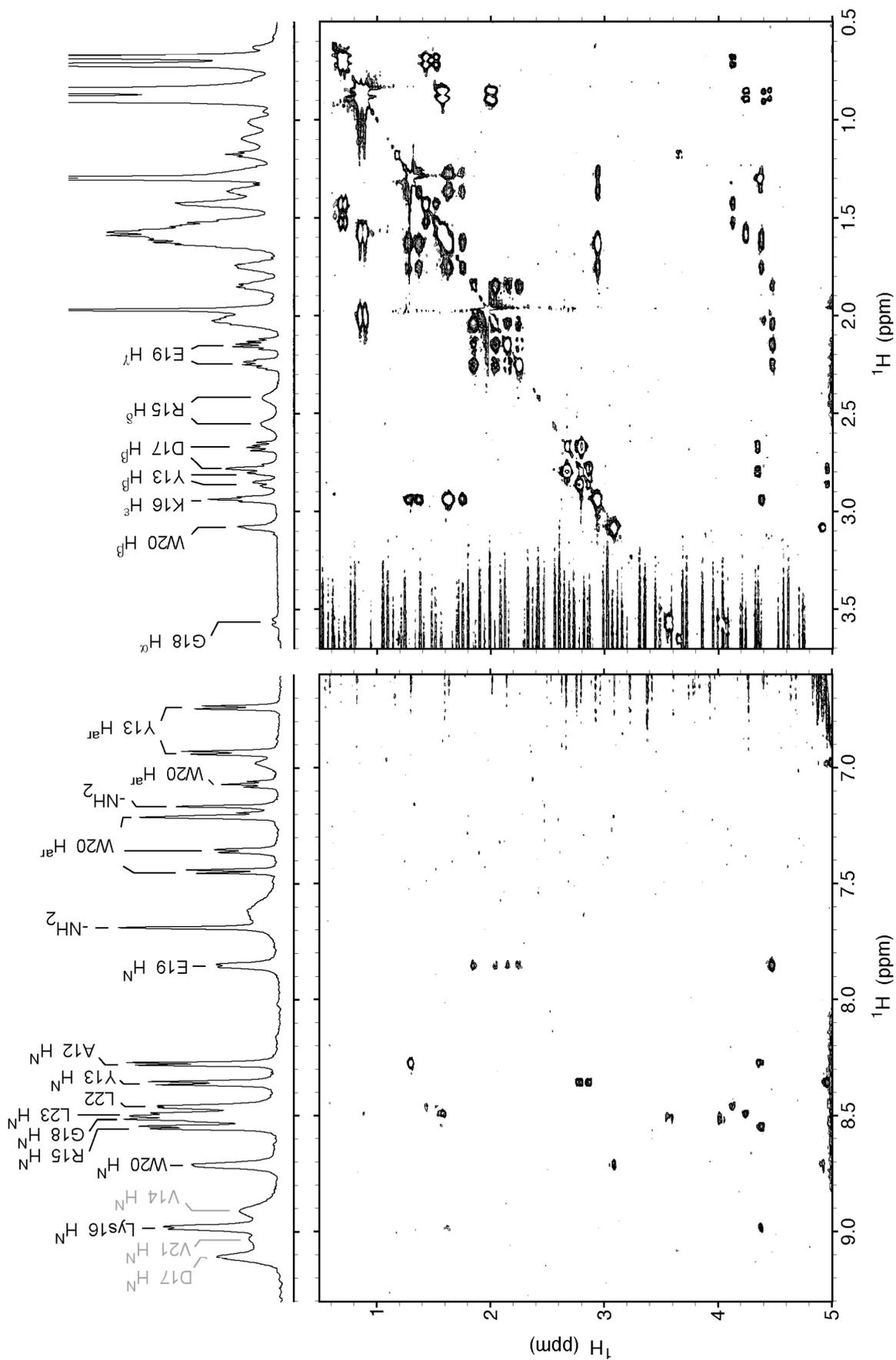


Table 4.3. Resonance Assignment of Ac-Fhp-NH₂ at pH 7.0 and 2 °C.

Residue	H ^N	³ J _{H^NH^α}	H ^α	H ^β	Other
Ac-	-	-	1.974		
Ala 12	8.274	7.0 Hz	4.381	1.299	
Tyr 13	8.358	8.1 Hz	4.960	2.858 2.786	H _{2,6} 6.932 H _{3,5} 6.736
Val 14	8.909 ^a	n.d. ^a	4.451	1.985	H ^γ 0.887, 0.848
Arg 15	8.548	7.0 Hz	4.382	1.557	H ^γ 1.136, 1.047 H ^δ 2.554, 2.418
Lys 16	8.982	7.6 Hz	4.378	1.751 1.603	H ^γ 1.362, 1.270 H ^δ 1.626 H ^ε 2.936
Asp 17	9.111 ^a	n.d. ^a	4.348	2.790 2.665	
Gly 18	8.510	-	4.029 3.569		
Glu 19	7.849	7.7 Hz	4.475	2.041 1.846	H ^γ 2.248, 2.145
Trp 20	8.713	7.6 Hz	4.919	3.083	H _{2,6} 7.208 H ₄ 7.355 H ₅ 7.069 H ₇ 7.443 HN 10.152
Val 21	9.039 ^a	n.d. ^a	4.396	2.018	H ^γ 0.902, 0.855
Leu 22	8.461	6.5 Hz	4.126	1.521	H ^γ 1.425 H ^δ 0.714, 0.676
Leu 23	8.492	n.d. ^b	4.242	1.58 ^b	H ^γ 1.58 ^b H ^δ 0.891, 0.849
-NH ₂	7.686 7.163	-			

Chemical shifts in ppm relative to internal DSS. Measurements were performed in 10 mM potassium phosphate with 5 % (v/v) D₂O.

^a Very broad resonance due to chemical exchange.

^b Not resolved.

The chemical shifts of the H^N and alpha protons (H^α) in **H-Fhp-OH** and **Ac-Fhp-NH₂** closely resemble those found in the monomeric foldon variant E5R (see 10.D) and in the A-state monomer.¹⁹⁹ The interpretation of the chemical shifts, however, is complicated by the presence of aromatic residues. In β -hairpins especially inward pointing H^α are strongly affected by ring currents. Nevertheless, strong downfield shifts are observed for the H^α resonances of Tyr13, Val14, Glu19, Trp20 and Val21 as expected for β -sheet conformations.²⁴⁵ A strong upfield shift is observed for the H^N resonance of Glu19, which is typical for the first strand position following type I' turns. Typical is as well the positive CSD of the H^N resonance of the first turn residue Asp17.²⁴⁶ The H^N resonances of Val14, Lys16 and Val21 show downfield shifts, as expected for amide protons engaged in strong and intact H-bonds (see Figure 4.9 and Figure 4.10).²⁴⁷ H-bonding is also likely the reason for the slow exchange process leading to the line broadening of the Val14 and Val21 H^N resonances (Figure 4.7). A further similarity between the spectra is the strong broadening of the Asp17 H^N resonance and the large splitting of the H^α resonances of Gly18 in the turn region (0.46 ppm for **Ac-Fhp-NH₂** at 2 °C, pH 7.0, see Table 4.3).

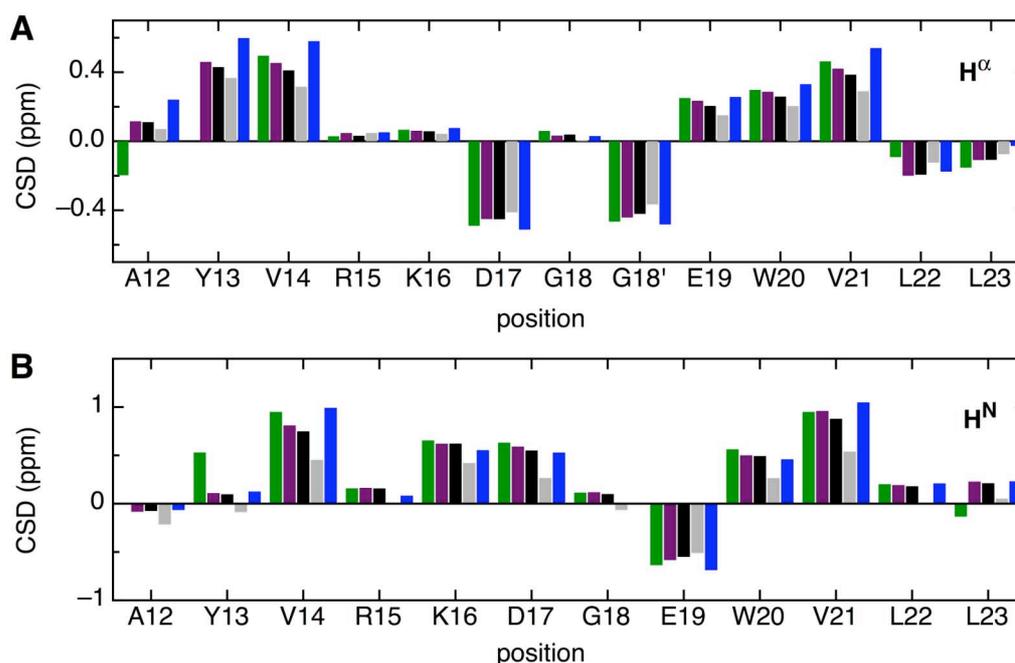


Figure 4.8. Chemical shift deviations (CSD) for (A) alpha protons and (B) amide protons in the foldon hairpin region. Compared are **H-Fhp-OH** at 2 °C and pH 7.0 (green), **Ac-Fhp-NH₂** at 2 °C and pH 2.0 (violet), **Ac-Fhp-NH₂** at 2 °C and pH 7.0 (black), **Ac-Fhp-NH₂** at 25 °C and pH 7.0 (grey) and the foldon variant E5R at 25 °C and pH 7.0 (blue, see 10.D). Random coil values were taken from Schwarzsinger *et al.*^{248,249}

Another level of structural information is provided by inter-residue NOE cross peaks, discussed here for **Ac-Fhp-NH₂**. Strong sequential NOEs are found for all $d_{\alpha\text{N}}(i, i + 1)$ distances (Figure 4.10), typical for extended conformations like anti-parallel β -sheets.^{14,250} Sequential H^N-H^N NOEs are only observed from Lys16 to Trp20, i.e. in the turn region, where short $d_{\text{NN}}(i, i + 1)$ distances are expected.^{14,250}

Next to the sequential information, cross-strand NOEs provide information on backbone interactions and the orientation of side chains (Figure 4.10). The only non-sequential H^N-H^N NOE in **Ac-Fhp-NH₂** is observed between Lys16 and Glu19 across the turn region. A short $d_{\text{NN}}(i, i + 3)$ distance at this position has also been found for other peptides with a type I' turn.^{206,211} The proximity of these regions is further supported by the H ^{α} -H^N NOE between Trp20 and Lys16. The determination of H ^{α} -H ^{α} NOEs should provide more detailed information on the alignment of the strands, but affords measurements in D₂O, which have not been performed yet. However, a number of non-sequential side chain – backbone and side chain – side chain NOEs has been observed. The medium- and long-range NOEs mainly involve Tyr13, Trp20 and Leu22 (Figure 4.9). These long-range NOEs clearly confirm that the peptide chain adopts an anti-parallel conformation. Tyr13, Trp20 and Leu22 are part of the hydrophobic core of the native foldon trimer²²⁶ and form the key interactions in the A-state hairpin¹⁹⁹ and the E5R hairpin (see 10.D). No NOEs were detected that would point to a trimerisation of the hairpin peptides.

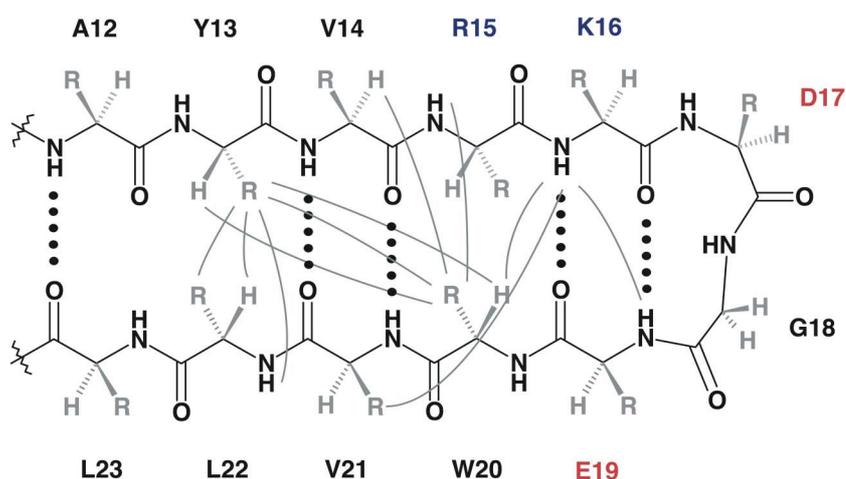


Figure 4.9. Schematic drawing of the foldon β -hairpin found in the foldon trimer²²⁶, the acidic state monomer¹⁹⁹ and the monomeric foldon variant E5R (see 10.D). The two anti-parallel β -strands are linked by type I' β -turn, consisting of Asp¹⁷-Gly¹⁸. The dots indicate the H-bonds present in the native trimer. The grey lines show medium- and long-range NOE cross peaks detected in **Ac-Fhp-NH₂** at 2 °C and pH 7.0 (see text and Figure 4.10).

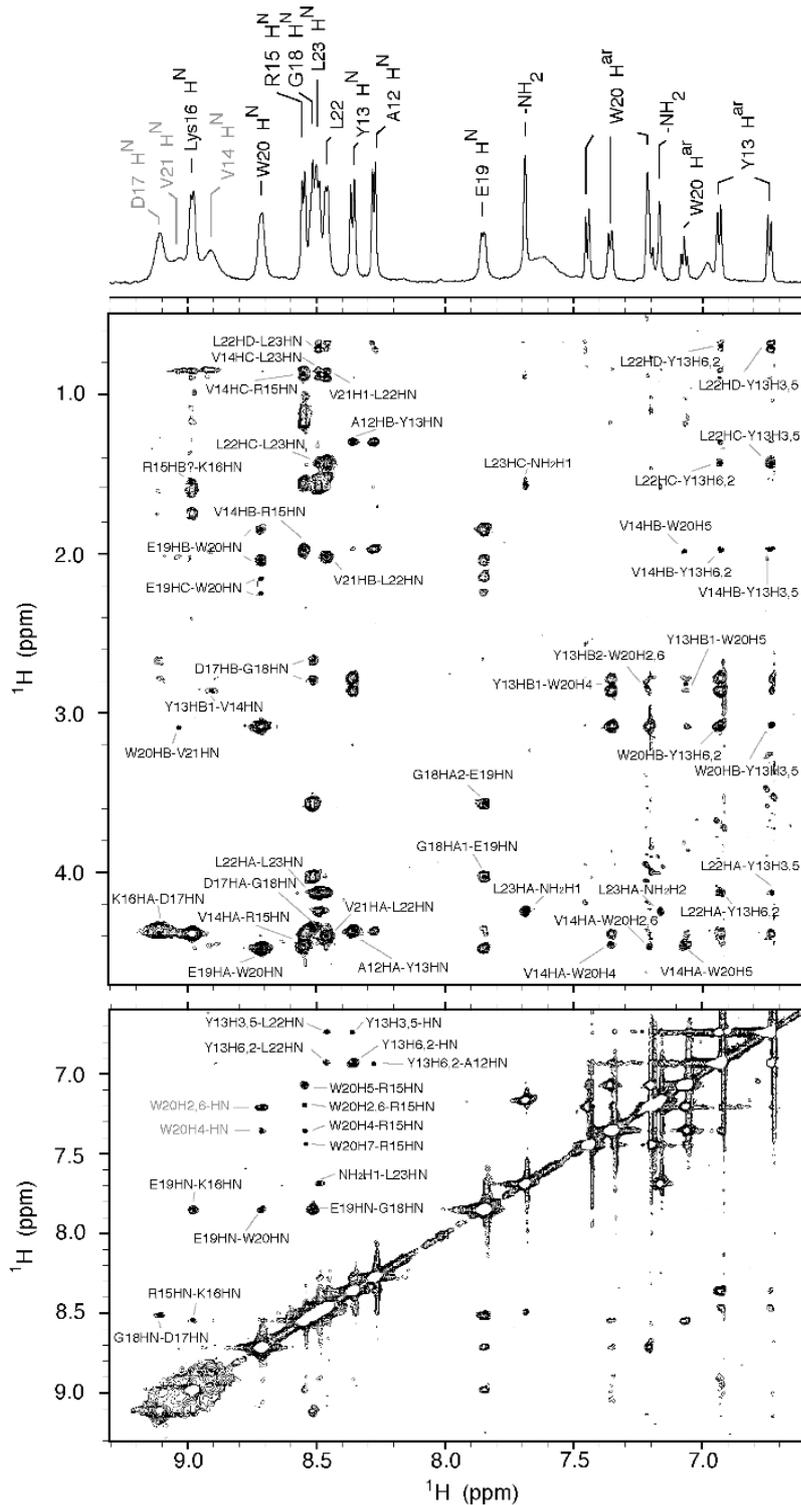


Figure 4.10. Regions of a NOESY spectrum of **Ac-Fhp-NH₂** at 2 °C in 10 mM potassium phosphate, pH 7.0, 5 % (v/v) D₂O. The spectrum was recorded with a mixing time of 350 ms and Watergate solvent suppression. The upper panel shows H^N/aromatic – side chain proton cross peaks, the lower panel H^N/aromatic – H^N/aromatic cross peaks. Unambiguous inter-residue NOE cross peaks are labelled. At the top the amide region of the 1D spectrum is shown, cf. Figure 4.7. The assignment is given in Table 4.3.

Discussion. The CD and NMR data indicate that the twelve amino acid residue peptides **H-Fhp-OH** and **Ac-Fhp-NH₂** form a β -hairpin structure in aqueous solution (Figure 4.9), similar to the hairpin in the A-state monomer,¹⁹⁹ the monomeric foldon variant E5R (see 10.D) and the wild type trimer.²²⁶ It is not possible to obtain thermodynamic data on the stability of this structure, since the unfolding transitions are broad and are not necessarily well approximated by two-state behaviour. However, qualitative comparisons can be made, based on the strength of the CD band at 228 nm (Figure 4.6) and the chemical shift deviations from the random coil values (Figure 4.8). According to these criteria **H-Fhp-OH** appears to be slightly more structured than **Ac-Fhp-NH₂**. In addition, the hairpin formation appears to be more favoured at pH 2.0 than at pH 7.0. The first observation can be readily explained by the uncapped termini in **H-Fhp-OH**, which are in proximity and thus can add a favourable charge-charge interaction. The stabilisation at pH 2.0 has to originate from protonation of Asp17 and/or Glu19. An explanation would be an unfavourable interaction between the negatively charged forms of these residues, located in close vicinity on the same side of the turn (see 10.D). Favourable electrostatic interactions with Arg15 and Lys16 seem to be of minor relevance. However, NOE cross peaks have neither been detected between Asp17 and Glu19, nor between the side chains of Lys16 and Asp17, which are in close proximity in the hairpin of the native trimer.²²⁶ A pH-titration experiment on the native trimer followed by NMR spectroscopy revealed only slightly shifted pK_a values of Asp17 and Glu19.¹⁹⁹ These, however, would be consistent with an unfavourable interaction between Asp17 and Glu19 and a favourable interaction of Asp17 with a positively charged group, likely Lys16.

Several other sequences derived from proteins are known to form stable β -hairpins in isolation^{200,201,204,208} and many designed peptides have been characterised to study their stability determinants.²⁰⁵⁻²¹⁴ The foldon hairpin sequence (Table 4.2) comprises several features that have been described in this context. Of particular importance is the turn sequence.^{207,211,251} Most tight two residue turns have a Xaa-Gly sequence, since only glycine allows to form type I' and type II' turns that are compatible with the right-handed twist of β -sheets.^{193,211} Moreover, Xaa is often Asn or Asp, which makes the foldon hairpin to a typical exponent of this class.^{207,210} The expected type I' conformation, characterised by an α_L - γ conformation of the turn residues, is present in the native foldon trimer.^{199,221} The current NOE data are however not sufficient to fully prove that this conformation is present in **Ac-Fhp-NH₂** as well. It should be noted that for peptides with similar turn sequences conformational heterogeneity was observed, with a fraction of peptides adopting type I', the other type II' conformation.^{205,251} The strands of the foldon

hairpin are mainly composed of aromatic and β -branched residues (Tyr13, Val14, Trp20, Val21, Leu22 and Leu23), which have high intrinsic propensities for β -sheet ϕ, ψ -angles.^{208,212} In addition, these residues are well suited to form hydrophobic interactions across the strands.²¹⁴ Cross-strand salt-bridge interactions²¹³ seem to play a minor role in the natural foldon hairpin and protonation of Asp17 and Glu19 even appears to stabilise the hairpin conformation. However, free termini of Ala12 and Leu23 apparently add a stabilising interaction, as observed with **H-Fhp-OH**.

4.2.2 TTET Measurements to Study β -Hairpin Dynamics

The dynamics of β -hairpins have been studied in several relaxation experiments.²¹⁷⁻²¹⁹ The results indicate that conformational transitions mainly occur on the timescale of microseconds. Dynamics in this time range can be probed by TTET (see 3.1), which would yield more information about the nature of the transitions and would provide an independent measurement of the rate constants of folding and unfolding. Similar to the experiments on helix folding and unfolding (see 3.1), the triplet donor and acceptor labels have to be placed at distant sites of the hairpin in order to prevent contact formation in the native state. The foldon hairpin should be particularly suited for TTET measurements, since it is well soluble and its high stability does not result from a large cluster of aromatic residues, as in other model systems.^{200,207,208,214} It is important that tyrosine and tryptophan can be replaced by the triplet labels or other hydrophobic amino acids, as they are known to interfere with TTET.¹³²

Trp20, which is part of the central hydrophobic cluster, was substituted with the triplet acceptor 1-naphthylalanine (Nal, Figure 4.11, Table 4.2), having a high structural similarity to tryptophan. The exchange should therefore have a minor influence on the hairpin stability and should lead to partial shielding of Nal in the hydrophobic cluster on the inner side of the twisted hairpin structure. Tyr13 was replaced by phenylalanine (**D/A-Fhp-1**) or valine (**D/A-Fhp-2**). The triplet donor xanthonic acid (Xan) was attached at the N-terminus replacing Ala12. In the natural foldon hairpin the Tyr13 side chain is involved in a hydrophobic cluster with Trp20 and Leu22. The amino group of Tyr13, respectively Phe13, which was used to attach Xan via a rigid amide bond, points away from the hairpin region. At this position Xan should not interfere with the hairpin structure and might be sufficiently separated to form no contact with Nal as long as the hairpin is folded.

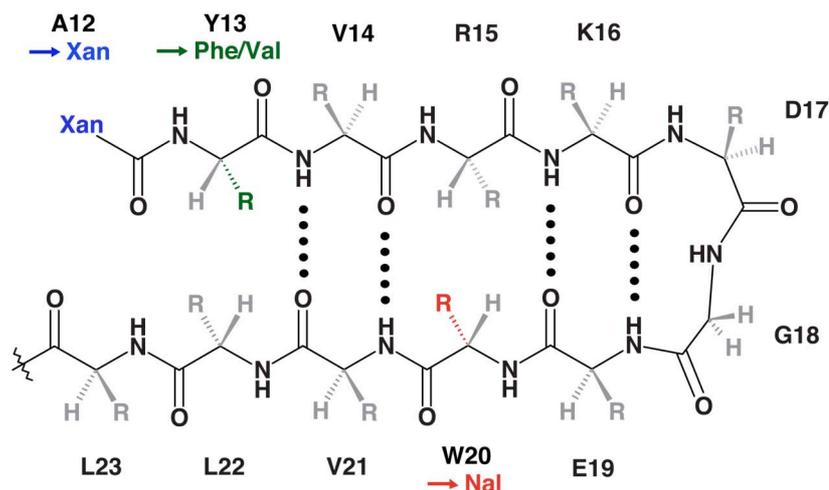


Figure 4.11. Scheme of the foldon β -hairpin with the amino acid replacements to facilitate TTET measurements. The N-terminus was acylated with the triplet donor xanthonic acid (blue), replacing Ala12. The triplet acceptor was introduced by substitution of Trp20 with 1-naphthylalanine (red). In order to avoid triplet quenching Tyr13 was replaced by Phe (**D/A-Fhp-1**) or Val (**D/A-Fhp-1**). See also Table 4.2.

After excitation of **D/A-Fhp-1** at 355 nm the intensity of the Xan triplet absorption band at 590 nm reached only 25 % of the expected signal (data not shown). One explanation would be, that a large fraction of molecules has the labels already in contact or forms contact within the 12 ns dead-time of the experiment. However, addition of GdmCl, which effectively disrupts the hairpin structure (see below), led only to a small increase of the Xan absorption band.

As alternative explanation for the low Xan triplet yield quenching by the directly neighbored Phe13 side chain has to be considered. In bimolecular quenching experiments phenylalanine at concentrations up to 50 mM had no effect of on the Xan lifetime,¹³² but in **D/A-Fhp-1** the local concentration of phenylalanine is probably much higher. To test this hypothesis, a second peptide was synthesised, in which Tyr13 was replaced by valine instead of phenylalanine (**D/A-Fhp-2**, Table 4.2).

CD and NMR-spectroscopy indicate that **D/A-Fhp-2** still forms a β -hairpin structure. The Far-UV CD spectrum is not longer dominated by contributions from tryptophan but from the triplet labels Xan and Nal (Figure 4.12A). A negative band at 206 nm, a positive band at 229 nm and broad negative band at 254 nm are observed. The strong near-UV band, either arising from Xan or Nal, is a strong indication for an asymmetric environment of these chromophores. Both the addition of denaturant (Figure 4.12A) and the increase of temperature (Figure 4.12B) disrupt the hairpin structure. The addition of GdmCl leads to

a strong reduction of the band at 229 nm and in 8 M GdmCl a negative signal is observed (Figure 4.12A), typical for denatured peptides in general and for peptides with the labels Xan and Nal, as well. Also the near-UV band changes drastically upon denaturation confirming structural change in the environment of the labels.

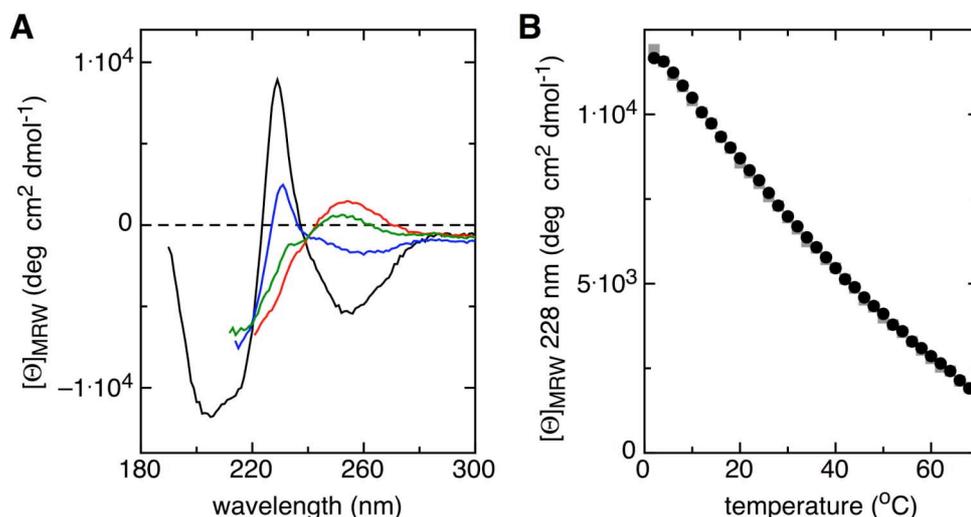


Figure 4.12. (A) Far-UV CD spectra of **D/A-Fhp-2** at 22.5 °C without denaturant (black) and in the presence of 2 M (blue), 4 M (green) and 8 M GdmCl (red). (B) Temperature-induced equilibrium unfolding of **D/A-Fhp-2** monitored by the ellipticity change at 228 nm (heating: black, cooling: grey). The measurements were performed in 10 mM potassium phosphate, pH 7.0.

The 1D $^1\text{H-NMR}$ spectrum of **D/A-Fhp-2** shows well-dispersed signals in the amide region, indicative for structure formation (Figure 7.1, Appendix). A direct comparison of the NMR spectrum with those of **H-Fhp-OH** or **Ac-Fhp-NH₂** is not possible, since the replacements of Trp20 by Nal and of Phe13 by Val change the chemical environment drastically and further resonances are added by the incorporation of Xan. Some vicinal $^3J_{\text{H}^{\text{N}}\text{H}^{\alpha}}$ coupling constants reach values of 8.5 - 9.5 Hz, confirming highly extended β -sheet conformations.

TTET measurements with **D/A-Fhp-2** reveal again a low intensity of the Xan triplet absorption band (30 %) in the absence of denaturant, but in this case the addition of GdmCl led to an increase close to the absorption expected from donor-only peptides (Figure 4.13A). The missing amplitude in the folded state probably results from fast or permanent contact formation between the labels. This would also explain the fluorescence observed in the early time region, arising from triplet-triplet annihilation. The resolved part of the Xan triplet decays can be attributed to TTET, since it is much faster than the intrinsic Xan lifetime and the formation of the Nal triplet states is confirmed by a

concomitant increase of their absorption band at 420 nm (data not shown). The positions of the triplet donor and acceptor (Figure 4.11) do therefore not separate the labels in the folded hairpin, but appear to favour contact formation.

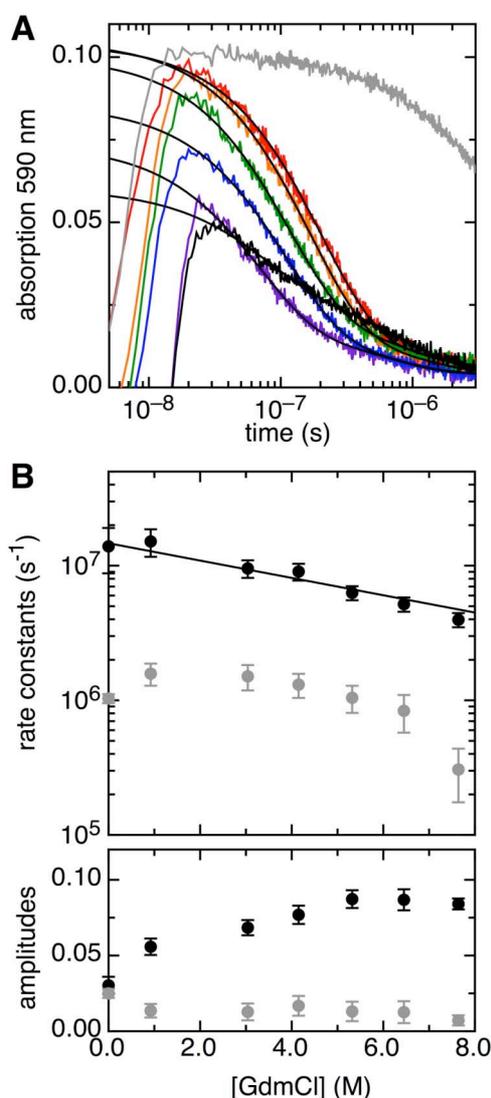


Figure 4.13. TTET measurements in **D/A-Fhp-2**. (A) Representative decays of the Xan triplet absorption band at different GdmCl concentrations (black: 0 M, violet to red: 0.9 M, 3.0 M, 4.2 M, 5.3 M, 6.5 M). At low GdmCl concentrations fluorescence is observed in the early time region up to 20 ns. For comparison the Xan triplet decay in a donor-only peptide is shown (grey). The solid lines show fits of double exponential functions. The corresponding rate constants and amplitudes are given in (B). The fast contact formation process (black circles) can be described with Eq. 12 (p. 40) and $k_c^0 = (1.5 \pm 0.1) \cdot 10^7 \text{ s}^{-1}$, $m_c = (343 \pm 55) \text{ J mol}^{-1} \text{ M}^{-1}$ (solid line). The measurements were performed with 50 μM peptide in 10 mM potassium phosphate, pH 7.0 at 22.5 $^\circ\text{C}$.

At all GdmCl concentrations the Xan triplet decays can be well described by the sum of two exponential functions (Figure 4.13A). The addition of GdmCl increases the amplitude of the fast phase, whereas the slow process becomes almost negligible at higher concentrations of GdmCl. It is therefore likely, that the fast phase with $k_c^0 = (1.5 \pm 0.1) \cdot 10^7 \text{ s}^{-1}$ is mainly related to contact formation in unfolded **D/A-Fhp-2**. This is supported by the sensitivity of the reaction towards GdmCl (Figure 4.13B), which is with $m_c = (343 \pm 55) \text{ J mol}^{-1} \text{ M}^{-1}$ similar to that found for contact formation in unstructured model peptides.¹³⁵ The contact formation process in **D/A-Fhp-2** has an apparent activation energy of 28 - 31 kJ mol^{-1} (not corrected for the change in solvent viscosity with temperature), both in the absence of denaturant and in 4 M GdmCl (data

not shown). For contact formation in unstructured peptides similar barrier heights were observed.^{133,134} The small amplitude of the slow phase does not change upon denaturation of the hairpin (Figure 4.13B), which makes it unlikely that it is related to slow conformational transitions, as it has been the case in the $i, i + 6$ labelled α -helical peptides (see 10.A and 10.B). The origin of this phase is presently unclear.

The fast phase, which we assigned to contact formation in unfolded **D/A-Fhp-2** is considerably slower than contact formation in unstructured model peptides. In poly(Ser) peptides contact formation between labels separated by eight peptide bonds is three-fold faster with a rate constant of $k_c^0 = 5.3 \cdot 10^7 \text{ s}^{-1}$.¹⁰⁶ End-to-end contact formation in Ser₄-Gly-Ser₄, which is similar to **D/A-Fhp-2**, having glycine between the labels, $k_c^0 = 4.8 \cdot 10^7 \text{ s}^{-1}$ was measured.¹³³

A somewhat reduced rate constant of contact formation $k_c^0 = 3.2 \cdot 10^7 \text{ s}^{-1}$ has been observed in the isolated six residue loop region of the GB1 hairpin.¹³⁴ There the slower contact formation was attributed to the presence of large side chains. Also in that case a major part of the Xan triplet decay occurred within the dead-time of the experiment. Structural data on the donor/acceptor-labelled GB1 loop peptide are not available. A loss of amplitude due to fast contact formation processes was also observed in a number of unstructured peptides.¹³⁷ Rate constants on the order of $1.2 \cdot 10^7 \text{ s}^{-1}$ have been observed for $i, i + 6$ contact formation in peptides consisting mainly of alanine (see 10.A), but this can be in part attributed to the presence of tails. We cannot exclude that the slow contact formation in **D/A-Fhp-2** originates from local structures opposing contact formation. Since the m -value is only somewhat lower than in unstructured model peptides,^{135,136} these structures would be largely insensitive to GdmCl.

A more detailed interpretation of the TTET data, including the slow process, is complicated by the observed fluorescence in the early time region and the change in the total amplitude. Apparently the nanosecond dead-time of the experiment is not sufficient to capture the major part of the contact formation process in structured **D/A-Fhp-2**. Furthermore, it is likely that the efficient contact formation arises from contact formation between the labels in the folded state. In order to assess the kinetics of hairpin folding and unfolding, it is however necessary to avoid instantaneous contact formation in the folded state. For this purpose, alternative variants of the foldon hairpin or other β -hairpins have to be investigated in which the labels are placed at more distant positions.

4.2.3 Material and Methods

Solid-phase peptide synthesis. The peptides were synthesised on an Applied Biosystems 433A synthesiser, using 9-fluorenylmethyloxycarbonyl (Fmoc) chemistry. Couplings were performed with 2-(1H-benzotriazole-1-yl)-*N,N,N',N'*-tetramethyl-uronium hexafluorophosphate and *N*-hydroxy-benzotriazole (HBTU/HOBt) on Tentagel S RAM or preloaded Tentagel S PHB resin (Rapp Polymere). Backbone protected glycine building blocks (Fmoc-Gly(FmocHmb)-OH, Novabiochem) were used to prevent aspartimide formation at Asp-Gly sequences. Nal was incorporated via Fmoc-protected 1-naphthylalanine (Bachem).

The xanthone derivative 9-oxoxanthene-2-carboxylic acid was synthesised according to Graham and Lewis,²⁵² activated with benzotriazol-1-yl-oxytripyrrolidinophosphonium hexafluorophosphate (PyBOP) and coupled to the β -amino group of selectively deprotected α,β -diaminopropionic acid (Dpr). Methyltrityl was used as orthogonal side chain protection group, which was selectively removed with 3 % (v/v) trifluoro acetic acid (TFA) in dichloromethane.

Final cleavage from the resin and deprotection of the side chains was achieved with 94/2/2/2 (v/v/v/v) TFA / triisopropylsilane / phenole / H₂O. The peptides were purified by preparative HPLC on a RP-8 column. The identity was verified by MALDI mass spectrometry and the purity exceeded 95 % according to analytical HPLC.

Concentration determination and CD measurements. All measurements were performed in 10 mM potassium phosphate, pH 2.0 or pH 7.0. Concentrations were determined via UV absorption spectroscopy, using an extinction coefficient of $\epsilon_{280} = 8250 \text{ M}^{-1} \text{ cm}^{-1}$ for peptides containing tryptophan²⁵³ and of $\epsilon_{343} = 3900 \text{ M}^{-1} \text{ cm}^{-1}$ for peptides containing Xan.¹²⁷ The GdmCl concentrations were calculated from the refractive index.²⁵⁴ CD measurements were performed with an Aviv DS62 spectropolarimeter. Spectra were recorded with $\sim 50 \mu\text{M}$ protein in a 0.1 cm cuvette. Temperature-induced equilibrium unfolding was measured at 228 nm with $7 \mu\text{M}$ peptide in a 1 cm cuvette.

NMR measurements. NMR measurements were performed on samples of 700 - 900 μM **H-Fhp-OH** or **Ac-Fhp-NH₂** at pH 2.0 or pH 7.0 in 10 mM potassium phosphate, containing 0.02 % (w/v) NaN₃ and 5 % (v/v) D₂O. 2,2-Dimethyl-2-silapentane-5-sulfonic acid (DSS) was added as internal standard.

Proton NMR spectra were recorded on a Bruker Avance 600 MHz spectrometer, processed with XWINNMR 2.6 (Bruker) and evaluated with Sparky 3.112

(T. D. Goddard and D. G. Kneller, UCSF). The temperature was calibrated with methanol and ethylene glycol reference samples according to the instructions given by Bruker.

1D NMR spectra were recorded with jump-and-return water suppression.²⁵⁵ Here 1024 FIDs with 8192 complex data points were recorded and multiplied with a phase shifted, squared sine window function before Fourier transformation.

TOCSY spectra were acquired with a Watergate sequence for solvent suppression and the MLEV-17 sequence was applied during the mixing time of 110 ms.²⁵⁶ In indirect dimension 600 and in acquisition 3072 complex data points were recorded (16 FIDs), zero-filled and multiplied with phase shifted, squared sine bell window functions in both dimensions.

NOESY spectra were acquired with a Watergate sequence for solvent suppression and a mixing time of 350 ms.²⁵⁷ In indirect dimension 512 and in acquisition 1024 complex data points were recorded (64 FIDs) and multiplied with phase shifted, squared sine bell functions in both dimensions.

TTET measurements. TTET was measured with an Applied Photophysics LKS.60 Laser Flash Photolysis Reaction Analyzer. Xan triplet states were generated with a 4 ns laser pulse at 355 nm (Nd:YAG, Quantel Brilliant) and the absorbance bands of the Xan and Nal triplet states were monitored at 590 nm and 420 nm, respectively. The measurements were performed in degassed solutions with $\sim 50 \mu\text{M}$ peptide. At this concentration TTET arising from intermolecular contact formation is slower than the intrinsic lifetime of the Xan triplet state. For every experiment four Xan triplet decays were averaged and a sum of two exponential functions was used to describe the data ($A_1, \lambda_1; A_2, \lambda_2$). The signal intensity was corrected for small differences in the peptide concentration. All fitting procedures were performed with ProFit (QuantumSoft) using non-linear least-square methods.

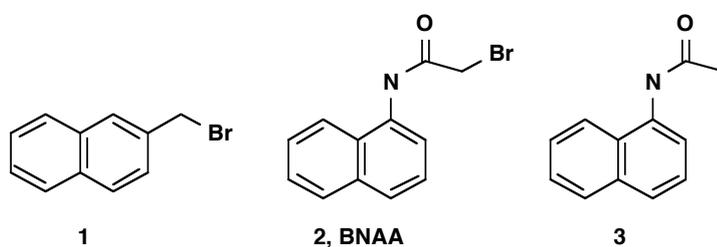
4.3 Incorporation of Triplet Donor and Acceptor Labels via Cysteines

TTET has been intensively used to characterise contact formation dynamics in unfolded polypeptide chains,^{106,113,131-138} α -helices (see 3.1 and 4.1) and in the denatured and native state of a small protein (see 3.4). The peptides for these studies were prepared by solid-phase peptide synthesis, facilitating the specific attachment of donor and acceptor groups. An extension of TTET measurements to larger systems is highly desirable, e.g. to investigate protein association reactions, native state dynamics or the dynamic properties of folding intermediates. Due to the size limitations in solid-phase peptide synthesis alternative labelling strategies have to be developed that allow the specific incorporation of two labels, triplet donor and acceptor, into proteins. Particularly native chemical ligation techniques, which make it possible to combine expressed or synthesised fragments to complete proteins, should be helpful in this context.^{258,259}

In principal, several amino acid side chain functionalities are suited for labelling. Most often the thiol group of cysteine is used, due its low abundance and the unique nucleophilic reactivity of its thiolate form ($pK_a \sim 8.7$). Up to now, no cysteine reactive naphthalene and xanthone derivatives have been described that are compatible with TTET measurements and are suited for labelling in aqueous solution.

4.3.1 Incorporation of Naphthalene

We first envisaged 2-(bromomethyl)-naphthalene **1** as reactive label to attach a naphthalene moiety to the nucleophilic thiol group of cysteine. Since the electronic structure is similar to that of routinely used 1-naphthyl acetic acid and 1-naphthylalanine, the photophysical properties of the product are likely to be the same. Another advantage of compound **1** is the short linker formed between naphthalene and the cysteine side chain.



However, coupling experiments revealed serious problems. Compound **1** is not well soluble in organic solvents that allow solubilising peptides (DMF, NMP, DMSO) and hardly soluble in aqueous solution, which is the solvent of choice for protein

modifications. No reaction with a cysteine containing model peptide or with Ac-Cys-OH could be observed in buffer solutions or aqueous mixtures with organic solvents.

The search for alternative compounds that are soluble in water, have high reactivity and produce short linkers was mainly restricted by the photophysical properties and the synthetic access. Many substituted naphthalenes show an unfavourable bathochromic shift of the ground state absorption or the substituents can undergo photo-induced rearrangements. 2-Bromo-*N*-(1-naphthyl)-acetamide **2** (BNAA) was identified as promising compound, since the starting compound 1-naphththylamine is commercially available and the synthesis of compound **2** has been described.²⁶⁰ The absorption maximum of *N*-(1-naphthyl)-acetamide **3**, which serves as reference compound is at 280 nm,²⁶¹ i.e. well separated from the $^1\pi\pi^*$ transition used to excite the triplet donor Xan. The solubility of BNAA **2** in water was found to be > 1.5 mM and the reactivity of alkyl bromides with electron withdrawing substituents in the α -position is generally quite high.²⁶²

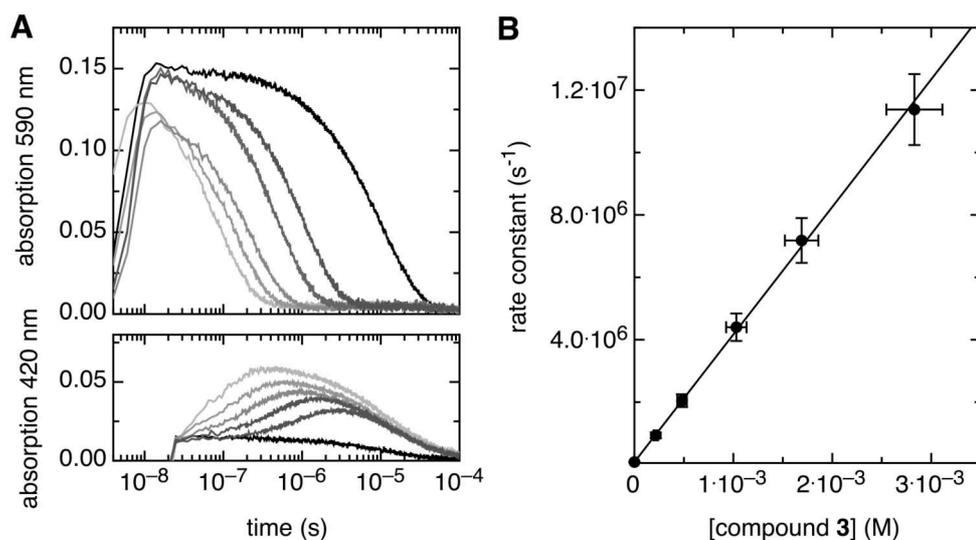


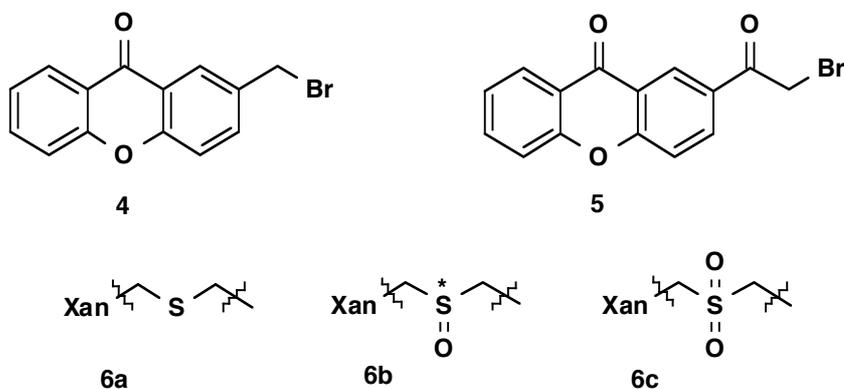
Figure 4.14 Bimolecular TTET experiment with $50 \mu\text{M}$ 9-oxoxanthene-2-carboxylic acid as donor and varying concentrations of the acceptor *N*-(1-naphthyl)-acetamide **3**. (A) Decay of the Xan triplet absorption band at 590 nm (upper panel) and concomitant increase of the naphthalene triplet absorption band at 420 nm (lower panel). Black to grey indicates increasing concentrations of compound **3** (0 – 2.8 mM). (B) Rate constant of the Xan triplet decay under pseudo first-order conditions. The slope corresponds to a bimolecular TTET rate constant of $(4.1 \pm 0.3) \cdot 10^9 \text{ M}^{-1} \text{ s}^{-1}$. The measurements were performed in 10 mM potassium phosphate, pH 7.0 at 22.5°C .

Another prerequisite for the use as acceptor in contact formation experiments is that TTET from xanthone occurs in a diffusion controlled reaction.²⁶³ In order to test for the efficiency of the transfer, a bimolecular TTET experiment with 9-oxoxanthene-2-carboxylic acid and the reference compound **3** was performed under pseudo first-order conditions (Figure 4.14). The experiment gave a bimolecular rate constant of $(4.1 \pm 0.3) \cdot 10^9 \text{ M}^{-1} \text{ s}^{-1}$, identical to the rate constant measured with 1-naphthyl acetic acid as acceptor.¹⁰⁶ Absorption measurements at 420 nm show, that like in alkyl naphthalene derivatives, the triplet transfer can be also followed by the increase of the triplet state absorption band of the acceptor.

Efficient labelling of a cysteine containing peptide with BNAA **2** worked with 1:1 stoichiometry both in organic and aqueous solution (pH 8.5). The effect of the longer linker produced upon coupling of BNAA **2** to a cysteine side chain was addressed in a first study.²⁶⁴ The rate constant of end-to-end contact formation in a (GlySer)₄-peptide labelled with BNAA **2** was found to be only slightly slower than in a peptide with 1-naphthylalanine.²⁶⁴ In summary, BNAA **2** is well suited for incorporating naphthalene as triplet acceptor into proteins. This has been recently exploited to obtain a double-labelled carp parvalbumin variant (108 amino acids) after native chemical ligation.²⁶⁴ Compound **2** might also be useful to introduce naphthalene for FRET measurements.¹⁰⁹

4.3.2 Incorporation of Xanthone

In order to couple the triplet donor xanthone to proteins 2-(bromomethyl)-9-oxoxanthene **4** had been suggested as cysteine-reactive labelling agent.²⁶⁵ The compound was synthesised in three steps according to Eckstein and Marona²⁶⁶ with a minor modification of the bromination procedure.



Labelling of cysteine-containing model peptides with compound **4** was readily achieved in water/DMF mixtures. The water solubility of compound **4** was found to be low and therefore 2-(bromoacetyl)-9-oxoxanthene **5** should be considered as an alternative. Furthermore, measurements on 2-methyl-9-oxoxanthene revealed that the triplet absorption is somewhat weaker as in 9-oxoxanthene-2-carboxylic acid. Using compound **5** should again give a higher signal due to the adjacent carbonyl group. The synthesis of compound **5** via an Ullmann condensation²⁶⁷ turned out to be less favourable as in the case of compound **4**, but an alternative route to 2-acetyl-9-oxoxanthene might be used.²⁶⁸

A different problem was encountered when the photophysical properties of a donor-only labelled peptide prepared with compound **4** were investigated. After excitation of the xanthone moiety at 355 nm the triplet absorption, which has still a broad maximum around 590 nm, is reduced to 50 % compared to free 2-methyl-9-oxoxanthene (Figure 4.15A). The resolved triplet absorption decays with a time constant of 500 ns, whereas the triplet state in free 2-methyl-9-oxoxanthene has a lifetime of $\sim 20 \mu\text{s}$ under these conditions (Figure 4.15A). In the presence of 1 M GdmCl the lifetime of **6a** was even further reduced to around 30 ns.

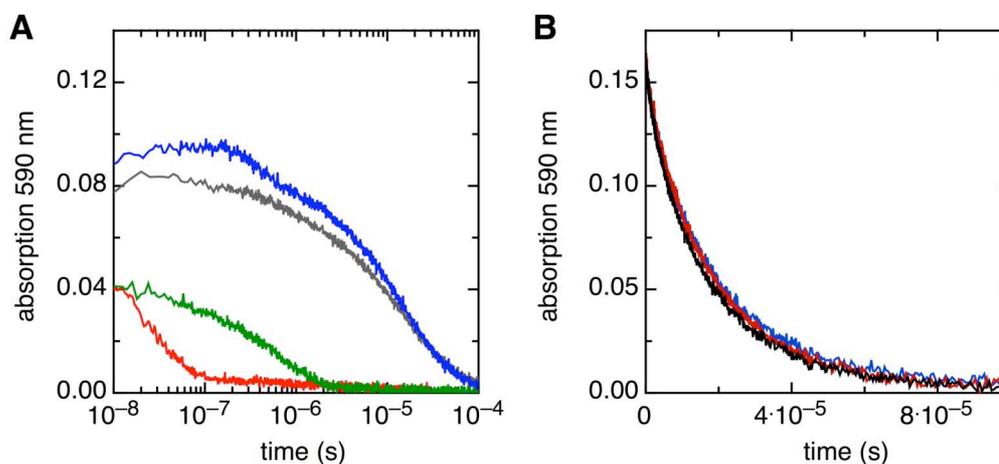


Figure 4.15. Xan triplet lifetime in the presence of thioethers, sulfoxides and sulfones. (A) Triplet decay after coupling **4** to a cysteine-containing peptide, which gives a thioether **6a** in 0 M GdmCl (green) and 1 M GdmCl (red). For comparison the triplet decay of free 2-methyl-9-oxoxanthene is shown (grey), as well as the kinetics after oxidation of the thioether to the sulfone (**6c**, blue). (B) Triplet lifetime of 9-oxoxanthene-2-carboxylic acid alone (black) and in the presence of 1 M dimethylsulfoxide (red) or 1 M dimethylsulfone (blue). The measurements were performed at 22.5 °C with 50 μM of xanthone compounds in degassed buffer, 10 mM potassium phosphate, pH 7.0.

The short triplet lifetime of thioether-linked Xan severely limits its use to investigate contact formation in proteins. The fast decay and reduced triplet yield is likely the consequence of triplet quenching by the nearby thioether moiety **6a**, since also methionine has been found to quench the triplet state of xanthone.¹³² To circumvent quenching of the Xan triplet state, a subsequent oxidation of the thioether **6a** to a sulfoxide **6b** or a sulfone **6c** was considered. Indeed, it was found that neither dimethylsulfoxide nor dimethylsulfone affect the triplet lifetime of 9-oxoxanthene-2-carboxylic acid, even when added in concentrations up to 1 M (Figure 4.15B). Oxidation of the thioether in the labelled peptide **6a** to a sulfone **6c** was achieved using performic acid²⁶⁹ and led to a full recovery of the triplet yield and lifetime (Figure 4.15A). This approach should be also applicable to eliminate triplet quenching after cysteine labelling in proteins.

4.3.3 Material and Methods

Labelling compounds. 2-(Bromomethyl)-naphthalene **1** was purchased from Fluka (No. 17617). 2-Bromo-*N*-(1-naphthyl)-acetamide **2** (BNAA) could be obtained from Scientific Exchange, Inc. (Ossipee, NH, USA; M-696303). Its purity and identity was verified by 1D-NMR and ESI mass spectrometry:

¹H-NMR (CDCl₃), δ (ppm) in CDCl₃: 4.20 (s, 2 H, CH₂Br); 7.48 - 7.61 (m, 3 H), 7.76 (d, 1 H, $J=8.2$ Hz); 7.86 - 7.92 (m, 2H); 7.97 (d, 1H, $J=7.1$ Hz); 8.69 (s, broad, 1 H, HN).

ESI-MS, $[M+H]^+$: $m/z_{\text{calc}}=264.002, 266.000$ (1:1), $m/z_{\text{meas}}=264.000, 265.998$ (1:1).

2-(Bromomethyl)-9-oxoxanthene **4** was synthesised according to Eckstein and Marona²⁶⁶ with minor modifications in three steps, NMR assignments are given according to Pickert and Frahm²⁷⁰:

2-(4'-Methylphenoxy)-benzoic acid. 10.80 g sodium methanolate (0.20 mol) were solved in a mixture of 100 ml dry methanol and 15.66 g 2-chlorobenzoic acid (0.10 mol), 32.4 g *p*-cresole (0.30 mol), 0.3 g Cu (< 5 micron) and 0.1 g Cu₂O were added. 100 ml liquid paraffin were added, the mixture heated to 80 °C until no further evaporation was observed and then stirred for 2 h at 190 – 205 °C. The mixture was cooled to r.t. and 100 ml toluene were added. The solid material was collected and washed with another 150 ml toluene to remove the paraffin. The residuum was dissolved in 200 ml 10 % Na₂CO₃ and the solution purified over charcoal. The product was obtained by fractionated precipitation with 10 % H₂SO₄. Recrystallisation from 70 % ethanol gave 10.1 g pure product (0.044 mol, 44 % yield).

$^1\text{H-NMR}$ (CDCl_3), δ in ppm: 2.38 (s, 3 H, CH_3); 6.82 (dd, 1 H, $J = 8.3, 0.8$ Hz, H3); 7.02 (m, 2 H, H2'); 7.17 – 7.27 (m, 3 H, H5, H3'); 7.46 (ddd, 1 H, $J = 8.3, 7.3, 1.8$ Hz, H4); 8.20 (dd, 1 H, $J = 7.8, 1.8$ Hz, H6).

$^{13}\text{C-NMR}$ (CDCl_3), δ in ppm: 21.2 (CH_3); 118.4 (C3); 120.2, 120.3 (C1, 2 x C2'); 123.6 (C5); 131.0 (2 x C3'); 133.6 (C6); 135.1, 135.0 (C4, C4'); 153.3 (C1'), 158.3 (C2); 168.6 (C=O).

2-(Methyl)-9-oxoxanthene. 4.6 g 2-(4'-Methylphenoxy)-benzoic acid (0.020 mol) were solved in 15 ml H_2SO_4 (conc.), heated for 2 h on a 90 °C water bath and then kept at room temperature for another hour. The mixture was poured into 70 ml ice water and the precipitate was collected. To remove the acidic starting material the precipitate was washed with 10 % Na_2CO_3 solution. Crystallisation from ethanol gave white needles (1.45 g, 0.007 mol, 35 % yield).

UV (EtOH), λ_{max} : 205 nm, 233 nm, 242 nm, 261 nm, 293 nm, 344 nm.

$^1\text{H-NMR}$ (CDCl_3), δ in ppm: 2.45 (s, 3 H, CH_3); 7.32 - 7.39 (m, 2 H, H4, H7); 7.45 (dd, 1 H, $J = 8.6, < 1$ Hz, H5); 7.51 (dd, 1 H, $J = 8.7, 2.0$ Hz, H3); 7.69 (m, 1 H, H6); 8.11 (d, 1 H, $J = 2.0$ Hz, H1); 8.30 (dd, 1 H, $J = 8.1, 1.8$ Hz, H8).

$^{13}\text{C-NMR}$ (CDCl_3), δ in ppm: 21.2 (CH_3); 118.1, 118.3 (C4, C5); 121.8, 122.2 (C8a, C8b); 124.1, 126.4, 127.1 (C1, C7, C8); 134.0 (C2); 135.0, 136.4 (C3, C6); 154.8, 156.6 (C4a, C4b); 177.7 (C=O).

2-(Bromomethyl)-9-oxoxanthene 4. 0.55 g 2-(methyl)-9-oxoxanthene (2.6 mmol), 0.41g *N*-Bromosuccinimide (NBS, 2.3 mmol) and 1 mg azobisisobutyronitrile (AIBN, 0.16 mmol) were solved in 10 ml CCl_4 . NBS had been dried over P_4O_{10} , CCl_4 over molecular sieve (4A). The mixture was slowly heated and then refluxed for 12 h on a water bath. The warm suspension was filtered to remove precipitated succinimide and then kept for 1 h at 5 °C. The precipitate was collected and recrystallised from ethanol. Preparative RP-HPLC was used to separate unbrominated starting material from product (110 mg, 0.4 mmol, 15 % yield).

EI-MS, $[\text{M}]^+$: $m/z_{\text{calc}} = 287.98, 289.98$ (1:1), $m/z_{\text{meas}} = 287.0, 289.0$ (1:1).

$^1\text{H-NMR}$ (CDCl_3), δ in ppm: 4.60 (s, 2H, CH_2Br); 7.37 - 7.42 (m, 1H); 7.47 - 7.52 (m, 2H); 7.71 - 7.79 (m, 2H); 8.31 - 8.35 (m, 2H).

$^{13}\text{C-NMR}$ (CDCl_3), δ in ppm: 32.7 (CH_2Br); 118.4, 119.3 (C4, C5); 122.1, 122.2 (C8a, C8b); 124.6, 127.2, 127.3 (C1, C7, C8); 134.1 (C2); 135.4, 136.1 (C3, C6); 156.3, 156.5 (C4a, C4b); 177.2 (C=O).

The first two steps of the 2-(bromoacetyl)-9-oxoxanthene **5** synthesis were performed according to Re *et al.*²⁶⁷ but with low yield due to problems with the Ullmann reaction.

Peptide synthesis and labelling reactions. For labelling experiments and donor-only measurements the peptide Ac-RSGCGSGRGFG-OH was chosen, containing a single cysteine and a phenylalanine for spectrometric detection. It was synthesised by Fmoc solid-phase peptide synthesis as described in section 4.2.3. Alternatively Ac-Cys-OH (Sigma A-8199) was used as model compound for labelling experiments. Preparative labelling of fully deprotected peptide was performed with a slight excess of compound **2** or **4** in 1/1 (v/v) DMF/buffer mixtures or in pure buffer (100 mM Tris-HCl, pH 8.4).

Performic acid oxidation. Performic acid for the thioether oxidation²⁶⁹ was generated by 1 h pre-incubation of a mixture of 50 μl H_2O_2 and 950 μl formic acid at 4 °C. 7 mg labelled peptide, dissolved in 500 μl formic acid, were added. After 1 h incubation at 4 °C the mixture was frozen in liquid N_2 , lyophilized and purified by RP-HPLC. Analysis of the reaction mixture by RP-HPLC and subsequent MALDI-MS spectroscopy indicated complete conversion of the starting material to the sulfone with formation of some side product. The observed mass shift of $\Delta m = +28$ indicates that it probably arises from formylation.

Time-resolved absorption spectroscopy. Bimolecular TTET experiments were performed under pseudo first-order conditions with 50 μM 9-oxoxanthene-2-carboxylic acid as triplet donor, synthesised according to Graham and Lewis.²⁵² The concentration of the acceptor *N*-(1-naphthyl)acetamide **3** (Acros 347050050) was determined spectroscopically ($\epsilon_{280} = 6800 \text{ M}^{-1} \text{ cm}^{-1}$,²⁶¹) and varied between 500 μM and 3 mM. Dimethylsulfoxide (Fluka 41641) and dimethylsulfone (Fluka 41631) were added in concentrations up to 1 M to test for Xan triplet quenching. All triplet measurements were performed at 22.5 °C in 10 mM potassium phosphate, pH 7.0 in degassed solution using an Applied Photophysics LKS.60 Laser Flash Photolysis Reaction Analyzer (see Section 4.2.3).

5 Summary

The aim of this thesis was to obtain deeper insight into the dynamics of structure formation in peptides and proteins with emphasis on the early stages of protein folding. On the route to the native state many conformations have to be sampled, marginally stable secondary structures may form, as well as intermediates that are often only populated to a low extent. These conformational dynamics were investigated in suitable model systems.

Dynamics in peptides and partially folded proteins are fast. For an adequate description, new approaches are required that provide the necessary time resolution and, quite important, allow a microscopic interpretation of the events.

Measurement of contact formation between triplet donor and acceptor groups by triplet-triplet energy transfer (TTET) has previously been used to characterise loop formation dynamics in unstructured polypeptide chains. In this work, we extended TTET measurements to structured systems. The aim was on the one hand to characterise their chain dynamics and on the other hand to use contact formation as irreversible probing reaction to study fast conformational transitions in equilibrium. This gave valuable new insights into structural dynamics on the hundreds of nanosecond timescale.

We used TTET coupled to the helix-coil transition to systematically investigate the local dynamics in α -helical peptides. For this purpose, triplet donor and acceptor groups were introduced at positions i and $i + 6$ in a number of alanine-based model helices. Contact formation in the coil state is with about 100 ns similarly fast as local helix folding and unfolding, which allows to obtain microscopic rate constants of these processes. By placing the labels at different positions along a 21 residue helical peptide, the local dynamics at different positions were investigated. Local folding occurs with time constants of $1/k_f \approx 400$ ns, almost independently of the position. In contrast, local unfolding $1/k_u$ shows a strong position dependence, accelerating from 1.4 μ s in the helix centre up to 250 ns at the termini. The low equilibrium stability at the ends, also known as helix fraying, is hence a consequence of fast unfolding at the termini.

Since local folding and unfolding are well-defined events, it is possible to compare the experimental results to theoretical descriptions of the helix-coil transition. As experiments of this type had not been discussed so far, we modeled our experimental

system in kinetic Ising simulations using a Monte-Carlo approach. These simulations fully reproduced the observed position dependence and showed that local folding and unfolding in our helices are dominated by helix elongation and unwinding at the boundaries. Nucleation processes and complete loss of helical structures appear to be of minor relevance under these conditions. From a quantitative comparison we can estimate that the addition of a segment to an existing helix occurs with a time constant of about 50 ns.

Urea has a strong denaturing effect on helices. We found that this results from a strong deceleration of local helix folding upon addition of urea. Temperature accelerates local unfolding, whereas folding shows non-Arrhenius behaviour, first becoming faster, then slower with rising temperatures. Our comparison with simulations suggests that this behaviour has to be the consequence of high activation energies of about 40 kJ mol⁻¹ of both basic steps, helix elongation and unwinding. Data from a number of peptides of different length show that the dynamics at the N-terminus of helices are independent of the peptide length. In the centre, however, unfolding becomes considerably slower with increasing peptide length. These data and further experiments should significantly contribute to understand the formation and fluctuations of α -helices.

Besides α -helices we wanted to investigate the dynamics of hairpin formation using the same approach. The central β -hairpin of the foldon domain from bacteriophage T4 fibrin appears to be well suited for that purpose. We synthesised peptides comprising the hairpin region and could show by nuclear magnetic resonance (NMR) spectroscopy that a 12mer indeed adopts the native-like type I' hairpin conformation. First experiments to find suitable positions for the triplet donor and acceptor groups were conducted.

Photoswitchable groups that impose conformational restraints are promising tools to induce conformational transitions. Thioamides, which have a photosensitive *cis-trans* equilibrium and are nearly isosteric to amide bonds, might be excellent candidates to trigger substantial helix folding and unfolding reactions. It was however unclear, whether backbone thioamide bonds are at all compatible with α -helix formation. We therefore synthesised alanine-based model helices with single amide-to-thioamide substitutions and characterised them by circular dichroism (CD) and NMR spectroscopy. We found that single thioamide substitutions have a strong destabilising effect of about 7 kJ mol⁻¹, but that they can be integrated into the backbone of regular α -helices.

It would be highly desirable to get access to folding intermediates under physiological conditions in order to characterise their structure and dynamics. The foldon domain of bacteriophage T4 fibrin is optimised for fast trimerisation and forms an intermediate on the stage of the monomer. This monomer assembles with high efficiency to the dimer and trimer. By disrupting an intersubunit salt-bridge through a single amino acid substitution (E5R) we were able to obtain a model of the monomeric folding intermediate for equilibrium studies. Optical spectroscopy confirmed that the foldon variant E5R resembles the kinetic folding intermediate and a structural model could be obtained by NMR spectroscopy. We found that the foldon monomer is folded into a compact, native-like structure in which the N-terminal segment forms a hydrophobic cluster with the central hairpin. The preorganisation of the monomer explains why foldon can associate so efficiently to the dimer and trimer. The dynamics in this naturally occurring folding intermediate will be addressed in future studies.

Several proteins have been reported to fold on the timescale of a few microseconds, among them the villin headpiece subdomain (HP35). The fast folding of this model protein, extensively studied in experiments and simulations, might be a consequence of residual structure in the unfolded state. We introduced triplet donor and acceptor groups at four different positions in order to investigate the chain dynamics of HP35. This allowed us to obtain for the first time absolute rate constants of contact formation in a small protein. In the denatured state contact formation between the different positions occurs with time constants of $1/k_c \approx 100 - 200$ ns, which is similar to loop formation in unstructured model peptides. Even under conditions strongly favouring the native state, the measured dynamics give no evidence for residual structure or a native-like topology in the unfolded state of HP35. At low concentrations of denaturant the native state becomes populated and the observed kinetics are in full agreement with the two-state equilibrium transition obtained by CD measurements. Global unfolding of the native state has to be slower than $10 \mu\text{s}$, but we detected local unfolding events at the C-terminal helix.

In order to extend TTET measurements to larger proteins it is necessary to introduce triplet donor and acceptor groups into expressed proteins. For this purpose we developed cysteine-reactive derivatives of naphthalene and xanthone and invented a method to eliminate triplet quenching by thioethers.

6 References

1. Creighton, T. E. (1992). *Proteins: Structures and Molecular Properties*. 2nd edit, W.H. Freeman, New York, USA.
2. Anfinsen, C. B. (1973). Principles that govern the folding of protein chains. *Science* **181**, 223-230.
3. Young, J. C., Agashe, V. R., Siegers, K. & Hartl, F. U. (2004). Pathways of chaperone-mediated protein folding in the cytosol. *Nat. Rev. Mol. Cell. Biol.* **5**, 781-791.
4. Messens, J. & Collet, J.-F. (2006). Pathways of disulfide bond formation in *Escherichia coli*. *Int. J. Biochem. & Cell Biol.* **38**, 1050-1062.
5. Lang, K., Schmid, F. X. & Fischer, G. (1987). Catalysis of protein folding by prolyl isomerase. *Nature* **329**, 268-270.
6. Lehn, J.-M. (2002). Toward self-organization and complex matter. *Science* **295**, 2400.
7. Pain, R. H. (2000). *Mechanisms of Protein Folding*. 2nd edit, Oxford University Press, Oxford, United Kingdom.
8. Finkelstein, A. V. & Ptitsyn, O. B. (2002). *Protein Physics: A Course of Lectures*. Academic Press, San Diego, USA.
9. Buchner, J. & Kiefhaber, T., Eds. (2005). *Protein Folding Handbook*. Wiley-VCH, Weinheim, Germany.
10. Dyson, H. J. & Wright, P. E. (2005). Intrinsically unstructured proteins and their functions. *Nat. Rev. Mol. Cell. Biol.* **6**, 197-208.
11. *RCSB Protein Data Bank*.
12. Perutz, M. F., Rossmann, M. G., Cullis, A. F., Muirhead, H., Will, G. & North, A. C. T. (1960). Structure of haemoglobin. A three-dimensional Fourier synthesis obtained by X-ray analysis. *Nature* **185**, 416-422.
13. Kendrew, J. C., Dickerson, R. E., Strandberg, B. E., Hart, R. G., Davies, D. R., Phillips, D. C. & Shore, V. C. (1960). Structure of myoglobin: a three-dimensional Fourier synthesis at 2 Å resolution. *Nature* **185**, 422-427.
14. Wüthrich, K. (1986). *NMR of Proteins and Nucleic Acids*, J. Wiley, New York, USA.
15. Schellman, J. A. (1955). The stability of hydrogen-bonded peptide structures in aqueous solution. *C.R. Trav. Lab Carlsberg Ser. Chim.* **29**, 230-259.
16. Kauzmann. (1959). Some factors in the interpretation of protein denaturation. *Adv. Protein Chem.* **14**, 1-63.
17. Israelachvili, J. N. (1991). *Intermolecular Forces and Surface Forces*. Academic Press, London, United Kingdom.
18. Matthews, B. W. (1995). Studies on protein stability with T4 lysozyme. *Adv. Protein Chem.* **46**, 249-278.
19. Stickle, D. F., Presta, L. G., Dill, K. A. & Rose, G. D. (1992). Hydrogen bonding in globular proteins. *J. Mol. Biol.* **226**, 1143-1159.
20. Perl, D., Mueller, U., Heinemann, U. & Schmid, F. X. (2000). Two exposed amino acid residues confer thermostability on a cold shock protein. *Nat. Struct. Biol.* **7**, 380-383.
21. Anderson, D. E., Bechtel, W. J. & Dahlquist, F. W. (1990). pH-Induced denaturation of proteins: a single salt bridge contributes 3-5 kcal/mol to the free energy of folding of T4 lysozyme. *Biochemistry* **29**, 2403-2408.

22. Baldwin, R. L. (2007). Energetics of protein folding. *J. Mol. Biol.* **371**, 283-301.
23. Modig, K., Liepinsh, E., Otting, G. & Halle, B. (2004). Dynamics of protein and peptide hydration. *J. Am. Chem. Soc.* **126**, 102-114.
24. Eckert, B., Martin, A., Balbach, J. & Schmid, F. X. (2005). Prolyl isomerization as a molecular timer in phage infection. *Nat. Struct. Mol. Biol.* **12**, 619-623.
25. Yankeelov, J. A. & Koshland, D. E. (1965). Evidence for conformation changes induced by substrates of phosphoglucosmutase. *J. Biol. Chem.* **240**, 1593-1602.
26. Ansari, A., Jones, C. M., Henry, E. R., Hofrichter, J. & Eaton, W. A. (1992). The role of solvent viscosity in the dynamics of protein conformational changes. *Science* **256**, 1796-1798.
27. Doster, W., Cusack, S. & Petry, W. (1989). Dynamical transition of myoglobin revealed by inelastic neutron scattering. *Nature* **337**, 754-756.
28. Frauenfelder, H., Petsko, G. A. & Tsernoglou, D. (1979). Temperature-dependent X-ray diffraction as a probe of protein structural dynamics. *Nature* **280**, 558-563.
29. Yang, H., Luo, G., Karnchanaphanurach, P., Louie, T. M., Rech, I., Cova, S., Xun, L. & Xie, X. S. (2003). Protein conformational dynamics probed by single-molecule electron transfer. *Science* **302**, 262-266.
30. Cavanagh, J., Fairbrother, W. J., Palmer, A. G. I., Rance, M. & Skelton, N. J. (2007). *Protein NMR Spectroscopy - Principles and Practice*. 2nd edit, Elsevier Academic Press, San Diego, USA.
31. Eisenmesser, E. Z., Millet, O., Labeikovsky, W., Korzhnev, D. M., Wolf-Watz, M., Bosco, D. A., Skalicky, J. J., Kay, L. E. & Kern, D. (2005). Intrinsic dynamics of an enzyme underlies catalysis. *Nature* **438**, 117-121.
32. Boehr, D. D., McElheny, D., Dyson, H. J. & Wright, P. E. (2006). The dynamic energy landscape of dihydrofolate reductase catalysis. *Science* **313**, 1638-1642.
33. Anson, M. L. & Mirsky, A. E. (1931). The reversibility of protein coagulation. *J. Phys. Chem.* **35**, 185-193.
34. Wu, H. (1931). Studies on denaturation of proteins. XIII. A theory of denaturation. *Chin. J. Physiol.* **5**, 321-344.
35. Neurath, H., Greenstein, J. P., Putnam, F. W. & Erickson, J. O. (1944). The chemistry of protein denaturation. *Chem. Rev.* **34**, 157-265.
36. Tanford, C. (1968,1970). Protein denaturation. *Adv. Protein Chem.* **23**, **24**, 122-282, 1-95.
37. Jackson, S. E. (1998). How do small single-domain proteins fold? *Fold. & Des.* **3**, R81-91.
38. Privalov, P. L. (1979). Stability of proteins. Small globular proteins. *Adv. Protein Chem.* **33**, 167-241.
39. Schellman, J. A. (1987). Selective binding and solvent denaturation. *Biopolymers* **25**, 549-559.
40. Greene, R. F. & Pace, C. N. (1974). Urea and guanidine hydrochloride denaturation of ribonuclease, lysozyme, α -chymotrypsin and β -lactoglobulin. *J. Biol. Chem.* **249**, 5388-5393.
41. 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.
42. Lumry, R., Biltonen, R. & Brandts, J. F. (1966). Validity of the "two-state" hypothesis for conformational transitions of proteins. *Biopolymers* **4**, 917-944.
43. Levinthal, C. (1969). How to fold graciously. In *Mössbauer Spectroscopy in Biological Systems (Conference Proceedings)*, Vol. 67, pp. 22-24. Univ. of Illinois Bulletin, Monticello, Ill, USA.

44. Zwanzig, R., Szabo, A. & Bagchi, B. (1992). Levinthal's paradox. *Proc. Natl. Acad. Sci. USA* **89**, 20-22.
45. Wolynes, P. G., Onuchic, J. N. & Thirumalai, D. (1995). Navigating the folding routes. *Science* **267**, 1619-1620.
46. Kubelka, J., Hofrichter, J. & Eaton, W. A. (2004). The protein folding 'speed limit'. *Curr. Opin. Struct. Biol.* **14**, 76-88.
47. Kubelka, J., Chiu, T. K., Davies, D. R., Eaton, W. A. & Hofrichter, J. (2006). Sub-microsecond protein folding. *J. Mol. Biol.* **359**, 546-553.
48. Dyer, R. B. (2007). Ultrafast and downhill protein folding. *Curr. Opin. Struct. Biol.* **17**, 38-47.
49. Jaenicke, R. (1987). Folding and association of proteins. *Prog. Biophys. Mol. Biol.* **49**, 117-237.
50. Anfinsen, C. B., Haber, E., Sela, M. & White, F. H. (1961). The kinetics of formation of native ribonuclease during oxidation of the reduced polypeptide chain. *Proc. Natl. Acad. Sci. USA* **47**, 1309-1314.
51. Creighton, T. E. (1997). Protein folding coupled to disulphide bond formation. *Biol. Chem.* **378**, 731-744.
52. 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.
53. Balbach, J. & Schmid, F. X. (2000). Proline isomerization and its catalysis in protein folding. In *Mechanisms of Protein Folding*, pp. 212-249. Oxford University Press, Oxford, United Kingdom.
54. Zwanzig, R. (1997). Two-state models of protein folding kinetics. *Proc. Natl. Acad. Sci. USA* **94**, 148-150.
55. Jackson, S. E. & Fersht, A. R. (1991). Folding of chymotrypsin inhibitor 2. 1. Evidence for a two-state transition. *Biochemistry* **30**, 10428-10435.
56. Bachmann, A., Segel, D. & Kiefhaber, T. (2002). Test for cooperativity in the early kinetic intermediate in lysozyme folding. *Biophys. Chem.* **96**, 141-151.
57. Ikai, A. (1971). *The Kinetics of the Unfolding and Refolding of Proteins*. PhD thesis, Duke University.
58. Bachmann, A. & Kiefhaber, T. (2005). Kinetic mechanisms in protein folding. In *Handbook of Protein Folding* (Buchner, J. & Kiefhaber, T., eds.). Wiley-VCH, Weinheim, Germany.
59. Schmid, F. X. (1983). Mechanism of folding of ribonuclease A. Slow refolding is a sequential reaction via structural intermediates. *Biochemistry* **22**, 4690-4696.
60. Wildegger, G. & Kiefhaber, T. (1997). Three-state model for lysozyme folding: triangular folding mechanism with an energetically trapped intermediate. *J. Mol. Biol.* **270**, 294-304.
61. Ikai, A. & Tanford, C. (1971). Kinetic evidence for incorrectly folded intermediate states in the refolding of denatured proteins. *Nature* **230**, 100-102.
62. Sanchez, I. E. & Kiefhaber, T. (2003). Evidence for sequential barriers and obligatory intermediates in apparent two-state protein folding. *J. Mol. Biol.* **325**, 367-376.
63. Kiefhaber, T., Labhardt, A. M. & Baldwin, R. L. (1995). Direct NMR evidence for an intermediate preceding the rate-limiting step in the unfolding of ribonuclease A. *Nature* **375**, 513-515.
64. Eyring, H. (1935). The activated complex in chemical reactions. *J. Chem. Phys.* **3**, 107-115.

65. Leffler, J. E. (1953). Parameters for the description of transition states. *Science* **117**, 340-341.
66. Pappenberger, G., Saudan, C., Becker, M., Merbach, A. E. & Kiefhaber, T. (2000). Denaturant-induced movement of the transition state of protein folding revealed by high-pressure stopped-flow measurements. *Proc. Natl. Acad. Sci. USA* **97**, 17-22.
67. Pohl, F. M. (1976). Temperature-dependence of the folding kinetics of folding of chymotrypsinogen A. *FEBS Letters* **65**, 293-296.
68. Schönbrunner, N., Pappenberger, G., Scharf, M., Engels, J. & Kiefhaber, T. (1997). Effect of preformed correct tertiary interactions on rapid two-state tendamistat folding: evidence for hairpins as initiation sites for β -sheet formation. *Biochemistry* **36**, 9057-9065.
69. Schätzle, M. & Kiefhaber, T. (2006). Shape of the free energy barriers for protein folding probed by multiple perturbation analysis. *J. Mol. Biol.* **357**, 655-664.
70. Fersht, A. R. (1999). *Structure and Mechanism in Protein Science. A Guide to Enzyme Catalysis and Protein Folding*. Palgrave Macmillan, Basingstoke, United Kingdom.
71. Sanchez, 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.
72. Kramers, H. A. (1940). Brownian motion in a field of force and the diffusion model of chemical reactions. *Physica* **4**, 284-304.
73. Haenggi, P., Talkner, P. & Borkovec, M. (1990). Reaction-rate theory: fifty years after Kramers. *Rev. Mod. Phys.* **62**, 251-341.
74. Jacob, M., Schindler, T., Balbach, J. & Schmid, F. X. (1997). Diffusion control in an elementary protein folding reaction. *Proc. Natl. Acad. Sci. USA* **94**, 5622-5627.
75. Chrnyk, B. A. & Matthews, C. R. (1990). Role of diffusion in the folding of the alpha subunit of tryptophan synthase from *Escherichia coli*. *Biochemistry* **29**, 2149-2154.
76. Wetlaufer, D. B. (1973). Nucleation, rapid folding, and globular intrachain regions in proteins. *Proc. Natl. Acad. Sci. USA* **70**, 697-701.
77. Karplus, M. & Weaver, D. L. (1976). Protein-folding dynamics. *Nature* **260**, 404-406.
78. Burton, R. E., Myers, J. K. & Oas, T. G. (1998). Protein folding dynamics: quantitative comparison between theory and experiment. *Biochemistry* **37**, 5337-5343.
79. Fersht, A. R. (1995). Optimization of rates of protein folding: the nucleation-condensation mechanism and its implications. *Proc. Natl. Acad. Sci. USA* **92**, 10869-10873.
80. Ptitsyn, O. B. (1995). Molten globule and protein folding. *Adv. Protein Chem.* **47**, 83-229.
81. Sali, A., Shakhnovich, E. & Karplus, M. (1994). How does a protein fold? *Nature* **369**, 248-251.
82. Brown, J. E. & Klee, W. A. (1971). Helix-coil transition of the isolated amino terminus of ribonuclease. *Biochemistry* **10**, 470-476.
83. Blanco, F., Jiménez, M. A., Rico, M., Santoro, J., Herranz, J. & Nieto, J. L. (1991). Tendamistat (12-26) fragment. NMR characterisation of isolated β -turn folding intermediates. *Eur. J. Biochem.* **200**, 345-351.
84. de Prat Gay, G. & Fersht, A. R. (1994). Generation of a family of protein fragments for structure-folding studies. 1. Folding complementation of two

- fragments of chymotrypsin inhibitor-2 formed by cleavage at its unique methionine residue. *Biochemistry* **33**, 7957-7963.
85. Alexandrescu, A. T., Abeygunawardana, C. & Shortle, D. (1994). Structure and dynamics of a denatured 131-residue fragment of staphylococcal nuclease: a heteronuclear NMR study. *Biochemistry* **33**, 1063-1072.
 86. Tang, Y., Rigotti, D. J., Fairman, R. & Raleigh, D. P. (2004). Peptide models provide evidence for significant structure in the denatured state of a rapidly folding protein: the villin headpiece subdomain. *Biochemistry* **43**, 3264-3272.
 87. Religa, T. L., Markson, J. S., Mayor, U., Freund, S. M. & Fersht, A. R. (2005). Solution structure of a protein denatured state and folding intermediate. *Nature* **437**, 1053-1056.
 88. Ramachandran, G. N., Ramakrishnan, C. & Sasisekharan, V. (1963). Stereochemistry of polypeptide chain configurations. *J. Mol. Biol.* **7**, 95-99.
 89. Miller, W. G., Brant, D. A. & Flory, P. J. (1967). Random coil configurations of polypeptide copolymers. *J. Mol. Biol.* **23**, 67-80.
 90. Flory, P. J. (1969). *Statistical Mechanics of Chain Molecules*. Hanser, Munich, Germany.
 91. Fitzkee, N. C. & Rose, G. D. (2005). Sterics and solvation window accessible conformational space for unfolded proteins. *J. Mol. Biol.* **353**, 873-887.
 92. Shi, Z., Chen, K., Liu, Z. & Kallenbach, N. R. (2006). Conformation of the backbone in unfolded proteins. *Chem. Rev.* **106**, 1877-1897.
 93. Chen, K., Liu, Z. & Kallenbach, N. R. (2004). The polyproline II conformation in short alanine peptides is noncooperative. *Proc. Natl. Acad. Sci. USA* **101**, 15352-15357.
 94. Damaschun, G., Damaschun, H., Gast, K. & Zirwer, D. (1998). Denatured states of yeast phosphoglycerate kinase. *Biochemistry (Moscow)* **63**, 259-275.
 95. des Gennes, P.-G. (1979). *Scaling Concepts in Polymer Physics*. Cornell University Press, Ithaca, USA.
 96. Chan, H. S. & Dill, K. A. (1991). Polymer principles in protein structure and stability. *Annu. Rev. Biophys. Biophys. Chem.* **20**, 447-490.
 97. Kohn, J. E., Millett, 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. USA* **101**, 12491-12496.
 98. Lumb, K. J. & Kim, P. S. (1994). Formation of a hydrophobic cluster in denatured bovine pancreatic trypsin inhibitor. *J. Mol. Biol.* **236**, 412-420.
 99. 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.
 100. Wirmer, J., Schlörb, C., Klein-Seetharaman, J., Hirano, R., Ueda, T., Imoto, T. & Schwalbe, H. (2004). Modulation of compactness and long-range interactions of unfolded lysozyme by single point mutations. *Angew. Chem. Int. Ed. Engl.* **43**, 5780-5785.
 101. Meier, S., Strohmeier, M., Blackledge, M. & Grzesiek, S. (2007). Direct observation of dipolar couplings and hydrogen bonds across a β -hairpin in 8 M urea. *J. Am. Chem. Soc.* **129**, 754-755.
 102. Arai, M. & Kuwajima, K. (2000). Role of the molten globule state in protein folding. *Adv. Protein Chem.* **53**, 209-282.

103. Jennings, P. A. & Wright, P. E. (1993). Formation of a molten globule intermediate early in the kinetic folding pathway of apomyoglobin. *Science* **262**, 892-896.
104. Sauder, J. M. & Roder, H. (1998). Amide protection in an early folding intermediate of cytochrome *c*. *Fold. & Des.* **3**, 293-301.
105. Cho, J. H., Sato, S. & Raleigh, D. P. (2004). Thermodynamics and kinetics of non-native interactions in protein folding: a single point mutant significantly stabilizes the N-terminal domain of L9 by modulating non-native interactions in the denatured state. *J. Mol. Biol.* **338**, 827-837.
106. 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.
107. Haas, E., Katchalski-Katzir, E. & Steinberg, I. Z. (1978). Brownian motion of the ends of oligopeptide chains in solution as estimated by energy transfer between the chain ends. *Biopolymers* **17**, 11-31.
108. Lakowicz, J. R., Kusba, J., Szmajcinski, H., Gryczynski, I., Eis, P. S., Wiczak, W. & Johnson, M. L. (1991). Resolution of end-to-end diffusion coefficients and distance distributions of flexible molecules using fluorescent donor-acceptor and donor-quencher pairs. *Biopolymers* **31**, 1363-1378.
109. Möglich, A., Joder, K. & Kiefhaber, T. (2006). End-to-end distance distributions and intrachain diffusion constants in unfolded polypeptide chains indicate intramolecular hydrogen bond formation. *Proc. Natl. Acad. Sci. USA* **103**, 12394-12399.
110. Nettels, D., Gopich, I. V., Hoffmann, A. & Schuler, B. (2007). Ultrafast dynamics of protein collapse from single-molecule photon statistics. *Proc. Natl. Acad. Sci. USA* **104**, 2655-2660.
111. 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.
112. 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. USA* **100**, 3838-3840.
113. Fierz, B. & Kiefhaber, T. (2005). Dynamics of unfolded polypeptide chains. In *Handbook of Protein Folding* (Buchner, J. & Kiefhaber, T., eds.), pp. 805-851. Wiley-VCH, Weinheim, Germany.
114. 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.
115. Lapidus, L. J., Eaton, W. A. & Hofrichter, J. (2002). Measuring dynamic flexibility of the coil state of a helix-forming peptide. *J. Mol. Biol.* **319**, 19-25.
116. 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.
117. Buscaglia, M., Lapidus, L. J., Eaton, W. A. & Hofrichter, J. (2006). Effects of denaturants on the dynamics of loop formation in polypeptides. *Biophys. J.* **91**, 276-288.
118. Buscaglia, M., Schuler, B., Lapidus, L. J., Eaton, W. A. & Hofrichter, J. (2003). Kinetics of intramolecular contact formation in a denatured protein. *J. Mol. Biol.* **332**, 9-12.

119. Buscaglia, M., Kubelka, J., Eaton, W. A. & Hofrichter, J. (2005). Determination of ultrafast protein folding rates from loop formation dynamics. *J. Mol. Biol.* **347**, 657-664.
120. Singh, V. R., Kopka, M., Chen, Y., Wedemeyer, W. J. & Lapidus, L. J. (2007). Dynamic similarity of the unfolded states of proteins L and G. *Biochemistry* **46**, 10046-10054.
121. Huang, C. Y., Getahun, Z., Zhu, Y., Klemke, J. W., DeGrado, W. F. & Gai, F. (2002). Helix formation via conformation diffusion search. *Proc. Natl. Acad. Sci. USA* **99**, 2788-2793.
122. Neuweiler, H., Schulz, A., Bohmer, M., Enderlein, J. & Sauer, M. (2003). Measurement of submicrosecond intramolecular contact formation in peptides at the single-molecule level. *J. Am. Chem. Soc.* **125**, 5324-5330.
123. Neuweiler, H., Lollmann, M., Doose, S. & Sauer, M. (2007). Dynamics of unfolded polypeptide chains in crowded environment studied by fluorescence correlation spectroscopy. *J. Mol. Biol.* **365**, 856-869.
124. Turro, N. J. (1991). *Modern Molecular Photochemistry*. University Science Books, Sausalito, USA.
125. Satzger, H., Schmidt, B., Root, C., Zinth, W., Fierz, B., Krieger, F., Kiefhaber, T. & Gilch, P. (2004). Ultrafast quenching of the xanthone triplet by energy transfer: new insight into the intersystem crossing kinetics. *J. Phys. Chem. A* **108**, 10072-10079.
126. Heinz, B., Schmidt, B., Root, C., Satzger, H., Milota, F., Fierz, B., Kiefhaber, T., Zinth, W. & Gilch, P. (2006). On the unusual fluorescence properties of xanthone in water. *Phys. Chem. Chem. Phys.* **8**, 3432-3439.
127. Murov, S. L., Carmichael, I. & Hug, G. L. (1993). *Handbook of Photochemistry*, Marcel Dekker Inc., New York, USA.
128. Dexter, D. L. (1953). A theory of sensitized luminescence in solids. *J. Chem. Phys.* **21**, 836-850.
129. Förster, T. (1948). Zwischenmolekulare Energiewanderung und Fluoreszenz. *Ann Phys.* **2**, 55-75.
130. Lakowicz, J. R. (2006). *Principles of Fluorescence Spectroscopy*. 3rd edit, Springer, Berlin, Germany.
131. 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.
132. 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.
133. Krieger, F., Möglich, A. & Kiefhaber, T. (2005). Effect of proline and glycine residues on dynamics and barriers of loop formation in polypeptide chains. *J. Am. Chem. Soc.* **127**, 3346-3352.
134. Krieger, F. (2004). *Dynamics in Unfolded Polypeptide Chains as Model for Elementary Steps in Protein Folding*. PhD thesis, Universität Basel.
135. Möglich, A., Krieger, F. & Kiefhaber, T. (2005). Molecular basis for the effect of urea and guanidinium chloride on the dynamics of unfolded polypeptide chains. *J. Mol. Biol.* **345**, 153-162.
136. Möglich, A. (2005). *Structure and Dynamics of Unfolded Polypeptide Chains*. PhD thesis, Universität Basel.

137. Fierz, B., Satzger, H., Root, C., Gilch, P., Zinth, W. & Kiefhaber, T. (2007). Loop formation in unfolded polypeptide chains on the picoseconds to microseconds time scale. *Proc. Natl. Acad. Sci. USA* **104**, 2163-2168.
138. Fierz, B. & Kiefhaber, T. (2007). End-to-end vs. interior loop formation kinetics in unfolded polypeptide chains. *J. Am. Chem. Soc.* **129**, 672-679.
139. Szabo, A., Schulten, K. & Schulten, Z. (1980). First passage time approach to diffusion controlled reactions. *J. Chem. Phys.* **72**, 4350-4357.
140. Jacobson, H. & Stockmayer, W. H. (1950). Intramolecular reaction in polycondensations. I. The theory of linear systems. *J. Chem. Phys.* **18**, 1600-1606.
141. 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.
142. Pauling, L., Corey, R. B. & Branson, H. R. (1951). The structure of proteins: two hydrogen-bonded helical configurations of the polypeptide chain. *Proc. Natl. Acad. Sci. USA* **37**, 205-211.
143. Barlow, D. J. & Thornton, J. M. (1988). Helix geometry in proteins. *J. Mol. Biol.* **201**, 601-619.
144. Waltho, J. P., Feher, V. A., Lerner, R. A. & Wright, P. E. (1989). Conformation of a T cell stimulating peptide in aqueous solution. *FEBS Lett.* **250**, 400-404.
145. Kammerer, R. A., Schulthess, T., Landwehr, R., Lustig, A., Engel, J., Aebi, U. & Steinmetz, M. O. (1998). An autonomous folding unit mediates the assembly of two-stranded coiled coils. *Proc. Natl. Acad. Sci. USA* **95**, 13419-13424.
146. Raschke, T. M. & Marqusee, S. (1997). The kinetic folding intermediate of ribonuclease H resembles the acidic molten globule and partially unfolded molecules detected under native conditions. *Nat. Struct. Biol.* **4**, 298-304.
147. Marqusee, S. & Baldwin, R. L. (1987). Helix stabilization by Glu...Lys⁺ salt bridges in short peptides of *de novo* design. *Proc. Natl. Acad. Sci. USA* **84**, 8898-8902.
148. Marqusee, S., Robbins, V. H. & Baldwin, R. L. (1989). Unusually stable helix formation in short alanine-based peptides. *Proc. Natl. Acad. Sci. USA* **86**, 5286-5290.
149. Koradi, R., Billeter, M. & Wüthrich, K. (1996). MOLMOL: a program for display and analysis of macromolecular structures. *J. Mol. Graph.* **14**, 51-5, 29-32.
150. Doig, A. J., MacArthur, M. W., Stapley, B. J. & Thornton, J. M. (1997). Structures of N-termini of helices in proteins. *Protein Sci.* **6**, 147-155.
151. Doig, A. J., Errington, N. & Iqbalsyah, T. M. (2005). Stability and design of α -helices. In *Protein Folding Handbook* (Buchner, J. & Kiefhaber, T., eds.), pp. 247-313. Wiley-VCH, Weinheim, Germany.
152. Bolin, K. A. & Millhauser, G. L. (1999). α and 3_{10} : the split personality of polypeptide helices. *Acc. Chem. Res.* **32**, 1027-1033.
153. Poland, D. & Scheraga, H. A. (1970). *Theory of Helix-Coil Transitions in Biopolymers*. Academic Press, New York, USA.
154. Zimm, B. H., Doty, P. & Iso, K. (1959). Determination of the parameters for helix formation in poly-gamma-benzyl-L-glutamate. *Proc. Natl. Acad. Sci. USA* **45**, 1601-1607.
155. Zimm, B. H. & Bragg, J. K. (1959). Theory of the phase transition between helix and random coil in polypeptide chains. *J. Chem. Phys.* **31**, 526-535.
156. Qian, H. & Schellman, J. A. (1992). Helix-coil theories: a comparative study for finite length polypeptides. *J. Phys. Chem.* **96**, 3987-3994.

157. Lifson, S. & Roig, A. (1961). On the theory of helix-coil transition in polypeptides. *J. Chem. Phys.* **34**, 1963-1974.
158. Hill, T. L. (1960). *An Introduction to Statistical Thermodynamics*. Dover Publications, New York, USA.
159. Lyu, P. C., Marky, L. A. & Kallenbach, N. R. (1989). The role of ion pairs in α -helix stability: two new designed helical peptides. *J. Am. Chem. Soc.* **111**, 2733-2734.
160. Bradley, E. K., Thomason, J. F., Cohen, F. E., Kosen, P. A. & Kuntz, I. D. (1990). Studies of synthetic helical peptides using circular dichroism and nuclear magnetic resonance. *J. Mol. Biol.* **215**, 607-622.
161. Scholtz, J. M., Qian, H., York, E. J., Stewart, J. M. & Baldwin, R. L. (1991). Parameters of helix-coil transition theory for alanine-based peptides of varying chain lengths in water. *Biopolymers* **31**, 1463-1470.
162. Chakrabartty, A., Schellman, J. A. & Baldwin, R. L. (1991). Large differences in the helix propensities of alanine and glycine. *Nature* **351**, 586-588.
163. Rohl, C. A. & Baldwin, R. L. (1994). Exchange kinetics of individual amide protons in ^{15}N -labeled helical peptides measured by isotope-edited NMR. *Biochemistry* **33**, 7760-7767.
164. 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.
165. 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.
166. Doig, A. J. & Baldwin, R. L. (1995). N- and C-capping preferences for all 20 amino acids in α -helical peptides. *Protein Sci.* **4**, 1325-1336.
167. Smith, J. S. & Scholtz, J. M. (1996). Guanidine hydrochloride unfolding of peptide helices: separation of denaturant and salt effects. *Biochemistry* **35**, 7292-7297.
168. Scholtz, J. M., Marqusee, S., Baldwin, R. L., York, E. J., Stewart, J. M., Santoro, M. & Bolen, D. W. (1991). Calorimetric determination of the enthalpy change for the α -helix to coil transition of an alanine peptide in water. *Proc. Natl. Acad. Sci. USA* **88**, 2854-2858.
169. Wieprecht, T., Apostolov, O., Beyermann, M. & Seelig, J. (1999). Thermodynamics of the α -helix-coil transition of amphipathic peptides in a membrane environment: implications for the peptide-membrane binding equilibrium. *J. Mol. Biol.* **294**, 785-794.
170. Lopez, M. M., Chin, D. H., Baldwin, R. L. & Makhatadze, G. I. (2002). The enthalpy of the alanine peptide helix measured by isothermal titration calorimetry using metal-binding to induce helix formation. *Proc. Natl. Acad. Sci. USA* **99**, 1298-1302.
171. Wójcik, J., Altmann, K. H. & Scheraga, H. A. (1990). Helix-coil stability constants for the naturally occurring amino acids in water. XXIV. Half-cystine parameters from random poly(hydroxybutylglutamine-*S*-methylthio-L-cysteine). *Biopolymers* **30**, 121-134.
172. Myers, J. K., Pace, C. N. & Scholtz, J. M. (1997). A direct comparison of helix propensity in proteins and peptides. *Proc. Natl. Acad. Sci. USA* **94**, 2833-2837.
173. Muñoz, V. & L., S. (1994). Elucidating the folding problem of helical peptides using empirical parameters. *Nat. Struct. Biol.* **1**, 399-409.

174. Miick, S. M., Casteel, K. M. & Millhauser, G. L. (1993). Experimental molecular dynamics of an alanine-based helical peptide determined by spin label electron spin resonance. *Biochemistry* **32**, 8014-8021.
175. Venyaminov, S. Y., Hedstrom, J. F. & Prendergast, F. G. (2001). Analysis of the segmental stability of helical peptides by isotope-edited infrared spectroscopy. *Proteins* **45**, 81-89.
176. Liff, M. I., Lyu, P. C. & Kallenbach, N. R. (1991). Analysis of asymmetry in the distribution of helical residues in peptides by proton nuclear magnetic resonance. *J. Am. Chem. Soc.* **113**, 1014-1019.
177. Jaravine, V. A., Alexandrescu, A. T. & Grzesiek, S. (2001). Observation of the closing of individual hydrogen bonds during TFE-induced helix formation in a peptide. *Protein Sci.* **10**, 943-950.
178. Eigen, M. (1967). *Immeasurably fast reactions*. Nobel Lecture.
179. 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.
180. Gruenewald, B., Nicola, C. U., Lustig, A. & Schwarz, G. (1979). Kinetics of the helix-coil transition of a polypeptide with non-ionic side groups, derived from ultrasonic relaxation experiments. *Biophys. Chem.* **9**, 137-147.
181. Schwarz, G. (1965). On the kinetics of the helix-coil transition of polypeptides in solution. *J. Mol. Biol.* **11**, 64-77.
182. Schwarz, G. (1968). General theoretical approach to the thermodynamic and kinetic properties of cooperative intramolecular transformations of linear biopolymers. *Biopolymers* **6**, 873-897.
183. Thompson, P. A., Eaton, W. A. & 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.
184. Thompson, P. A., Munoz, V., Jas, G. S., Henry, E. R., Eaton, W. A. & Hofrichter, J. (2000). The helix-coil kinetics of a heteropeptide. *J. Phys. Chem. B* **104**, 378-389.
185. Lednev, I. K., Karnoup, A. S., Sparrow, M. C. & Asher, S. A. (1999). Nanosecond UV resonance Raman examination of initial steps in α -helix secondary structure evolution. *J. Am. Chem. Soc.* **121**, 4076-4077.
186. Williams, S., Causgrove, T. P., Gilmanishin, R., Fang, K. S., Callender, R. H., Woodruff, W. H. & Dyer, R. B. (1996). Fast events in protein folding: helix melting and formation in a small peptide. *Biochemistry* **35**, 691-697.
187. Wang, T., Zhu, Y., Getahun, Z., Du, D., Huang, C. Y., DeGrado, W. F. & Gai, F. (2004). Length dependent helix-coil transition kinetics of nine alanine-based peptides. *J. Phys. Chem. B* **108**, 15301-15310.
188. Werner, J. H., Dyer, R. B., Fesinmeyer, R. M. & Andersen, N. H. (2002). Dynamics of the primary processes of protein folding: helix nucleation. *J. Phys. Chem. B* **106**, 487-494.
189. Pozo Ramajo, A., Petty, S. A., Starzyk, A., Decatur, S. M. & Volk, M. (2005). The α -helix folds more rapidly at the C-terminus than at the N-terminus. *J. Am. Chem. Soc.* **127**, 13784-13785.
190. Nesmelova, I., Krushelnitsky, A., Idiyatullin, D., Blanco, F., Ramirez-Alvarado, M., Daragan, V. A., Serrano, L. & Mayo, K. H. (2001). Conformational exchange on the microsecond time scale in α -helix and β -hairpin peptides measured by ^{13}C NMR transverse relaxation. *Biochemistry* **40**, 2844-2853.

191. Bredenbeck, J., Helbing, J., Kumita, J. R., Woolley, G. A. & Hamm, P. (2005). α -Helix formation in a photoswitchable peptide tracked from picoseconds to microseconds by time-resolved IR spectroscopy. *Proc. Natl. Acad. Sci. USA* **102**, 2379-2384.
192. Nelson, R., Sawaya, M. R., Balbirnie, M., Madsen, A. O., Riek, C., Grothe, R. & Eisenberg, D. (2005). Structure of the cross- β spine of amyloid-like fibrils. *Nature* **435**, 773-778.
193. Sibanda, B. L., Blundell, T. L. & Thornton, J. M. (1989). Conformation of β -hairpins in protein structures. A systematic classification with applications to modelling by homology, electron density fitting and protein engineering. *J. Mol. Biol.* **206**, 759-777.
194. McCallister, E. L., Alm, E. & Baker, D. (2000). Critical role of β -hairpin formation in protein G folding. *Nat. Struct. Biol.* **7**, 669-673.
195. Grantcharova, V. P., Riddle, D. S. & Baker, D. (2000). Long-range order in the src SH3 folding transition state. *Proc. Natl. Acad. Sci. USA* **97**, 7084-7089.
196. Deechongkit, S., Nguyen, H., Powers, E. T., Dawson, P. E., Gruebele, M. & Kelly, J. W. (2004). Context-dependent contributions of backbone hydrogen bonding to β -sheet folding energetics. *Nature* **430**, 101-105.
197. Forge, V., Hoshino, M., Kuwata, K., Arai, M., Kuwajima, K., Batt, C. A. & Goto, Y. (2000). Is folding of β -lactoglobulin non-hierarchical? Intermediate with native-like β -sheet and non-native α -helix. *J. Mol. Biol.* **296**, 1039-1051.
198. Walkenhorst, W. F., Edwards, J. A., Markley, J. L. & Roder, H. (2002). Early formation of a β -hairpin during folding of staphylococcal nuclease H124L as detected by pulsed hydrogen exchange. *Protein Sci.* **11**, 82-91.
199. Meier, S., Güthe, S., Kiefhaber, T. & Grzesiek, S. (2004). Foldon, the natural trimerization domain of T4 fibritin, dissociates into a monomeric A-state form containing a stable β -hairpin: atomic details of trimer dissociation and local β -hairpin stability from residual dipolar couplings. *J. Mol. Biol.* **344**, 1051-1069.
200. 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.
201. Cox, J. P., Evans, P. A., Packman, L. C., Williams, D. H. & Woolfson, D. N. (1993). Dissecting the structure of a partially folded protein. Circular dichroism and nuclear magnetic resonance studies of peptides from ubiquitin. *J. Mol. Biol.* **234**, 483-492.
202. Schätzle, M. (2005). *Properties of the Free Energy Barriers for Folding of the α -Amylase Inhibitor Tendamistat*. PhD thesis, Universität Basel.
203. Searle, M. S., Zerella, R., Williams, D. H. & Packman, L. C. (1996). Native-like β -hairpin structure in an isolated fragment from ferredoxin: NMR and CD studies of solvent effects on the N-terminal 20 residues. *Protein Eng.* **9**, 559-665.
204. Zerella, R., Evans, P. A., Ionides, J. M., Packman, L. C., Trotter, B. W., Mackay, J. P. & Williams, D. H. (1999). Autonomous folding of a peptide corresponding to the N-terminal β -hairpin from ubiquitin. *Protein Sci.* **8**, 1320-1331.
205. de Alba, E., Jiménez, M. A., Rico, M. & Nieto, J. L. (1996). Conformational investigation of designed short linear peptides able to fold into β -hairpin structures in aqueous solution. *Fold. & Des.* **1**, 133-144.
206. de Alba, E., Rico, M. & Jiménez, M. A. (1997). Cross-strand side-chain interactions versus turn conformation in β -hairpins. *Protein Sci.* **6**, 2548-2560.
207. Cochran, A. G., Skelton, N. J. & Starovasnik, M. A. (2001). Tryptophan zippers: stable, monomeric β -hairpins. *Proc. Natl. Acad. Sci. USA* **98**, 5578-5583.

208. Searle, M. S. (2005). Design and stability of peptide β -sheets. In *Handbook of Protein Folding* (Buchner, J. & Kiefhaber, T., eds.). Wiley-VCH, Weinheim, Germany.
209. Haque, T. S. & Gellman, S. H. (1997). Insights on β -hairpin stability in aqueous solution from peptides with enforced type I' and type II' β -turns. *J. Am. Chem. Soc.* **119**, 2303-2304.
210. Ramirez-Alvarado, M., Blanco, F. J. & Serrano, L. (1996). *De novo* design and structural analysis of a model β -hairpin peptide system. *Nat. Struct. Biol.* **3**, 604-612.
211. Ramirez-Alvarado, M., Blanco, F. J., Niemann, H. & Serrano, L. (1997). Role of β -turn residues in β -hairpin formation and stability in designed peptides. *J. Mol. Biol.* **273**, 898-912.
212. Griffiths-Jones, S. R., Sharman, G. J., Maynard, A. J. & Searle, M. S. (1998). Modulation of intrinsic phi, psi propensities of amino acids by neighbouring residues in the coil regions of protein structures: NMR analysis and dissection of a β -hairpin peptide. *J. Mol. Biol.* **284**, 1597-1609.
213. Ciani, B., Jourdan, M. & Searle, M. S. (2003). Stabilization of β -hairpin peptides by salt bridges: role of preorganization in the energetic contribution of weak interactions. *J. Am. Chem. Soc.* **125**, 9038-9047.
214. Espinosa, J. F., Munoz, V. & Gellman, S. H. (2001). Interplay between hydrophobic cluster and loop propensity in β -hairpin formation. *J. Mol. Biol.* **306**, 397-402.
215. Muñoz, V., Thompson, P. A., Hofrichter, J. & Eaton, W. A. (1997). Folding dynamics and mechanism of β -hairpin formation. *Nature* **390**, 196-199.
216. Du, D., Zhu, Y., Huang, C. Y. & Gai, F. (2004). Understanding the key factors that control the rate of β -hairpin folding. *Proc. Natl. Acad. Sci. USA* **101**, 15915-15920.
217. Dyer, R. B., Maness, S. J., Peterson, E. S., Franzen, S., Fesinmeyer, R. M. & Andersen, N. H. (2004). The mechanism of β -hairpin formation. *Biochemistry* **43**, 11560-11566.
218. Olsen, K. A., Fesinmeyer, R. M., Stewart, J. M. & Andersen, N. H. (2005). Hairpin folding rates reflect mutations within and remote from the turn region. *Proc. Natl. Acad. Sci. USA* **102**, 15483-15487.
219. Chen, R. P., Huang, J. J., Chen, H. L., Jan, H., Velusamy, M., Lee, C. T., Fann, W., Larsen, R. W. & Chan, S. I. (2004). Measuring the refolding of β -sheets with different turn sequences on a nanosecond time scale. *Proc. Natl. Acad. Sci. USA* **101**, 7305-7310.
220. Schrader, T. E., Schreier, W. J., Cordes, T., Koller, F. O., Babitzki, G., Denschlag, R., Renner, C., Löweneck, M., Dong, S.-L., Moroder, L., Tavan, P. & Zinth, W. (2007). Light-triggered β -hairpin folding and unfolding. *Proc. Natl. Acad. Sci. USA* **104**, 15729-15734.
221. Tao, Y., Strelkov, S. V., Mesyanzhinov, V. V. & Rossmann, M. G. (1997). Structure of bacteriophage T4 fibritin: a segmented coiled coil and the role of the C-terminal domain. *Structure* **5**, 789-798.
222. Kostyuchenko, V. A., Chipman, P. R., Leiman, P. G., Arisaka, F., Mesyanzhinov, V. V. & Rossmann, M. G. (2005). The tail structure of bacteriophage T4 and its mechanism of contraction. *Nat. Struct. Mol. Biol.* **12**, 810-813.
223. Letarov, A. V., Londer, Y. Y., Boudko, S. P. & Mesyanzhinov, V. V. (1999). The carboxy-terminal domain initiates trimerization of bacteriophage T4 fibritin. *Biochemistry (Moscow)* **64**, 817-823.

224. Frank, S., Kammerer, R. A., Mechling, D., Schulthess, T., Landwehr, R., Bann, J., Guo, Y., Lustig, A., Bächinger, H. P. & Engel, J. (2001). Stabilization of short collagen-like triple helices by protein engineering. *J. Mol. Biol.* **308**, 1081-1089.
225. Yang, X., Lee, J., Mahony, E. M., Kwong, P. D., Wyatt, R. & Sodroski, J. (2002). Highly stable trimers formed by human immunodeficiency virus type 1 envelope glycoproteins fused with the trimeric motif of T4 bacteriophage fibrin. *J. Virol.* **76**, 4634-4642.
226. Güthe, S., Kapinos, L., Möglich, A., Meier, S., Grzesiek, S. & Kiefhaber, T. (2004). Very fast folding and association of a trimerization domain from bacteriophage T4 fibrin. *J. Mol. Biol.* **337**, 905-915.
227. Milla, M. E. & Sauer, R. T. (1994). P22 Arc repressor: folding kinetics of a single-domain, dimeric protein. *Biochemistry* **33**, 1125-1133.
228. McKnight, C. J., Doering, D. S., Matsudaira, P. T. & Kim, P. S. (1996). A thermostable 35-residue subdomain within villin headpiece. *J. Mol. Biol.* **260**, 126-134.
229. McKnight, C. J., Matsudaira, P. T. & Kim, P. S. (1997). NMR structure of the 35-residue villin headpiece subdomain. *Nat. Struct. Biol.* **4**, 180-184.
230. Chiu, T. K., Kubelka, J., Herbst-Irmer, R., Eaton, W. A., Hofrichter, J. & Davies, D. R. (2005). High-resolution X-ray crystal structures of the villin headpiece subdomain, an ultrafast folding protein. *Proc. Natl. Acad. Sci. USA* **102**, 7517-7522.
231. Frank, B. S., Vardar, D., Buckley, D. A. & McKnight, C. J. (2002). The role of aromatic residues in the hydrophobic core of the villin headpiece subdomain. *Protein Sci.* **11**, 680-687.
232. Kubelka, J., Eaton, W. A. & Hofrichter, J. (2003). Experimental tests of villin subdomain folding simulations. *J. Mol. Biol.* **329**, 625-630.
233. Brewer, S. H., Vu, D. M., Tang, Y., Li, Y., Franzen, S., Raleigh, D. P. & Dyer, R. B. (2005). Effect of modulating unfolded state structure on the folding kinetics of the villin headpiece subdomain. *Proc. Natl. Acad. Sci. USA* **102**, 16662-16667.
234. Wang, M., Tang, Y., Sato, S., Vugmeyster, L., McKnight, C. J. & Raleigh, D. P. (2003). Dynamic NMR line-shape analysis demonstrates that the villin headpiece subdomain folds on the microsecond time scale. *J. Am. Chem. Soc.* **125**, 6032-6033.
235. Zagrovic, B., Snow, C. D., Khaliq, S., Shirts, M. R. & Pande, V. S. (2002). Native-like mean structure in the unfolded ensemble of small proteins. *J. Mol. Biol.* **323**, 153-164.
236. Tang, Y., Goger, M. J. & Raleigh, D. P. (2006). NMR characterization of a peptide model provides evidence for significant structure in the unfolded state of the villin headpiece helical subdomain. *Biochemistry* **45**, 6940-6946.
237. Duan, Y. & Kollman, P. A. (1998). Pathways to a protein folding intermediate observed in a 1-microsecond simulation in aqueous solution. *Science* **282**, 740-744.
238. Lei, H., Wu, C., Liu, H. & Duan, Y. (2007). Folding free-energy landscape of villin headpiece subdomain from molecular dynamics simulations. *Proc. Natl. Acad. Sci. USA* **104**, 4925-4930.
239. Satzger, H., Root, C., Gilch, P., Zinth, W., Wildemann, D. & Fischer, G. (2005). Photoswitchable elements within a peptide backbone - ultrafast spectroscopy of thioxylated amides. *J. Phys. Chem. B* **109**, 4770-4775.
240. Helbing, J., Bregy, H., Bredenbeck, J., Pfister, R., Hamm, P., Huber, R., Wachtveitl, J., DeVico, L. & Olivucci, M. (2004). A fast photoswitch for

- minimally perturbed peptides: investigation of the *trans* → *cis* photoisomerization of *N*-methylthioacetamide. *J. Am. Chem. Soc.* **126**, 8823-8834.
241. Miwa, J. H., Pallivathucal, L., Gowda, S. & Lee, K. E. (2002). Conformational stability of helical peptides containing a thioamide linkage. *Org. Lett.* **4**, 4655-4657.
 242. Luo, P. & Baldwin, R. L. (1997). Mechanism of helix induction by trifluoroethanol: a framework for extrapolating the helix-forming properties of peptides from trifluoroethanol/water mixtures back to water. *Biochemistry* **36**, 8413-8421.
 243. Buchete, N.-V. & Straub, J. E. (2000). Mean first-passage time calculations for the coil-to-helix transition: the active helix Ising model. *J. Phys. Chem. B* **105**, 6684-6697.
 244. Woody, R. W. & Dunker, A. K. (1996). Aromatic and cysteine side-chain circular dichroism in proteins. In *Circular Dichroism and the Conformational Analysis of Biomolecules* (Fasman, G. D., ed.). Plenum Press, New York, USA.
 245. Wishart, D. S., Sykes, B. D. & Richards, F. M. (1991). Relationship between nuclear magnetic resonance chemical shift and protein secondary structure. *J. Mol. Biol.* **222**, 311-333.
 246. Fesinmeyer, R. M., Hudson, F. M., Olsen, K. A., White, G. W. N., Euser, A. & Andersen, N. H. (2005). Chemical shifts provide folded populations and register of β -hairpins and β -sheets. *J. Biomol. NMR* **33**, 213-231.
 247. Wagner, G., Pardi, A. & Wüthrich, K. (1983). Hydrogen bond length and ^1H NMR chemical shifts in proteins. *J. Am. Chem. Soc.* **105**, 5948-5949.
 248. Schwarzing, S., Kroon, G. J. A., Foss, T. R., Wright, P. E. & Dyson, H. J. (2000). Random coil chemical shifts in acidic 8 M urea: implementation of random coil chemical shifts. *J. Biomolec. NMR* **18**, 43-48.
 249. Schwarzing, S., Kroon, G. J. A., Foss, T. R., Chung, J., Wright, P. E. & Dyson, H. J. (2001). Sequence-dependent correction of random coil NMR chemical shifts. *J. Am. Chem. Soc.* **123**, 2970-2978.
 250. Wüthrich, K., Billeter, M. & Braun, W. (1984). Polypeptide secondary structure determination by nuclear magnetic resonance observation of short proton-proton distances. *J. Mol. Biol.* **180**, 715-740.
 251. de Alba, E., Jiménez, M. A. & Rico, M. (1997). Turn residue sequence determines β -hairpin conformation in designed peptides. *J. Am. Chem. Soc.* **119**, 175-183.
 252. Graham, R. & Lewis, J. R. (1978). Synthesis of 9-oxoxanthen-2-carboxylic acids. *J. Chem. Soc. Perkin. Trans. 1*, 876-881.
 253. Gill, S. C. & von Hippel, P. H. (1989). Calculation of protein extinction coefficients from amino acid sequence data. *Anal. Biochem.* **182**, 319-326.
 254. Pace, C. N. (1986). Determination and analysis of urea and guanidine hydrochloride denaturation curves. *Methods Enzymol.* **131**, 266-280.
 255. Hore, P. J. (1983). Solvent suppression in Fourier transform NMR. *J. Magn. Res.* **55**, 283-300.
 256. Bax, A. & Davis, D. G. (1985). MLEV-17 based two dimensional homonuclear magnetization transfer spectroscopy. *J. Magn. Res.* **65**, 355-360.
 257. Piotto, M., Saudek, V. & Sklenář, V. (1992). Gradient-tailored excitation for single-quantum NMR spectroscopy of aqueous solutions. *J. Biomolec. NMR* **2**, 661-666.
 258. Muir, T. W. (2003). Semisynthesis of proteins by expressed protein ligation. *Annu. Rev. Biochem.* **72**, 249-289.

259. David, R., Richter, M. P. & Beck-Sickinger, A. G. (2004). Expressed protein ligation. Method and applications. *Eur. J. Biochem.* **271**, 663-677.
260. Hopsu, V. K. & Glenner, G. G. (1964). Characterization of enzymes hydroxylising acyl naphthylamides: I. Mono -and dihalogen derivatives. *J. Histochem. Cytochem.* **12**, 674-686.
261. Rosenberry, T. L. & Bernhard, S. A. (1972). Catalysis by acetylcholinesterase. Synergistic effects of inhibitors during the hydrolysis of acetic acid esters. *Biochemistry* **11**, 4308-4321.
262. Dahl, K. H. & McKinley-McKee, J. S. (1981). The reactivity of affinity labels: a kinetic study of the reaction of alkyl halides with thiolate anions - a model reaction for protein alkylation. *Bioorg. Chem.* **10**, 329-341.
263. Fierz, B. (2006). *Dynamics of Unfolded and α -Helical Polypeptide Chains*. PhD thesis, Universität Basel.
264. Nyffenegger, C. (2007). *Dynamics in Different States of Parvalbumin Measured by Triplet-Triplet Energy Transfer*. Master thesis, Universität Basel.
265. Fierz, B. (2002). *Dynamics of Intramolecular Contact Formation in Model Peptides*. Diploma thesis, Universität Basel.
266. Eckstein, M. & Marona, H. (1980). Preparation and properties of some 2-mercaptomethylxanthenes. *Polish J. Chem.* **54**, 1281-1285.
267. Re, P. D., Valenti, P., Borraccini, A. & Primofiore, G. P. (1972). β -Adrenergic blocking agents of the chromone and xanthone groups. *J. Med. Chem.* **15**, 198-199.
268. Zhao, J. & Larock, R. C. (2005). One-pot synthesis of xanthenes and thioxanthenes by the tandem coupling-cyclization of arynes and salicylates. *Org. Lett.* **7**, 4273-4275.
269. Hirs, C. H. (1967). Performic acid oxidation. *Methods Enzymol.* **11**, 197-199.
270. Pickert, M. & Frahm, A. W. (1998). Substituted xanthenes as antimycobacterial agents, part 1: synthesis and assignment of $^1\text{H}/^{13}\text{C}$ NMR chemical shifts. *Arch. Pharm. Pharm. Med. Chem.* **331**, 177-192.

7 Appendix

- **Figure 7.1:** 1D ^1H -NMR spectrum of **D/A-Fhp 2** at 2 °C and pH 7.0
- **Table 7.1:** Assignments for **H-Fhp-OH** at 2 °C and pH 7.0
- **Table 7.2:** Assignments for **Ac-Fhp-NH₂** at 2 °C and pH 2.0
- **Table 7.3:** Assignments for **Ac-Fhp-NH₂** at 25 °C and pH 7.0

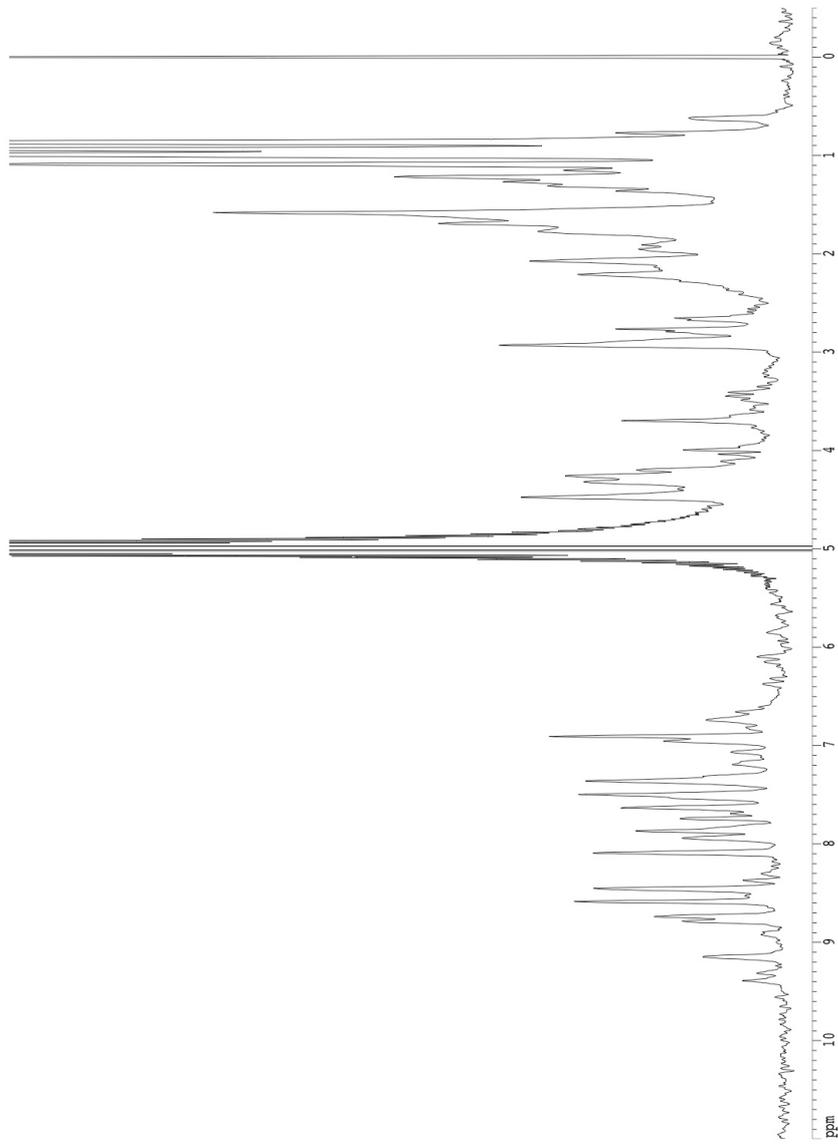


Figure 7.1. 1D ^1H -NMR spectrum of **D/A-Fhp 2** (see 4.2) at 2°C in 10 mM potassium phosphate, pH 7.0 with 5 % (v/v) D_2O . The spectrum was acquired with excitation sculpting (180°) water suppression. Chemical shifts relative to internal DSS.

Table 7.1. Resonance Assignment of H-Fhp-OH at pH 7.0 and 2 °C.

Residue	H ^N	H ^α	H ^β	Other
Ala 12	-	4.073	1.529	
Tyr 13	8.790	5.165	2.900 2.766	H _{2,6} 6.933 H _{3,5} 6.735
Val 14	9.11 ^{a,d}	4.536	2.000	H ^γ 0.898, 0.864
Arg 15	8.549	4.379	1.505	H ^γ 1.038, 0.903 H ^δ
Lys 16	9.015	4.387	1.734 1.633 ^u	H ^γ 1.356, 1.265 H ^δ 1.594 ^u H ^ε 2.933
Asp 17	9.192 ^a	4.310	2.813 2.658	
Gly 18	8.525	4.050 3.524		
Glu 19	7.763	4.521	2.056 1.837	H ^γ 2.263, 2.159
Trp 20	8.783	4.958	3.061 3.020	H _{2,6} 7.209 H ₄ 7.302 H ₅ 7.055 H ₇ 7.445 HN 10.154
Val 21	9.11 ^{a,d}	4.473	2.042	H ^γ 0.917, 0.870
Leu 22	8.482	4.228	1.608 1.364	H ^γ 1.469 H ^δ 0.715, 0.679
Leu 23	8.145	4.197	1.526	H ^γ 1.648 H ^δ 0.869, 0.843

Chemical shifts in ppm relative to internal DSS in 10 mM potassium phosphate with 5 % (v/v) D₂O.

^a Very broad resonance due to exchange.

^d Not resolved.

^u Uncertain assignment.

Table 7.2. Resonance Assignment of Ac-Fhp-NH₂ at pH 2.0 and 2 °C.

Residue	H ^N	³ J _{H^NH^α}	H ^α	H ^β	Other
Ac-	-	-	1.968		
Ala 12	8.266	6.8 Hz	4.386	1.300	
Tyr 13	8.370	8.6 Hz	4.990	2.860 2.769	H _{2,6} 6.923 H _{3,5} 6.728
Val 14	8.970 ^a	n.d. ^a	4.494	1.984	H ^γ 0.891, 0.841
Arg 15	8.554	7.0 Hz	4.398	1.547	H ^γ 1.141, 1.072 H ^δ 2.511, 2.363
Lys 16	8.981	7.6 Hz	4.381	1.743 1.586	H ^γ 1.363, 1.270 H ^δ 1.625 H ^ε 2.933
Asp 17	9.151 ^a	n.d. ^a	4.349	2.826 2.685	
Gly 18	8.529	-	4.023 3.548		
Glu 19	7.816	8.1 Hz	4.505	2.058 1.857	H ^γ 2.296, 2.215
Trp 20	8.721	7.6 Hz	4.947	3.068	H _{2,6} 7.207 H ₄ 7.336 H ₅ 7.053 H ₇ 7.436 HN 10.148
Val 21	9.120 ^a	n.d. ^a	4.431	2.028	H ^γ 0.901, 0.852
Leu 22	8.473	6.5 Hz	4.120	1.519	H ^γ 1.417 H ^δ 0.699, 0.665
Leu 23	8.509	n.d. ^d	4.241	1.571 ^d	H ^γ 1.571 ^d H ^δ 0.889, 0.849
-NH ₂	7.693 7.163	-			

Chemical shifts in ppm relative to internal DSS in 10 mM potassium phosphate with 5 % (v/v) D₂O.

^a Very broad and weak resonance due to exchange.

^d Not resolved.

Table 7.3. Resonance Assignment of Ac-Fhp-NH₂ at pH 7.0 and 25 °C.

Residue	H ^N	H ^α	H ^β	Other
Ac-	-	1.971		
Ala 12	8.135	4.341	1.282	
Tyr 13	8.173	4.897	2.904 2.816	H _{2,6} 6.964 H _{3,5} 6.750
Val 14	8.613	4.356	1.961	H ^γ 0.876, 0.839
Arg 15	8.387	4.398	1.604	H ^γ 1.263, 1.170 H ^δ 2.706, 2.608
Lys 16	8.781	4.363	1.757 1.624 ^u	H ^γ 1.366, 1.290 H ^δ 1.640 ^u H ^ε 2.945
Asp 17	8.826	4.388	2.771 2.667	
Gly 18	8.344	3.997 3.624		
Glu 19	7.889	4.421	2.019 1.862	H ^γ 2.228, 2.139
Trp 20	8.485	4.864	3.120	H _{2,6} 7.221 H ₄ 7.408 H ₅ 7.082 H ₇ 7.443 HN 10.148
Val 21	8.699 ^a	4.300	1.991	H ^γ 0.877, 0.836
Leu 22	8.269	4.196	1.551 1.451 ^d	H ^γ 1.451 ^d H ^δ 0.740
Leu 23	8.332	4.275	~ 1.59 ^d	H ^γ ~ 1.59 ^d H ^δ 0.885, 0.849
-NH ₂	7.529 7.047			

Chemical shifts in ppm relative to internal DSS in 10 mM potassium phosphate with 5 % (v/v) D₂O.

^a Significantly broadened resonance due to exchange.

^d Not resolved.

^u Uncertain assignment.

8 Curriculum Vitae

- Not Included in Electronic Version -

9 Acknowledgements

This work was carried out from September 2004 to November 2007 in the laboratory of Prof. Dr. Kiefhaber in the Division of Biophysical Chemistry at the Biozentrum of the University of Basel.

First of all, I want to thank Prof. Dr. Thomas Kiefhaber for his excellent supervision. I am very grateful for his support and encouragement and the balance of guidance and independency he gave to me. The time in his lab was a rewarding experience.

I further want to thank Prof. Dr. J. Seelig for his kind support after the major part of the group had moved to Munich, and for accepting to referee this thesis.

Many thanks to all members of the former Basel Lab – Annett Bachmann, Christophe Bodenreider, María Dolores Crespo, Beat Fierz, Sarah Güthe, Judith Habazettl, Kerstin Hoffmann-Jacobsen, Karin Joder, Florian Krieger, Robert Kübler, Stefan Langheld, Andreas Möglich, Rita Müller, Christian Nyffenegger, Manuela Schätzle, Therese Schulthess, Tobias Schümmer and Josef Wey.

I am obliged to Josef for the synthesis of xanthonic acid. I am particularly indebted to Judith who taught me the practical aspects of NMR spectroscopy and became engaged with the foldon project. Special thanks go to Beat for his enthusiasm and the close collaboration on the helix project, which included many discussions and great moments of insight.

I thank all my colleagues on the 6th floor, as well as the technical workshop, namely Hans Vogt, Dr. Gernot Hänisch, Simon Saner and Leo Faletti, for their valuable help. I also want to thank Dr. Daniel Häusinger from the Department of Organic Chemistry for help at the spectrometer and initial NMR measurements and Dr. Peter Henklein from the Humboldt University Berlin for precious expert knowledge on peptide synthesis.

For proofreading parts of this thesis I want to acknowledge Rita and Christine Kaimer.

I thank all those people with a deep interest in science and a passion for teaching who contributed to my education.

I would like to express my deepest gratitude to my friends, my sister and brother, my parents and Christine.

10 Manuscripts

10.A Local Conformational Dynamics in α -Helices Measured by Fast Triplet Transfer

with Supporting Information

- Manuscript Replaced by Article -

10.B Temperature Dependence of Local Equilibrium Folding and Unfolding of an α -Helix

10.C Effect of Thioxopeptide Bonds on α -Helix Structure and Stability

with Supporting Information

- Manuscript Replaced by Article -

10.D NMR Structure of a Monomeric Intermediate on the Evolutionarily Optimized Assembly Pathway of the Trimeric Foldon Domain

with Supporting Information

10.E Conformational Dynamics in the Native and Denatured State of the Villin Headpiece Subdomain Measured by Triplet-Triplet Energy Transfer

with Supporting Information

Local conformational dynamics in α -helices measured by fast triplet transfer

Beat Fierz^{a,1,2}, Andreas Reiner^{b,1}, and Thomas Kiefhaber^{b,3}

^bChemistry Department and Munich Center for Integrated Protein Science, Technische Universität München, Lichtenbergstrasse 4, D-85747 Garching, Germany; and ^aDivision of Biophysical Chemistry, Biozentrum der Universität Basel, Klingelbergstrasse 70, CH-4056 Basel, Switzerland

Edited by Robert L. Baldwin, Stanford University Medical Center, Stanford, CA, and approved November 26, 2008 (received for review September 3, 2008)

Coupling fast triplet–triplet energy transfer (TTET) between xanthone and naphthylalanine to the helix–coil equilibrium in alanine-based peptides allowed the observation of local equilibrium fluctuations in α -helices on the nanoseconds to microseconds time scale. The experiments revealed faster helix unfolding in the terminal regions compared with the central parts of the helix with time constants varying from 250 ns to 1.4 μ s at 5 °C. Local helix formation occurs with a time constant of \approx 400 ns, independent of the position in the helix. Comparing the experimental data with simulations using a kinetic Ising model showed that the experimentally observed dynamics can be explained by a 1-dimensional boundary diffusion with position-independent elementary time constants of \approx 50 ns for the addition and of \approx 65 ns for the removal of an α -helical segment. The elementary time constant for helix growth agrees well with previously measured time constants for formation of short loops in unfolded polypeptide chains, suggesting that helix elongation is mainly limited by a conformational search.

α -helix–coil transition | protein dynamics | protein folding | triplet–triplet energy transfer

Conformational dynamics is of fundamental importance for folding and function of biomolecules (1). Structural fluctuations in proteins occur on different time scales and involve various degrees of motions from side-chain rotations to complete folding/unfolding reactions. Only a few methods exist that allow the study of conformational transitions between different states in equilibrium. NMR spectroscopy, single-molecule fluorescence and hydrogen/deuterium (H/D) exchange experiments allow the characterization of different kinds of equilibrium dynamics, but they are insensitive for transitions on the nanoseconds to microseconds time scale (2, 3). On this time scale, large-scale backbone movements (4, 5) and structural transitions in α -helices (6–8) and β -hairpins (9) occur that play a key role in conformational transitions during protein folding, misfolding, and regulation. We therefore sought a method to study site-specific equilibrium transitions on the nanoseconds to microseconds time scale that allows us to gain insight into the dynamics of the α -helix-to-coil transition.

Isolated α -helices are multistate systems with higher helix content in the center compared with the helix ends (10–13). Numerous theoretical models for the helix–coil transition have been proposed (10, 11, 13–18). Theoretical models typically assume a nucleation–growth mechanism with the establishment of a first helical turn representing an entropically unfavorable, slow nucleation reaction. In the subsequent growth reactions, helical segments are added, which is supposed to be a fast process. Experimental studies on helix dynamics have been limited to perturbation methods starting from an ensemble of helical states (6–8, 19–23). Dielectric relaxation measurements (6) and ultrasonic absorption techniques (7) on long homopolymeric helices revealed relaxation times for helix unfolding on the hundreds of nanoseconds to microseconds time scale, depending on the nature of the solvent and on the amino acid sequence. From these experiments, time constants in the range of 0.1–10 ns were estimated for the growth reaction (7). Nanosecond temperature-jump (T-jump) methods on short alanine-based peptides also showed relaxation times for helix unfolding on the

hundreds of nanoseconds to microseconds time scale (8, 19–22, 24) and also yielded similar time constants for helix growth (8). T-jump relaxation dynamics using an N-terminal fluorescence probe revealed faster helix unfolding in the N-terminal region compared with the decrease in average helix content (8). Site-specific IR studies on ¹³C-labeled peptides, in contrast, suggested only little position dependence of the relaxation times after a T-jump (22, 25–27). These studies gave insight into the kinetics of perturbation-induced helix unfolding, but they did not yield information on the dynamic properties of α -helices under equilibrium conditions. Equilibrium dynamics may substantially differ from perturbation-induced kinetics because of the non-2-state nature of the helix–coil transition. This complexity may lead to different dynamics depending on the initial conditions as observed for different-size T-jumps to identical final conditions (22).

Linking an irreversible process to a conformational transition is a powerful kinetic approach to study stability and dynamics of chemical equilibria. A well-known example is the use of H/D exchange to gain information on individual hydrogen bonds in proteins (2, 28, 29). H/D exchange occurs on the milliseconds to hours time scale, depending on pH, and is thus not suited to measure dynamics on the nanoseconds to microseconds time scale. We have previously applied diffusion-controlled triplet–triplet energy transfer (TTET) between a xanthone donor and a naphthylalanine acceptor group to observe intrachain loop formation in unfolded polypeptide chains (4, 5, 30). TTET through loop formation is an irreversible process that is based on Van der Waals contact between donor and acceptor. It should thus enable us to monitor dynamics in folded and partially folded structures when linked to a folding/unfolding equilibrium. Because loop formation occurs on the 10–100 ns time scale, depending on loop length, amino acid sequence, and on the position of the loop within the chain (5, 30), this approach allows us to monitor dynamics that are 4–5 orders of magnitude faster than those accessible to H/D exchange. A prerequisite for the application of TTET is that contact between the TTET labels can only occur by loop formation in the coiled state and is prevented in the folded structure. Experiments by Lapidus *et al.* (31) had indicated that contact formation between the ends of an α -helix does not occur. To investigate local helix folding and unfolding dynamics, we introduced xanthone and naphthylalanine with *ij*+6 spacing at different positions along an α -helical peptide. This places the labels at opposite sides of the helix, which prevents TTET in the helical state and requires at least partial unfolding before triplet transfer can occur (Scheme 1; see also [Scheme S1](#) in

Author contributions: B.F., A.R., and T.K. designed research; B.F. and A.R. performed research; B.F., A.R., and T.K. analyzed data; and B.F., A.R., and T.K. wrote the paper.

The authors declare no conflict of interest.

This article is a PNAS Direct Submission.

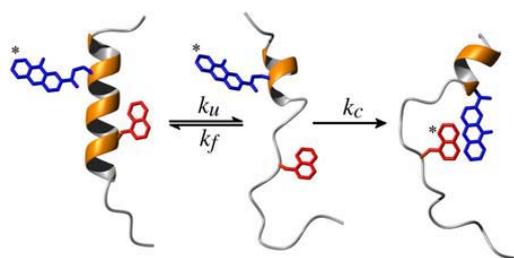
¹B.F. and A.R. contributed equally to this work.

²Present Address: Laboratory of Synthetic Protein Chemistry, The Rockefeller University, New York, New York 10065.

³To whom correspondence should be addressed. E-mail: t.kiefhaber@tum.de.

This article contains supporting information online at www.pnas.org/cgi/content/full/0808581106/DCSupplemental.

© 2009 by The National Academy of Sciences of the USA



Scheme 1.

the supporting information). This approach allowed us to determine the local unfolding and refolding rate constants for the helix-coil transition at equilibrium for different positions in the α -helix. The results were compared with simulations of the helix-coil dynamics using a linear Ising model, which gave insight into the basic dynamics of the helix-coil transition.

Results and Discussion

Design and Global Stability of the Helices. We used an alanine-based model peptide with the sequence Ac-(A)₅-(AAARA)₃-A-NH₂ to measure local α -helix dynamics and stability. Similar peptides were shown to exhibit $\approx 70\%$ helical content at low temperature (8, 32), and their unfolding kinetics have been probed in relaxation studies (8, 19–22). TTET labels were introduced at different positions along the peptide by using a xanthone moiety (Xan; X) as triplet donor and the nonnatural amino acid 1-naphthylalanine (Nal; Z) as triplet acceptor (see Table 1). The labels were placed with $i, i + 6$ spacing to prevent contact between the labels in the helical state (Fig. 1A). Introducing the labels in the N-terminal region (X1–Z7), in the interior (X5–Z11, X7–Z13, X11–Z17) or in the C-terminal region (X15–Z21) allowed us to monitor local dynamics and stability in different regions of the helix. To compare the local dynamics with global helix unfolding and refolding, the TTET labels were positioned at the ends of the helix in the X0–Z21 peptide.

Xanthone and naphthylalanine are expected to have a small helix-destabilizing effect compared with alanine. Fig. 1B shows that all peptides form similar amounts of helical structure as judged by the positive CD bands at 190 nm and the negative bands at ~ 208 and 222 nm. Contributions of Nal to the CD signal at ≈ 220 nm prevent a quantitative determination of the helix content, but the relative differences between the peptides can be assessed. In the X0–Z21 reference peptide, the effect of the labels on helix stability should be negligible because of the low helix content at the termini (12). Introduction of the labels with $i, i + 6$ spacing has only little effect on global helix stability relative to the X0–Z21 peptide. A slight destabilization is observed when the labels are placed near the helix center, as expected from helix-coil theory (11–13). Based on the differences in CD signal at 222 nm, the helix content of the different peptides varies by $<15\%$. This similarity in helical content

Table 1. Sequences of helical peptides used in TTET experiments

Peptide	Sequence
X0–Z21	X-AAAAA AAARA AAARA AAARA ZGG-NH ₂
X1–Z7	Ac-XAAAA AZARA AAARA AAARA A-NH ₂
X5–Z11	Ac-AAAAX AAARA ZAARA AAARA A-NH ₂
X7–Z13	Ac-AAAAA AXARA AAZRA AAARA A-NH ₂
X11–Z17	Ac-AAAAA AAARA XAARA AZARA A-NH ₂
X15–Z21	Ac-AAAAA AAARA AAARX AAARA Z-NH ₂

X, xanthonic acid (Xan) attached to the N-terminus (X0–Z21) or to α, β -Dpr (all other peptides); Z, 1-naphthylalanine (Nal).

was confirmed by thermal melting transitions, which showed similar changes in the CD signal upon unfolding for all peptides (Fig. S1).

Kinetics of TTET Coupled to a Helix-Coil Transition. We performed TTET experiments at 5 °C to measure dynamics in the different peptides. TTET was monitored by the change in xanthone triplet absorbance at 590 nm (5). All peptides exhibit double-exponential TTET kinetics (Fig. 1C), which was confirmed by an analysis of the distribution of time scales for the observed kinetics (Fig. S2). The slowest kinetics were found for the X0–Z21 peptide with a main kinetic phase of $\lambda_1 = 2.3 \cdot 10^5 \text{ s}^{-1}$ (90% amplitude) and a faster phase of $\lambda_2 = 2.6 \cdot 10^6 \text{ s}^{-1}$ (10% amplitude). In the peptides with local $i, i + 6$ spacing, faster TTET kinetics with a larger amplitude of the fast phase are observed when the labels are attached near the termini compared with the central positions (Table 2). In all peptides, the 2 observable reactions are faster than spontaneous donor triplet decay, which occurs with $\lambda_T = 2.5 \cdot 10^4 \text{ s}^{-1}$, measured in donor-only reference helices. This demonstrates that the observed triplet decay in the helical peptides is due to intrinsic dynamics in the helix-coil system and is not reflecting the triplet lifetime. The observed double-exponential kinetics suggest an equilibrium between 2 distinct populations of molecules. To test for the origin of the 2 kinetic phases, we stabilized the helical state by addition of 40% TFE (33), which results in slow single-exponential TTET kinetics with a rate constant of $\lambda = 1.2 \cdot 10^5 \text{ s}^{-1}$ for the X5–Z11 peptide (Fig. 1) indicating the absence of the fast process. In contrast, addition of urea, which destabilizes helical structure (34), leads to an increase in amplitude of the fast phase in all peptides (see below). These results suggest that the slow process originates from molecules that contain a critical amount of helical structure between the labels. In these molecules, TTET can only occur via helix unfolding. The fast phase is related to TTET in conformations that are at least partially unfolded between the labels. This model is supported by the magnitude of the observed time constants. The slow phase has time constants on the microseconds time scale in the X0–Z21 peptide, which is similar to the observed relaxation times for helix unfolding after temperature jump (8, 19–22). The fast phase is on the 100-ns time scale, which is in agreement with the time constant for loop formation in unfolded polypeptide chains (5, 30).

Rate Constants for Local Helix Formation and Unfolding. The observation of 2 observable rate constants, λ_1 and λ_2 , for TTET in the helical peptides allows the determination of the microscopic rate constants k_u and k_f reporting on unfolding and formation of helical structure and of k_c , the rate constant for loop formation in the unfolded state in a single experiment (see Scheme 1). We use the analytical solutions for the mechanism shown in Scheme 1 (see *SI Text*) to analyze the kinetics without any simplifying assumptions. It should be noted that this approach is different from the classical EX1 and EX2 limits commonly used to analyze H/D exchange kinetics, which allow the determination of either the equilibrium constant between a folded and an unfolded state (EX2) or the microscopic rate constant for the unfolding process (EX1) and is only applicable if the unfolded state is populated to very low amounts. It further uses simplified equations that are only valid under certain conditions (29). Determination of all microscopic rate constants from these experiments thus requires a change in the experimental conditions from the EX2 limit to the EX1 limit.

Because of the low amplitude of the fast phase in the central parts of the helix, the experiments were performed in the presence of different urea concentrations between 0 and 7 M. This allows a more reliable determination of all microscopic rate constants and gives additional mechanistic insight into the dynamics (35). We assumed that urea has a linear effect on $\ln(k_f)$ and $\ln(k_u)$ (34, 36) as well as on $\ln(k_c)$ (5, 37) according to

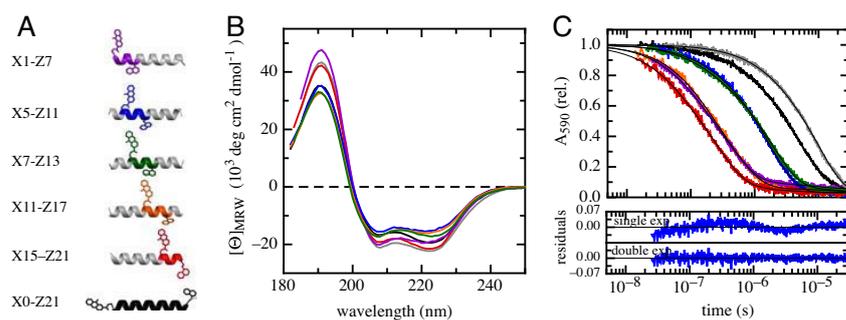


Fig. 1. Structure and dynamics of the helical peptides labeled with donor and acceptor groups for TTET. (A) Schematic position of the labels in the helix. For sequences, see Table 1. (B) Far-UV CD spectra. (C) TTET kinetics measured by the decrease in xanthone absorbance at 590 nm. The colors of the lines correspond to the colors of the helical segments probed by TTET as shown in A. The black lines represent the results from double-exponential fits. As an example, the residuals for single- and double-exponential fits are displayed for the TTET kinetics in the X5-Z11 helix. Results for the X5-Z11 helix in the presence of 40% TFE are additionally shown in gray. Experiments were performed at 5 °C.

$$\ln k_i = \ln k_i(\text{H}_2\text{O}) - \frac{m_i[\text{urea}]}{RT}, \quad [1]$$

where k_i denotes a given microscopic rate constant, $k_i(\text{H}_2\text{O})$ is the rate constant in the absence of urea, and m_i denotes the sensitivity of the respective reaction to urea (m value).

Double-exponential TTET kinetics are observed for all peptides and at all urea concentrations. The urea dependences of λ_1 and λ_2 and of the respective amplitudes A_1 and A_2 were fitted globally to the analytical solution of the 3-state model shown in Scheme 1 (38) to yield the microscopic rate constants $k_i(\text{H}_2\text{O})$ and their urea dependences (m_i) (see *SI Text*). Fig. 2 shows the results from the global fit for all helices, which reveal that λ_1 , λ_2 , A_1 , and A_2 contain contributions from all microscopic rate constants. This allows a reliable determination of k_f , k_u , and k_c in all helical peptides. The k_f and k_u values obtained for the different peptides reflect the local dynamics for helix folding and unfolding at different positions in the helix (Fig. 3 and Table 2). Comparison of the results for the different helices shows that local helix formation does not systematically depend on the position along the helix and has a time constant ($1/k_f$) ≈ 400 ns in all peptides (Fig. 3A). Local helix opening, in contrast, is ≈ 6 -fold faster at the termini ($1/k_u = 250$ ns) compared with the helix center ($1/k_u = 1.4 \mu\text{s}$; Fig. 3B). This difference does not arise from the effect of the labels on helix stability, because the destabilization is largest at the helix center (Fig. 1) (12) where the slowest rate constants for unfolding are observed. Attaching the labels at the ends of the helix in the X0-Z21 reference peptide yields slower dynamics compared with all local folding/unfolding reactions with time constants of 1.1 μs for helix formation and 3.3 μs for helix unfolding (Fig. 1 and Table 2). This observation suggests that end-to-end contact formation monitors more global structural transitions.

The urea dependence of k_f shows an almost position-independent m_f value of ≈ 800 (J/mol)/M (Table 2). The m_u values are much smaller and show complex behavior, with positive values

at the helix termini and negative values for the central positions. In terms of a classical interpretation of protein folding m values, this indicates that the transition state for helix formation has native-like solvent-accessible surface area (35).

The microscopic rate constants for contact formation k_c in the coil state are ≈ 100 ns in all peptides (Table 2). These values are slightly lower than the rate constants observed for formation of $i, i + 6$ interactions in polyserine model chains at 5 °C (39), which is expected, taking into account that loop formation in the helical peptides involves interior parts of the chain, which are less flexible than the ends (30). The m_c values are approximately 200 (J/mol)/M and virtually position independent. This value is identical to the m_c values found for loop formation in unfolded model polypeptide chains (5, 37). These results indicate that in the alanine-based helical peptides the ensemble of conformations that can form contact between the labels has similar properties as an ensemble of unfolded polypeptide chains.

Position Dependence of Local α -Helix Stability. The results from the TTET experiments reveal higher helix content in the central part of the helix compared with the termini, which is in agreement with helix-coil theory (11, 10) and has previously been observed experimentally by amide hydrogen exchange (32, 40), by NMR spectroscopy (41, 42), and by electron spin resonance (43). The k_f and k_u values determined for the different peptides allow the determination of local helix stability. The equilibrium constant, $K_{\text{eq}} = k_f/k_u$ obtained from the TTET experiments is ≈ 3.9 in the central part of the helix and ≈ 0.5 near the ends (Fig. 3C). These equilibrium constants for local helix stability are slightly lower in all regions of the helix compared with equilibrium constants for individual hydrogen bonds measured in H/D exchange experiments under the same conditions (32). This difference can be explained by the helix-destabilizing effect of the TTET labels. The local equilibrium constants from the TTET experiments are in good agreement with predicted helix contents based on the AGADIR algorithm (44)

Table 2. Kinetic and thermodynamic parameters

Peptide	λ_1 (10^6 s^{-1})	λ_2 (10^6 s^{-1})	A_1 , %	A_2 , %	k_f (10^6 s^{-1})	k_u (10^6 s^{-1})	k_c (10^6 s^{-1})	m_f (J/mol)/M	m_u (J/mol)/M	m_c (J/mol)/M	K_{eq}
X1-Z7	2.4 ± 0.3	12 ± 4.0	71 ± 5	29 ± 5	3.0 ± 0.6	3.3 ± 0.3	11.1 ± 1.5	696 ± 224	199 ± 89	136 ± 102	0.9 ± 0.3
X5-Z11	0.78 ± 0.08	8.2 ± 3.2	88 ± 3	12 ± 3	4.0 ± 0.7	1.0 ± 0.1	7.4 ± 1.1	908 ± 131	-66 ± 71	149 ± 97	3.9 ± 0.9
X7-Z13	0.55 ± 0.04	5.6 ± 3.5	81 ± 2	19 ± 2	1.9 ± 0.4	0.7 ± 0.1	8.0 ± 0.7	831 ± 116	-222 ± 73	230 ± 52	2.5 ± 0.6
X11-Z17	2.0 ± 0.3	8.0 ± 2.4	67 ± 10	33 ± 10	2.5 ± 0.6	3.0 ± 0.4	8.1 ± 1.1	687 ± 254	83 ± 140	200 ± 83	0.8 ± 0.3
X15-Z21	2.7 ± 0.5	11 ± 4.0	58 ± 10	42 ± 10	1.8 ± 0.5	3.7 ± 0.4	9.0 ± 1.2	772 ± 330	321 ± 149	154 ± 99	0.5 ± 0.2
X0-Z21	0.23 ± 0.03	2.6 ± 0.9	90 ± 2	10 ± 2	0.9 ± 0.2	0.3 ± 0.1	1.6 ± 0.3	537 ± 166	-116 ± 63	90 ± 80	—

The Xan triplet decays of all peptides can be fitted by double exponential functions yielding the apparent rate constants λ_1 and λ_2 and the relative amplitudes A_i , which are given only for the decays in the absence of urea. The microscopic rate constants k_f , k_u , and k_c were obtained by a global analysis of all decays at different urea concentrations using the 3-state model (Scheme 1) and are given for 0 M urea. Their sensitivity toward urea is described by the m_i values. K_{eq} gives the local equilibrium constant between closed and open conformations between the labels.

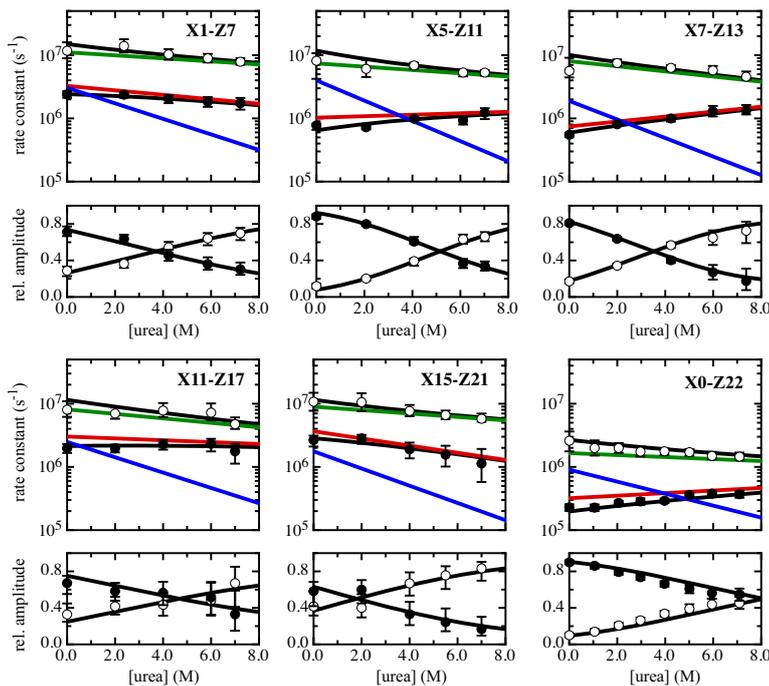
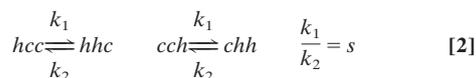


Fig. 2. Effect of urea on the TTET kinetics in different helical peptides at 5 °C. The observable rate constants λ_1 (●) and λ_2 (○) (Upper) and their corresponding amplitudes, A_1 (●) and A_2 (○) (Lower) obtained from double-exponential fits to the Xan triplet decays (see Fig. 1C) are shown. The results from the global fit to the 3-state model (Scheme 1) are displayed for the observable rate constants and their amplitudes (black lines) as well as for the microscopic rate constants k_u (red lines), k_f (blue lines), and k_c (green lines).

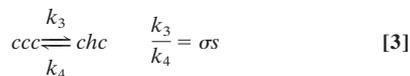
when the TTET labels are modeled with Trp residues. AGADIR predicts equilibrium constants ≈ 5 in the helix center and ≈ 1.5 near the termini. In further agreement with our experimental results, AGADIR predicts higher helix content in the N-terminal regions than in the C-terminal region.

Simulation of Local Helix–Coil Dynamics with a Linear Ising Model. To compare the experimental results with predictions from kinetic models for the underlying microscopic dynamics, we performed Monte Carlo simulations of helix–coil dynamics using a linear Ising model. The helix is represented by a finite sequence of 21 identical residues, which can be either in the helical state h or the coil state c . We use the statistical weights of Zimm and Bragg to assign the equilibrium constants for the following types of reactions (10):

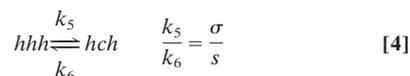
helix elongation:



helix nucleation:



coil nucleation:



This formalism considers only nearest-neighbor interaction as in the widely used 2×2 matrix approximation of Zimm and Bragg (10). Additional factors for helix nucleation, γ_h , and coil nucle-

ation, γ_c , were introduced by Schwarz (16) to allow for additional kinetic effects on the nucleation reactions relative to elongation:

$$k_3 = \gamma_h \cdot \sigma \cdot k_1 \quad k_4 = \gamma_h \cdot k_2 \quad [5]$$

$$k_5 = \gamma_c \cdot \sigma \cdot k_2 \quad k_6 = \gamma_c \cdot k_1 \quad [6]$$

Schwarz suggested that it is reasonable to assume that $1 \leq \gamma = 1/\sqrt{\sigma}$ (16). In the simulations, we used a σ value of 0.003 and a position-independent s value of 1.31, which is the average s value experimentally determined for a similar peptide at 5 °C (45). Varying γ_h and γ_c between 1 and 9 has only little effect on the results. Values of $\gamma_h = \gamma_c = 2$ were used for all simulations. The simulations were performed based on a reduced time, t' , by using k_1 as reference ($t' = t \cdot k_1$). All other rate constants were scaled to k_1 according to Eqs. 2–6. The simulations were started with an equilibrium ensemble of helices. The time evolution of the system was simulated with discrete time steps of $\Delta t' = 0.05$. These simulations are similar to the method applied by Poland (46) to simulate nonequilibrium dynamics of helix formation. A detailed description of the simulations is given in the *SI Text*.

Fig. 4 shows a typical result from the equilibrium simulations that yield an average helix content of 58.2%. This value agrees well with the experimentally determined helical content and with equilibrium Zimm and Bragg theory, which predicts 58.3% average helical content. A closer inspection of the results from the simulations reveals that the equilibrium dynamics are largely dominated by shrinking and growth of an existing helix, whereas helix and coil nucleation events are rare. This explains the insensitivity of the simulations to variations in γ_h and γ_c and suggests that the equilibrium TTET experiments do not yield information on the dynamics of helix nucleation events. The simulations further reveal that the commonly used single-sequence approximation (10, 13) holds most of the time, but 2 separate helical regions are sometimes observed (Fig. 4).

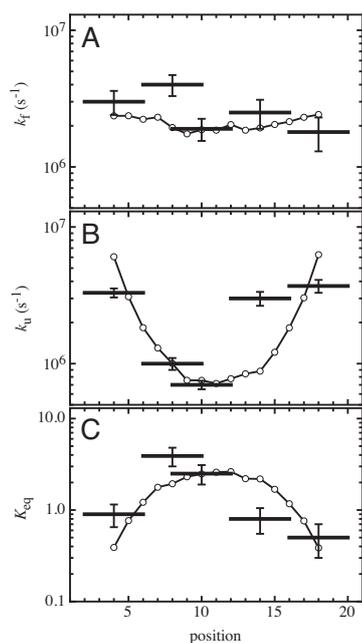


Fig. 3. Position dependence of local helix dynamics and stability. The horizontal bars indicate the regions of the helix probed in the different peptides. The microscopic rate constants for helix formation, k_f (A), and helix unfolding, k_u (B), obtained from the global fit of the TTET kinetics (see Table 1) are shown in addition to the equilibrium constant, K_{eq} (C). For comparison, the results from the simulations using the kinetic Ising model (C) with $s = 1.31$, $\sigma = 0.003$ and $\gamma_h = \gamma_c = 2$ are displayed. The rate constants used in the simulations were scaled with $k_1 = 2.1 \cdot 10^7 \text{ s}^{-1}$ (see Eqs. 2–6).

Comparison of Experiment with Simulation. For comparison with the experimental time constants, we extracted first-passage times (FPTs) for local folding and unfolding kinetics from the simulations. Histograms of the FPTs for folding and unfolding were used to calculate rate constants k_u/k_1 and k_f/k_1 for the local unfolding and refolding reactions (Fig. S3). Because TTET experiments probe a region of 5 residues between the labels, we calculated kinetics for segments of 5 residues, which are counted as helical if at least 4 of these residues are helical at a given time. However, the results do not critically depend on the segment length used in data analysis. Monitoring the dynamics of single segments gives nearly identical results (see Fig. S4). Fig. 3 shows that the simulations yield the same position dependence of the rate constants for local helix folding and unfolding as the TTET experiments. The resulting rate constants for helix formation are virtually position independent. Helix unfolding, in contrast, occurs faster at the ends compared with the helix center despite the assumption of position-independent microscopic rate constants in Eqs. 2–6. The rate constants obtained from the simulations match the experimental values when the reduced time is scaled with a constant factor of $2.1 \cdot 10^7 \text{ s}^{-1} = 1/(48 \text{ ns})$, which represents the elementary rate constant k_1 for helix elongation. In combination with the s value of 1.31, this yields an elementary time constant for helix unfolding, $1/k_2$, of $\approx 63 \text{ ns}$ (Eq. 2). The assumption of a single helical segment between the labels being sufficient to prevent loop formation yields values of $1/k_1 = 37 \text{ ns}$ and $1/k_2 = 49 \text{ ns}$. The local differences between the experimental data and the results from the simulations (Fig. 3) are most likely due to the use of a uniform sequence with an average s value in the simulations. The experimentally investigated model helices, in contrast, contain arginine residues

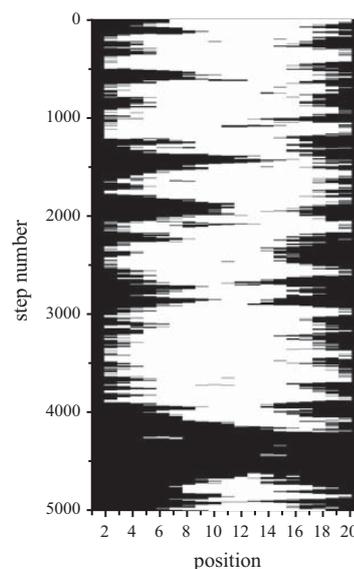


Fig. 4. Snapshot of a typical simulation showing 5,000 discrete steps ($\Delta t' = 0.05$, i.e., $t' = 250$) of a 21-residue helix. Helical segments h are represented as white boxes and coil segments c as black boxes. The parameters used in the simulations were $s = 1.31$, $\sigma = 0.003$ and $\gamma_h = \gamma_c = 2$.

to increase solubility and are asymmetric molecules with a macrodipole.

The results from TTET experiments and simulations reveal significantly larger differences in helix dynamics between the ends and the central part of a helix compared with T-jump relaxation kinetics. This is due to the different types of processes monitored by the different techniques. As discussed above, irreversible TTET coupled to a helix–coil equilibrium allows the determination of microscopic rate constants for helix opening and closing, whereas T-jump experiments yield macroscopic relaxation times that contain contributions from forward and backward reactions. The fast and position-independent rate constant for helix closing observed in TTET experiments contributes strongly to the observed relaxation rate in T-jump experiments at all positions. This explains the weak position dependence of the relaxation rates. A similar difference is observed when unfolding kinetics of ribonuclease A are monitored by either H/D exchange, which measures opening rate constants for hydrogen bonds, and by optical spectroscopy, which measures relaxation times with contributions from folding and unfolding rate constants (47). This explanation is further supported by simulations of helix relaxation kinetics by using a similar kinetic helix–coil model as shown in Eqs. 2–6, which reproduced the weak position dependence of the experimental relaxation rates measured in T-jump experiments (48).

Mechanism of Equilibrium Helix–Coil Dynamics. Comparison of the experimental results with the simulations gives a mechanistic explanation for the experimentally observed position dependence of the unfolding rate constant k_u . The comparison reveals that the observed position dependence for helix unfolding is a consequence of unzipping from the nearest helix–coil boundary as major mechanism for helix unfolding. This 1-dimensional diffusion process leads to longer diffusion distances for central positions compared with the terminal regions and thus to lower unfolding rate constants in the center of the helix. The elementary processes of adding and removing helical segments occur with position-independent time constants of ≈ 50 and $\approx 65 \text{ ns}$, respectively. These rate constants for

local equilibrium helix folding and unfolding measured by TTET set a limit on the time scales for conformational transitions involving α -helices during protein folding and misfolding.

TTET coupled to helix dynamics is sensitive for processes between ≈ 1 ns and 50 μ s. Therefore, the elongation time constant ($1/k_1$) of ≈ 50 ns reflects the time for formation of a helical segment that is at least stable for some nanoseconds. Much faster local processes, which do not lead to formation or decay of stable structure are not monitored, although they may occur during the experiments. The rate constant for helix elongation, k_1 , obtained from the comparison of the simulations with the experimental data are virtually identical to the experimentally determined rate constant for formation of short protein loops at 5 °C (5), indicating that helix elongation is mainly limited by a conformational search.

Materials and Methods

Peptide Synthesis. All peptides were synthesized by using standard Fmoc-chemistry on an Applied Biosystems 433A peptide synthesizer (5). The Xan derivative 9-oxoxanthene-2-carboxylic acid was synthesized and attached either to the N-terminus (in the X0–Z21 peptide) or to the β -amino group of a α,β -diaminopropionic-acid (Dpr) residue (in all peptides with $i, i + 6$ labeling) (5, 30). Naphthalene was incorporated as the nonnatural amino acid 1-naphthylalanine (Bachem). All peptides were purified to >95% purity by preparative HPLC on a

RP-8 column. Purity was checked by analytical HPLC, and the mass was determined by MALDI mass spectrometry.

TTET and CD Measurements. TTET measurements were performed on a Laser Flash Photolysis Reaction Analyzer (LKS.60) from Applied Photophysics with a Quantel Nd:YAG Brilliant laser (354.6 nm, ≈ 4 nm pulse width, ≈ 50 mJ). Transient absorption traces were recorded at 590 nm to monitor the xanthone triplet band. All measurements were performed in 10 mM phosphate buffer (pH 7.0) at 5 °C. Peptide concentrations were 50 μ M as determined by the Xan absorbance at 343 nm ($\epsilon_{343} = 3,900 \text{ M}^{-1}\text{cm}^{-1}$). All solutions were degassed before the measurements. Urea concentrations were determined by the refractive index according to Pace (49). CD measurements were performed on an Aviv DS62 spectropolarimeter. The experimental data were analyzed by using the ProFit software (QuantumSoft).

Simulations on the Helix–Coil Transition. The model is based on Eqs. 2–6 and the underlying assumptions as discussed in the article. Details of the simulations are given in the *SI Text*.

ACKNOWLEDGMENTS. We thank Josef Wey for the synthesis of 9-oxoxanthene-2-carboxylic acid and Robert L. Baldwin, Annett Bachmann, and Rudolph A. Marcus for comments on the manuscript. This work was supported by grants from the VolkswagenStiftung and the Deutsche Forschungsgemeinschaft (SFB 749) and by the Munich Center for Integrated Protein Science.

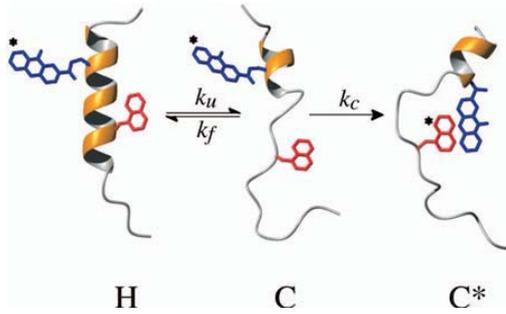
- Frauenfelder H, Sligar SG, Wolynes PG (1991) The energy landscapes and motions of proteins. *Science* 254:1598–1603.
- Arrington CB, Robertson AD (1997) Microsecond protein folding kinetics from native-state hydrogen exchange. *Biochemistry* 36:8686–8691.
- Palmer AG, III (2004) NMR characterization of the dynamics of biomacromolecules. *Chem Rev* 104:3623–3640.
- Bieri O, et al. (1999) The speed limit for protein folding measured by triplet–triplet energy transfer. *Proc Natl Acad Sci USA* 96:9597–9601.
- 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.
- Schwarz G, Seelig J (1968) Kinetic properties and the electric field effect of the helix–coil transition of poly(gamma-benzyl L-glutamate) determined from dielectric relaxation measurements. *Biopolymers* 6:1263–1277.
- Gruenewald B, Nicola CU, Lustig A, Schwarz G, Klump H (1979) Kinetics of the helix–coil transition of a polypeptide with non-ionic side groups, derived from ultrasonic relaxation measurements. *Biophys Chem* 9:137–147.
- 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.
- Munoz V, Thompson P, Hofrichter J, Eaton W (1997) Folding dynamics and mechanism of β -hairpin formation. *Nature* 390:196–199.
- Zimm BH, Bragg JK (1959) Theory of phase transition between helix and random coil in polypeptide chains. *J Chem Phys* 31:526–535.
- Lifson S, Roig A (1961) On the theory of helix–coil transition in polypeptides. *J Chem Phys* 34:1963–1974.
- Chakrabartty A, Schellman JA, Baldwin RL (1991) Large differences in the helix propensities of alanine and glycine. *Nature* 351:586–588.
- Qian H, Schellman JA (1992) Helix–coil theories: A comparative study for finite length polypeptides. *J Phys Chem* 96:3978–3994.
- Schellman JA (1955) The stability of hydrogen-bonded peptide structures in aqueous solutions. *Compt Rend Carlsberg Lab (Ser Chim)* 29:230–259.
- Schellman JA (1958) The factors affecting the stability of hydrogen-bonded polypeptide structures in solution. *J Phys Chem* 62:1485–1494.
- Schwarz G (1965) On the kinetics of the helix–coil transition of polypeptides in solution. *J Mol Biol* 11:64–77.
- Poland D, Scheraga H (1966) Kinetics of the helix–coil transition in polyamino acids. *J Chem Phys* 45:2071–2091.
- Poland D, Scheraga HA (1970) *Theory of Helix–Coil Transitions in Biopolymers* (Academic, New York).
- Williams S, et al. (1996) Fast events in protein folding: Helix melting and formation in a small peptide. *Biochemistry* 35:691–697.
- Lednev IK, Karnoup AS, Sparrow MC, Asher SA (1999) Alpha-helix peptide folding and unfolding activation barriers: A nanosecond UV resonance Raman study. *J Am Chem Soc* 121:8074–8086.
- Thompson PA, et al. (2000) The helix–coil kinetics of a heteropeptide. *J Phys Chem B* 104:378–389.
- Huang CY, et al. (2002) Helix formation via conformation diffusion search. *Proc Natl Acad Sci USA* 99:2788–2793.
- Bredenbeck J, Helbing J, Kumita JR, Woolley GA, Hamm P (2005) Alpha-helix formation in a photoswitchable peptide tracked from picoseconds to microseconds by time-resolved IR spectroscopy. *Proc Natl Acad Sci USA* 102:2379–2384.
- Huang CY, Klemke JW, Getahun Z, DeGrado WF, Gai F (2001) Temperature-dependent helix–coil transition of an alanine based peptide. *J Am Chem Soc* 123:9235–9238.
- Werner JH, Dyer RB, Fesinmeyer RM, Andersen NH (2002) Dynamics of the primary processes of protein folding: helix nucleation. *J Phys Chem B* 106:487–494.
- Pozo Ramajo A, Petty SA, Starzyk A, Decatur SM, Volk M (2005) The α -helix folds more rapidly at the C-terminus than at the N-terminus. *J Am Chem Soc* 127:13784–13785.
- Mikhonin AV, Asher SA, Bykov SV, Murza A (2007) UV Raman spatially resolved melting dynamics of isotopically labeled polyalanyl peptide: Slow α -helix melting follows 3_{10} -helices and π -bulges premelting. *J Phys Chem B* 111:3280–3292.
- Linderström-Lang K (1955) Deuterium exchange between peptides and water. *Chem Soc Spec Publ* 2:1–20.
- Hvidt A, Nielsen SO (1966) Hydrogen exchange in proteins. *Adv Protein Chem* 21:287–386.
- Fierz B, Kiefhaber T (2007) End-to-end vs. interior loop formation kinetics in unfolded polypeptide chains. *J Am Chem Soc* 129:672–679.
- Lapidus LJ, Eaton WA, Hofrichter J (2002) Measuring dynamic flexibility of the coil state of a helix-forming peptide. *J Mol Biol* 319:19–25.
- Rohl CA, Baldwin RL (1994) Exchange kinetics of individual amide protons in 15N-labeled helical peptides measured by isotope-edited NMR. *Biochemistry* 33:7760–7767.
- Luo P, Baldwin RL (1997) Mechanism of helix induction by trifluoroethanol: A framework for extrapolating the helix-forming properties of peptides from trifluoroethanol/water mixtures back to water. *Biochemistry* 36:8413–8421.
- Scholtz JM, Barrick D, York EJ, Stewart JM, Baldwin RL (1995) Urea unfolding of peptide helices as a model for interpreting protein unfolding. *Proc Natl Acad Sci USA* 92:185–189.
- Sánchez IE, 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.
- Santoro MM, Bolen DW (1988) Unfolding free energy changes determined by the linear extrapolation method. 1. Unfolding of phenylmethanesulfonyl alpha-chymotrypsin using different denaturants. *Biochemistry* 27:8063–8068.
- Möglich A, Krieger F, Kiefhaber T (2005) Molecular basis for the effect of urea and guanidinium chloride on the dynamics of unfolded polypeptide chains. *J Mol Biol* 345:153–162.
- Kiefhaber T, Kohler HH, Schmid FX (1992) Kinetic coupling between protein folding and prolyl isomerization. I. Theoretical models. *J Mol Biol* 224:217–229.
- Krieger F, Möglich A, Kiefhaber T (2005) Effect of proline and glycine residues on dynamics and barriers of loop formation in polypeptide chains. *J Am Chem Soc* 127:3346–3352.
- Zhou H-X, Hull LA, Kallenbach NR (1994) Quantitative evaluation of stabilizing interactions in a pre-nucleated α -helix by hydrogen exchange. *J Am Chem Soc* 116:6482–6483.
- Bradley EK, Thomason JF, Cohen FE, Kosen PA, Kuntz ID (1990) Studies of synthetic helical peptides using circular dichroism and nuclear magnetic resonance. *J Mol Biol* 215:607–622.
- Jaravine VA, Alexandrescu AT, Grzesiek S (2001) Observation of the closing of individual hydrogen bonds during TFE-induced helix formation in a peptide. *Protein Sci* 10:943–950.
- Miick SM, Todd AP, Millhauser GL (1991) Position-dependent local motions in spin-labeled analogues of a short alpha-helical peptide determined by electron spin resonance. *Biochemistry* 30:9498–9503.
- Munoz V, Serrano L (1995) Elucidating the folding problem of helical peptides using empirical parameters. III. Temperature and pH dependence. *J Mol Biol* 245:297–308.
- Scholtz JM, Qian H, York EJ, Stewart JM, Baldwin RL (1991) Parameters of helix–coil transition theory for alanine-based peptides of varying chain lengths in water. *Biopolymers* 31:1463–1470.
- Poland D (1996) Discrete step model of helix–coil kinetics: Distribution of fluctuation times. *J Chem Phys* 105:1242–1269.
- Kiefhaber T, Baldwin RL (1995) Kinetics of hydrogen bond breakage in the process of unfolding of ribonuclease A measured by pulsed hydrogen exchange. *Proc Natl Acad Sci USA* 92:2657–2661.
- Doshi UR, Munoz V (2004) The principles of α -helix formation: Explaining complex kinetics with nucleation–elongation theory. *J Phys Chem B* 108:8497–8506.
- Pace CN (1986) Determination and analysis of urea and guanidine hydrochloride denaturation curves. *Methods Enzymol* 131:266–280.

Supporting Information

Fierz et al. 10.1073/pnas.0808581106

SI Text

Analytical Solution of the Kinetic 3-State Model Used to Determine the Microscopic Rate Constants for Helix Formation and Unfolding. The



characteristic equation for the mechanism with the 3 states helix (H), coil (C), and coil with exchanged triplet state (C^*) (Scheme S1) is given by

$$f(\lambda) = \lambda^2 - \lambda(k_u + k_f + k_c) + k_u k_c$$

which yields 2 apparent rate constants λ_1 and λ_2

$$\lambda_{1,2} = \frac{k_u + k_f + k_c \pm \sqrt{(k_u + k_f + k_c)^2 - 4k_u k_c}}{2}$$

The time-dependent behavior of the 3 species for the case that the initial concentration of C_0^* is zero (i.e., the total concentration is $H_0 + C_0$) is given by

$$\begin{aligned} H(t) &= H_0 \left(\frac{k_u(\lambda_1 - k_c)}{\lambda_1(\lambda_1 - \lambda_2)} e^{-\lambda_1 t} + \frac{k_u(k_c - \lambda_2)}{\lambda_2(\lambda_1 - \lambda_2)} e^{-\lambda_2 t} \right) \\ &\quad + C_0 \left(\frac{-k_f \lambda_1}{\lambda_1(\lambda_1 - \lambda_2)} e^{-\lambda_1 t} + \frac{k_f \lambda_2}{\lambda_2(\lambda_1 - \lambda_2)} e^{-\lambda_2 t} \right) \\ C(t) &= H_0 \left(\frac{-\lambda_1 k_u}{\lambda_1(\lambda_1 - \lambda_2)} e^{-\lambda_1 t} + \frac{k_u \lambda_2}{\lambda_2(\lambda_1 - \lambda_2)} e^{-\lambda_2 t} \right) \\ &\quad + C_0 \left(\frac{-\lambda_1(k_u - \lambda_1)}{\lambda_1(\lambda_1 - \lambda_2)} e^{-\lambda_1 t} + \frac{\lambda_2(k_u - \lambda_2)}{\lambda_2(\lambda_1 - \lambda_2)} e^{-\lambda_2 t} \right) \\ C^*(t) &= H_0 \left(\frac{k_u k_c}{\lambda_1(\lambda_1 - \lambda_2)} e^{-\lambda_1 t} + \frac{-k_u k_c}{\lambda_2(\lambda_1 - \lambda_2)} e^{-\lambda_2 t} + \frac{k_u k_c}{\lambda_1 \lambda_2} \right) \\ &\quad + C_0 \left(\frac{k_c(k_u - \lambda_1)}{\lambda_1(\lambda_1 - \lambda_2)} e^{-\lambda_1 t} + \frac{k_c(\lambda_2 - k_u)}{\lambda_2(\lambda_1 - \lambda_2)} e^{-\lambda_2 t} + \frac{k_u k_c}{\lambda_1 \lambda_2} \right) \end{aligned}$$

Furthermore, the starting concentrations are constrained by $H_0/C_0 = k_f/k_u$.

Simulations on the Helix-Coil Transition. *Simulation in discrete time steps.* The simulation was set up in discrete time steps $\Delta t'$, with $t' = t k_1$. A starting conformation of 21 h or c was generated that satisfies the equilibrium probability and has all h in a row (single sequence). The state of the first and last residue was set to be c and did not change throughout the simulation (1). Within each time step $\Delta t'$, all 19 triplets of the sequence were read, and the probabilities $p_i = k_i \Delta t'$ for a transition of their central residues were calculated according to Eqs. 2-6. The resulting probabilities were compared with uniformly distributed random numbers r ($0 < r < 1$) to decide whether the particular transition occurs (if $k_i \Delta t' > r$) or not (if $k_i \Delta t' \leq r$). After evaluation of all triplets, the changes resulting from the transitions within the time step $\Delta t'$ were made in parallel to obtain the new conformation. Typical simulations consisted of 10^6 to 10^7 steps with $\Delta t' = 0.05$. Values of s (278.15 K) = 1.3059 and $\sigma = 0.003$ for alanine-based helices were taken from equilibrium experiments (2, 3). Simulations with σ values in the range of 0.001–0.1 revealed that the helix content of our simulation is influenced by this parameter but not the general results of the simulations.

Rate constants of local helix folding and unfolding. The local helix folding and unfolding dynamics at a single position p are given by the times that the residue remains in the c state (folding time) or the h state (unfolding time). We decided to analyze first-passage times (FPTs), which reduces the number of fast fluctuations arising from repeated growth and shrinking when a helix boundary oscillates around the selected site. This resembles our TTET experiment, which is not sensitive to very fast fluctuations and is based on an irreversible triplet-transfer step. To obtain FPTs of local folding and unfolding, we randomly selected starting conformations from a time course of 10^7 steps. This ensures that the calculations start from a number of unrelated molecules that represent the equilibrium of conformations present in the experiment. These separate simulations were only continued until the criterion for local folding (or unfolding) was met. In this way, we collected 4,000 FPTs for each event, which we stored in 50 equally spaced bins, adjusted to the longest observed FPT (Fig. S3). The distributions of our FPTs can be reasonably approximated by single-exponential functions to yield rate constants of local helix folding k_f/k_1 and unfolding k_u/k_1 (Fig. S3). The ratios of the obtained rate constants agree with the expected equilibrium helix content at the position, which confirms that FPTs and single-exponential decay curves provide a good description of the system. Performing this procedure for different positions of the helix gives the position dependence of folding and unfolding rate constants shown in Fig. S4. Because our experiments probe a region of 5 residues between the labels, we adapted the simulation procedure to monitor the dynamics of larger segments instead of single site fluctuations. A segment of 5 residues was counted as helical if at least 4 of its residues were helical at a certain time (Fig. S4). Assuming that only a single helical segment between the labels is required to prevent TTET between the labels gives very similar results (Fig. S4).

Technical details of the simulations and data analysis. The operations were performed using the framework of Matlab 7 R14 (MathWorks) in combination with compiled C code for time-consuming iterations. Random numbers were generated using a Lehmer type algorithm (4). Fitting of exponential functions to the distributions of FPTs was performed using nonlinear least-square fitting, applying the Matlab *lsqnonlin* algorithm.

1. Zimm BH, Bragg JK (1959) Theory of phase transition between helix and random coil in polypeptide chains. *J Chem Phys* 31:526–535.
2. Scholtz JM, Qian H, York EJ, Stewart JM, Baldwin RL (1991) Parameters of helix–coil transition theory for alanine-based peptides of varying chain lengths in water. *Biopolymers* 31:1463–1470.
3. Scholtz JM, Barrick D, York EJ, Stewart JM, Baldwin RL (1995) Urea unfolding of peptide helices as a model for interpreting protein unfolding. *Proc Natl Acad Sci USA* 92:185–189.
4. Park SK, Miller KW (1988) Random number generators: Good ones are hard to find. *Commun ACM* 31:1192–1201.
5. Provencher SW (1982) CONTIN: A general purpose constrained regularization program for inverting noisy linear algebraic and integral equations. *Comput Phys Commun* 27:229.

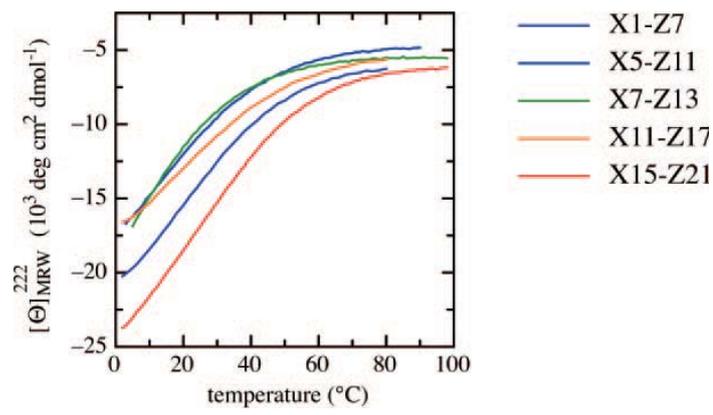


Fig. S1. Thermal melting curves of the different peptides indicated in the legend detected by the change in ellipticity at 222 nm. Measurements were performed in 10 mM phosphate buffer (pH 7).

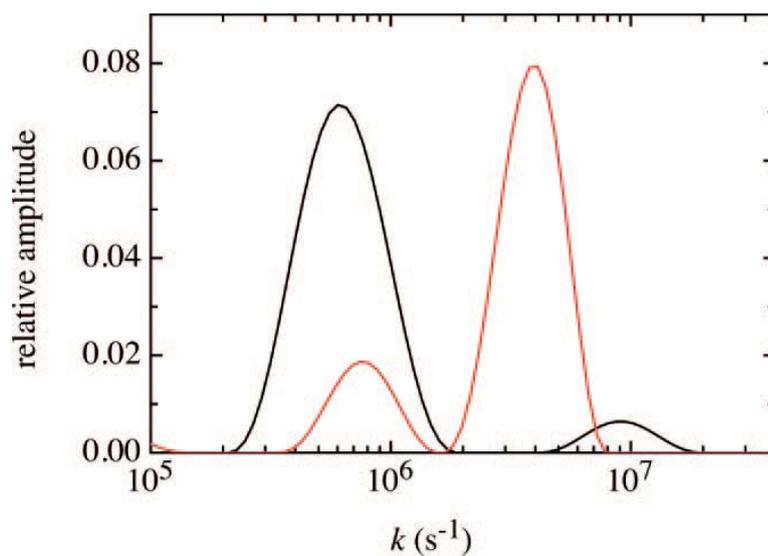


Fig. S2. Distributions of rate constants in the Xan triplet decay curves of X5-Z11 in the absence of urea (black) and in the presence of 7.0 M urea (red). The analysis was performed using the CONTIN package (5). The results from fitting double-exponential functions to the same data are shown in Fig. 2 of the article. Measurements were performed in 10 mM potassium phosphate (pH 7.0) at 5 °C.

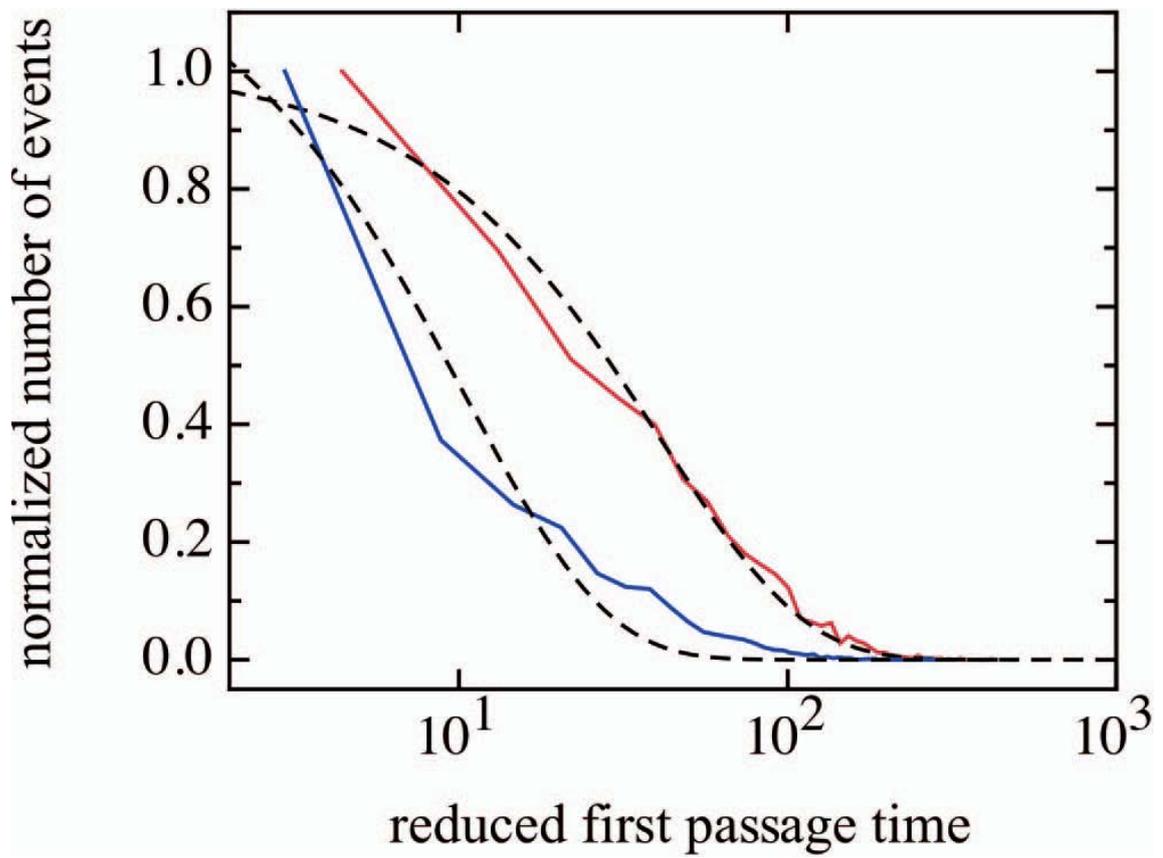


Fig. S3. Histogram of FPTs (normalized) for local folding (blue) and unfolding (red) in the helix center (position 10) evaluated for a single segment (segment length = 1). The simulation parameters were chosen as described in Fig. 4 of the article. The broken lines show single-exponential fits, which give an approximation of the kinetics. The kinetics are shown on a reduced time axis, t' (in units of k_1 ; $t' = t k_1$). The fits give $k_f = k_f/k_1 = 0.024$, and $k_u = k_u/k_1 = 0.097$.

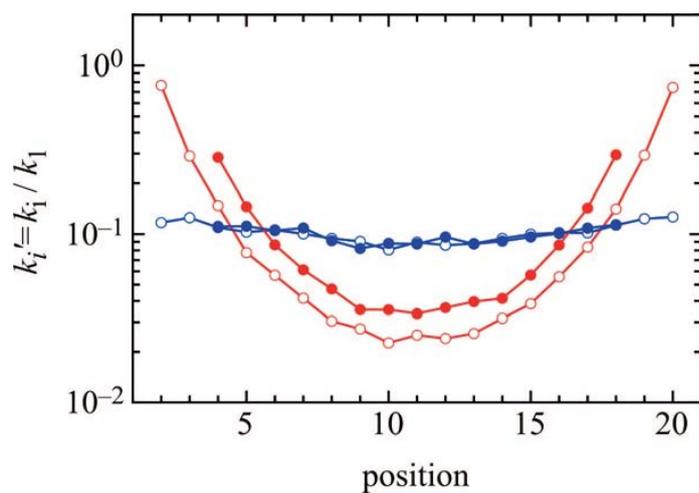


Fig. S4. Position dependence of the reduced folding and unfolding rate constants k_i (blue) and k_i^f (red) for a single segment (segment length = 1, open symbols) and segment length = 5 (closed symbols) obtained from the simulations. Segments of 5 residues were counted as closed if 4 or more residues were helical. The simulation parameters were chosen as described in Fig. 4 of the article.

Temperature Dependence of Local Equilibrium Folding and Unfolding of an α -Helix

Andreas Reiner¹, Beat Fierz¹, & Thomas Kiefhaber^{2*}

¹Biozentrum der Universität Basel, Division of Biophysical Chemistry,
Klingelbergstrasse 70, CH-4056 Basel, Switzerland

²Munich Center for Integrated Protein Science & Chemistry Department,
Technische Universität München, Lichtenbergstrasse 4, D-85747 Garching, Germany

*corresponding author. Phone: ++49 89 289 13420; fax: ++49 89 289 13416;
e-mail: t.kiefhaber@tum.de

Abstract

Local structure elements such as α -helices or β -hairpins can form in the absence of tertiary interactions and may hence be of importance in the early stages of protein folding. α -Helical structures in peptides are only marginally stable with the exception of alanine-based helices. Despite their structural simplicity and considerable experimental and theoretical efforts, the dynamics of these α -helical peptides remain poorly understood. We recently used triplet-triplet energy transfer to measure local equilibrium folding and unfolding at different positions within helical peptides. Here we extend this approach to investigate the effect of temperature on local helix dynamics by placing triplet labels with $i, i + 6$ spacing in the centre of an alanine-based model helix. The data are interpreted in terms of a classical two-state formalism and are compared to a more realistic multi-state description of the system, obtained from a kinetic Ising model. The results from the experiments suggest that the helix folding is an entropy controlled process, whereas helix unfolding encounters a mainly enthalpic barrier. Comparison with the simulations indicates that the apparent low activation energy for helix folding is a result of the multi-state character of the system and that both, the addition and the removal of a helical segment at a helix boundary are associated with significant enthalpic barriers of about 40 kJ/mol.

Introduction

The folding of proteins is a process of remarkable cooperativity, usually an all-or-none transition. Most single domain proteins do not populate any equilibrium intermediates and their folding kinetics can be described as a transition over a single energy barrier. The microscopic events on the path from the unfolded state to the fully folded structure remain therefore hidden and can only be assessed indirectly. The discovery and development of peptide sequences that form stable α -helices [1-4] or β -hairpins [5] in aqueous solution allows to study the formation of these basic structures individually and gives a new perspective to understand the process of structure formation. Isolated secondary structure elements have also been observed in unfolded states and folding intermediates of several proteins. Their local character suggests that they play a key role in the early stages of protein folding. A fundamental difference between secondary structure formation and protein folding is the lower cooperativity of secondary structure elements, which should make it possible to dissect different microscopic events.

The thermodynamics of α -helix formation have been worked out in the 1950s in the context of the helix-coil transition in long homopolymers [6-10]. The use of statistical weights accounting for the energetics of nucleation (boundary positions) and elongation (interior positions) has also proven to describe the stability of α -helical structures in small peptides quite well [2-4,11]. The concept of nucleation and elongation is also commonly used to discuss the kinetics of the helix-coil transition [12-17], although a direct experimental validation remains elusive. A number of molecular dynamic simulations gives little support for the sharp distinction between nucleation and elongation, but points to the importance of alternative conformations and non-native H-bonds at the initial stages of helix formation, which subsequently rearrange to the canonical structure [18-20].

The most widely used techniques to study the dynamics of helical systems are relaxation experiments. Early studies on long homopolymers made use of dielectric [21] and ultrasonic relaxation [22], which gave relaxation times of several hundred of nanoseconds. The kinetic theory proposed by Schwarz [12] relates the observed mean relaxation times to the time constant of helix growth, placing it on the timescale of 0.1 –

10 ns [22]. Short alanine-based model helices have mainly been studied by laser-induced temperature-jump (*T*-jump) measurements and sub-microsecond time resolution has been attained in a number of experiments [16,23-34]. The use of fluorophores [16,35] or the incorporation of isotope labels for IR [28,29,32] or UV resonance Raman detection [33,34] even enables site specific measurements. However, the *T*-jumps are often small and thus the multi-state equilibrium is only slightly shifted towards the coil state. The interpretation of the relaxation rate constants is difficult, as they cannot be attributed to specific events. Furthermore, in multi-state systems such as α -helices it is difficult to separate the contributions of forward and backward reactions to the relaxation behaviour.

An alternative approach to study helix dynamics is to extract kinetic information from thermal fluctuations without perturbing the helix-coil equilibrium. H/D-exchange, which was extensively used to directly probe the opening and closing of individual H-bonds in proteins, is unfortunately much too slow to capture the kinetics in α -helical peptides. By this technique only the local equilibrium stabilities can be obtained, which confirm the fraying of helices towards their termini [36]. Microsecond dynamics were detected in a ^{13}C -NMR relaxation study, but the microscopic origin of the exchange process remained unclear [37]. Alternatively, a probing reaction can be coupled to helix unfolding that is fast enough to compete with refolding. Hofrichter and coworkers employed triplet-quenching to study global helix unfolding in an alanine-based model helix [38]. We recently used triplet-triplet energy transfer (TTET) [39-41] to probe the local dynamics along a helical peptide of 21 residues [42]. Folding was found to have similar time constants ($1/k_f$) of about 400 ns at all positions, whereas the time constants for unfolding, ($1/k_u$), varied between 1.4 μs in the helix centre to 250 ns at the termini (at 5 °C). In addition, we addressed the influence of urea on the helix dynamics and were able to rationalize the results with Monte-Carlo simulations of a kinetic Ising model.

Here we report the temperature dependence of local helix folding and unfolding in the centre of an α -helical peptide. Despite the wealth of *T*-jump data on the helix-coil transition, the role of enthalpic contributions is intensely discussed [33,34]. We interpret our data in two different ways. First, we discuss the local dynamics in terms of apparent two-state behaviour, neglecting the influence of the other parts of the helix. This allows

comparison with previous experimental results from T -jump experiments. Second, we compare our results with a more realistic description of the multi-state system, obtained from a kinetic Ising model addressing the effect of temperature on the microscopic rate constants for helix growth and unwinding. Surprisingly, there are major differences in activation energies and enthalpies between the macroscopic and microscopic rate constants.

Results and Discussion

Measurement of local helix folding and unfolding.

We previously used TTET between i and $i + 6$ spaced triplet donor and acceptor labels to study local helix folding and unfolding [42]. To investigate the effect of temperature on helix dynamics we used an alanine-based helical peptide, in which the triplet donor xanthone (Xan) and the triplet acceptor naphthylalanine (Nal) are placed around the helix centre (Xan7-Nal13) [42]. The helix content reaches a maximum in the central part [11], which allows to investigate the dynamics of the helix-coil transition over a wide range of conditions. The sequence of the 21 residue peptide is Ac-AAAAA-AXanARA-AANalRA-AAARA-A-NH₂, for details see Material and Methods.

The peptide forms α -helical structure as indicated by its Far-UV CD spectrum, showing a characteristic maximum of the ellipticity at 190 nm and minima around 208 nm and 222 nm (Figure 1A), as well as a pronounced melting transition (Figure 1B). Due to unknown contributions of the labels Xan and Nal in the Far-UV it is not possible to determine the helix content of the peptide accurately. The $i, i + 6$ -spaced labels have in general a slight destabilising effect [42], especially when placed in the helix centre, which is most sensitive towards perturbations [11].

Figure 1

To characterise the dynamics in the helix centre, we measured TTET in the temperature range from 0 – 30 °C. Briefly, a short laser pulse (4 ns at 355 nm) is used to excite Xan to the singlet state, from where it undergoes fast and efficient intersystem crossing to the triplet manifold [41]. The triplet state resides on Xan (Eq. 1), as long as helical structure

is formed in this region, since the $i, i + 6$ spacing of the labels places them on opposite sides of the helix and prohibits contact formation [42].



Upon unfolding of the segment between the labels (k_u), the labels become free to form contact with a typical rate constant of contact formation (k_c), as long as the helical structure is not reformed (k_f). In unfolded model peptides the time constant $1/k_c$ for formation of an $i, i + 6$ loop was shown to be around 30 ns. In the mechanism depicted in Equation 1 contact formation measured by TTET serves as probing reaction, coupled to the helix-coil transition. This approach is similar to H/D-exchange experiments, which probe opening and closing of single H-bonds in thermodynamic equilibrium. There often single exponential kinetics are observed, since only the folded state is populated to a significant amount ($k_f \gg k_u$). Furthermore, usually conditions are chosen in which the H/D-exchange is either much faster than k_f and k_u to obtain k_u directly (EX1), or much slower than k_f and k_u , which allows to calculate the equilibrium constant $K_{eq} = k_u/k_f$ (EX2). In our case k_f , k_u and k_c are in a similar range and it is possible to determine all three microscopic rate constants from the two apparent rate constants and the corresponding amplitude information.

Figure 2 shows TTET kinetics between Xan and Nal placed in the centre of the 21 residue model helix. The time course of TTET can be directly measured by the decrease of the Xan triplet absorption band at 590 nm (Figure 2A) or the concomitant increase of the Nal triplet absorption band at 420 nm (data not shown). The Xan triplet decay can be well described by the sum of two exponential functions at all temperatures, which indicates that both, the helical state (H) and the coil state (C) are populated to significant amounts in equilibrium. With increasing temperature the Xan triplet absorption decays become faster (Figure 2A). The apparent rate constants λ_1 and λ_2 and the corresponding amplitudes are shown in Figure 2. Both, λ_1 and λ_2 increase with temperature and the faster process gains amplitude.

Figure 2

We additionally varied the urea concentration at each temperature between 0 M and 8 M, which destabilizes the helical state [43]. This allows to get reliable values for the microscopic rate constants k_u , k_f and k_c [42]. As example the urea dependencies at the extreme temperatures of 0 °C and 30 °C are shown in Figure 3.

Figure 3

Again all decay curves can be well described with the sum of two exponential functions with the apparent rate constants λ_1 and λ_2 and the amplitudes shown in Figure 3B and Figure 3C, respectively. For global fitting of all kinetics obtained at a single temperature we assume that the logarithms of k_f , k_u and k_c are linear functions of the molar concentration of urea (Eq. 2) [42].

$$\ln k_i = \ln k_i^0 - \frac{m_i \cdot [\text{urea}]}{RT} \quad (2)$$

Results of the global fitting procedure using the analytic solution of Eq. 1 (see Material and Methods) are shown for the kinetics at 0 °C and 30 °C (Figure 3). The apparent rate constants and amplitudes calculated from the global fits agree well with those determined from fitting the traces individually. The obtained microscopic rate constants k_f , k_u and k_c are shown as well. The contact formation process k_c is always faster than local folding k_f and unfolding k_u . In the absence of denaturant and low temperature (0 °C) folding $1/k_f^0 = (775 \pm 65)$ ns is around three times faster than unfolding $1/k_u^0 = (2220 \pm 90)$ ns. At 30 °C the local stability is reduced and folding $1/k_f^0 = (348 \pm 57)$ ns is 1.5 fold slower than unfolding $1/k_u^0 = (229 \pm 19)$ ns. At 0 °C helix unfolding is accelerated by urea ($m_u < 0$), but it is decelerated at 30 °C ($m_u > 0$). A full representation of the double perturbation experiment is given in Figure 4. Figure 4A shows an Arrhenius plot of k_f^0 , k_u^0 and k_c^0 , the rate constants extrapolated to 0 M urea. Figure 4B gives the corresponding m_i -values.

Figure 4

Similar results as shown in Figure 4A have been reported from a triplet-quenching experiment [38]. In that study a triplet label and a triplet quencher had been placed at the opposite termini of a 22 residue peptide, i.e. global helix dynamics were measured, whereas in our experiment local dynamics are probed with $i, i + 6$ spaced labels. The global probing complicates the interpretation, since contact formation can occur as soon as a substantial amount of the peptide is unfolded, but complete unfolding is not required. The authors do not further discuss the helix-coil dynamics but focus on the contact formation in the coil state.

Contact formation in the open state (k_c), which is used as probing reaction for the helix-coil transition (Eq. 1), is at all temperatures at least four-fold faster than k_f and k_u and shows linear Arrhenius behaviour with an apparent activation energy of 33 kJ mol^{-1} (Figure 4A). This value is in agreement with activation energies found for contact formation in unstructured polypeptides [44] and can be partly attributed to the change of solvent viscosity with temperature. In model peptides k_c has been shown to depend strongly on the inverse of the solvent viscosity η (Eq. 3 with $\beta = -0.7$ to -1.0) [39,44,45], as expected for diffusion-controlled reactions [46].

$$k_c = k_c^0 \cdot \left(\frac{\eta}{\eta_0} \right)^\beta \quad (3)$$

In the studied temperature range the change of solvent viscosity causes in first approximation linear Arrhenius behaviour with an apparent activation energy of 17.7 kJ mol^{-1} for $\beta = -1$ [47]. Correction for this value reduces the observed activation energy of k_c to 15 kJ mol^{-1} . The use of a two-state model as shown in Eq. 1 is justified to analyse our TTET data and to obtain rate constants for local helix folding and unfolding, since we monitor two macroscopic states of the system. The open state includes all conformations allowing contact formation between labels, whereas the folded state comprises all conformations prohibiting contact formation. For the interpretation of the temperature dependence we first keep the local two-state formalism, ignoring the influence of the other parts of the peptide. In the second part we will compare our results

to a more realistic description that takes the multi-state character of the whole system into account.

Interpretation of k_u and k_f in terms of a two-state description.

Helix formation in the centre of the peptide varies only moderately with temperature, with time constants $\tau = 1/k_f^0$ between (348 ± 57) ns and (775 ± 65) ns at 0 °C and 30 °C, respectively. Unfolding in contrast is accelerated by factor of ten, from $1/k_u^0 = (2220 \pm 90)$ ns at 0 °C to (229 ± 19) ns at 30 °C (Figure 4A). At around 22 °C the rate constants of helix folding k_f and unfolding k_u become equal, i.e. the equilibrium constant for local unfolding in the helix centre are unity at this temperature. The rate constant for helix formation, k_f , shows non-linear Arrhenius behaviour (Figure 4A) indicating a significant ΔC_p^\ddagger for this process. This is a unexpected finding since the equilibrium constant for helix formation was shown to have a linear temperature dependence, indicating an equilibrium ΔC_p close to 0. A less pronounced, but still significant curvature is also observed for k_u . In protein folding non-Arrhenius behaviour can be attributed to the differences in the heat capacities of the transition state and the folded or unfolded state [48-50]. The rate constants k_f and k_u are then expressed by the corresponding differences in enthalpy $\Delta H_i^{0\ddagger}$, entropy $\Delta S_i^{0\ddagger}$ and heat capacity $\Delta C_{p,i}^{0\ddagger}$ between the ground state and the transition state at a reference temperature T^0 according to Eq. 4.

$$\begin{aligned}
 k_i &= k_i^0 \cdot \exp\left(-\frac{\Delta G_i^{0\ddagger}}{RT}\right) = \\
 &= k_i^0 \cdot \exp\left[-\frac{\Delta H_i^{0\ddagger}(T^0)}{RT} + \frac{\Delta S_i^{0\ddagger}(T^0)}{R} - \frac{\Delta C_{p,i}^{0\ddagger}}{RT}\left(T - T^0 - T \ln \frac{T}{T^0}\right)\right]
 \end{aligned} \tag{4}$$

In the Arrhenius plots this leads to a convex curvature for k_f ($\Delta C_p^{0\ddagger}_{U-TS} < 0$) and a concave curvature for k_u ($\Delta C_p^{0\ddagger}_{N-TS} > 0$). Usually a major part of the non-polar surface is already buried in the transition state and hence the curvature for folding is typically more pronounced than the curvature for unfolding ($|\Delta C_p^{0\ddagger}_{U-TS}| > |\Delta C_p^{0\ddagger}_{N-TS}|$).

The temperature dependence of local helix folding k_f and unfolding k_u can be well described by Eq. 4 (solid lines in Figure 4A). The parameters obtained from a non-linear least-square fit are summarized in Table 1. It should be noted that $\Delta H_i^{0\ddagger}$ represents the actual enthalpy difference between ground state and transition state, whereas $\Delta S_i^{0\ddagger}$, and consequently also $\Delta G_i^{0\ddagger}$, are apparent values depending on the pre-exponential factor in Eq. 4. For folding reactions involving motions of a peptide chain in solution k_i^0 is likely in the range of $10^7 - 10^8 \text{ s}^{-1}$ [40]. The values in Table 1 were calculated with $k_i^0 = 5 \cdot 10^7 \text{ s}^{-1}$, which is the rate constant for formation of an $i, i + 4$ contact in unfolded polyserine chains [40]. The values obtained for $\Delta H_i^{0\ddagger}$, $\Delta S_i^{0\ddagger}$ and $\Delta C_{p,i}^{0\ddagger}$ are associated with rather large standard deviations (Table 1), but they nevertheless describe the nature of the barriers. At 25 °C unfolding is clearly limited by a large enthalpic barrier of $\sim 40 \text{ kJ mol}^{-1}$, close to the change in enthalpy for equilibrium unfolding. This is consistent with the view that the rate-limiting step for helix unfolding includes the breaking of H-bonds, which appear to be fully lost in the transition state. However, only one third of the overall entropy change upon unfolding has been gained at this point. It is therefore likely that the ϕ, ψ -angles are still severely restricted in the transition state, compared to the coil state. Helix formation at 25 °C, in contrast, appears to be purely limited by entropic barriers. For the equilibrium between the helical state and unfolded state the gain in enthalpy and the loss in entropy largely compensate each other and at 25 °C a ΔG_{eq}^0 of $\sim 1.4 \text{ kJ mol}^{-1}$ remains.

For both reactions a negative $\Delta C_{p,i}^{0\ddagger}$ is obtained (Table 1), also evident from the convex curvatures in the Arrhenius plot (Figure 4A). The determined values, $\Delta C_p^{0\ddagger}_{\text{U-TS}} = (-2.0 \pm 0.8) \text{ kJ mol}^{-1} \text{ K}^{-1}$ and $\Delta C_p^{0\ddagger}_{\text{N-TS}} = (-1.1 \pm 0.4) \text{ kJ mol}^{-1} \text{ K}^{-1}$ are associated with large errors, allowing only a rough estimate of the heat capacity change upon unfolding in the helix centre ($\Delta C_{p,\text{eq}}^0 = (0.9 \pm 1.0) \text{ kJ mol}^{-1} \text{ K}^{-1}$). From calorimetric measurements, which found a temperature independent ΔH^0 [51,52] it was estimated that $\Delta C_{p,\text{eq}}^0$ has to be below $\pm 0.032 \text{ kJ mol}^{-1} \text{ K}^{-1}$ per residue. A large $\Delta C_{p,\text{eq}}^0$ for unfolding of the segment between the $i, i + 6$ spaced labels is therefore not expected from the calorimetric measurements. Since $\Delta C_{p,i}^{0\ddagger}$ of both, the forward and the backward reaction are negative, the transition state appears to have a lower heat capacity than both, the coil state and the helical state.

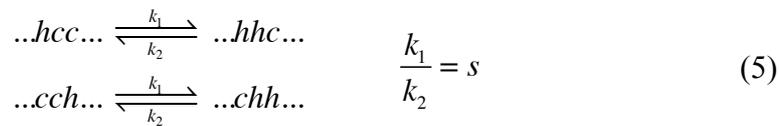
As already discussed for k_c , the change of the water viscosity with temperature has to be taken into account as well. In first approximation $\ln(\eta_0/\eta)$ is a linear function of $1/T$ and $\Delta H_i^{0\ddagger}$ might have to be reduced by up to 17.7 kJ mol^{-1} [47], in the case of a fully viscosity dependent reaction ($\beta = -1$ in Eq. 3). The small deviation from non-linear behaviour leads to a small apparent $\Delta C_{p,i}^{0\ddagger}$ of minimally $-0.15 \text{ kJ mol}^{-1} \text{ K}^{-1}$, which is by far not sufficient to explain the observed curvatures. Temperature significantly affects the sensitivity of the kinetics towards urea (Figure 4B). At low temperatures urea accelerates unfolding and at high temperatures it reduces k_u with $m_u(0^\circ) = (-246 \pm 54) \text{ J mol}^{-1} \text{ M}^{-1}$ and $m_u(30^\circ\text{C}) = (261 \pm 101) \text{ J mol}^{-1} \text{ M}^{-1}$, respectively. In contrast, the sensitivity of k_f towards urea remains unchanged or increases only slightly with temperature ($m_f \approx 880 \text{ J mol}^{-1} \text{ M}^{-1}$). The m -value for contact formation k_c does not change with temperature ($m_c \approx 220 \text{ J mol}^{-1} \text{ M}^{-1}$) and is close to the value found earlier in the coil state of helical peptides [42].

These results show that a two-state model is able to describe the temperature dependence of local helix folding and unfolding, as it is also often done in the interpretation of T -jump experiments. Isolated α -helices are however systems of low cooperativity and should therefore be examined with a formalism more appropriate to the helix-coil transition.

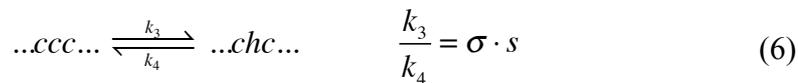
Effect of temperature on the local dynamics of a multi-state system.

Our experiments report on a microscopically well-defined event, namely local helix unfolding/folding dynamics between an ensemble of helical conformations, in which contact formation between the side chain labels is prevented, and an ensemble of open conformations, in which contact formation can occur. We have previously used a linear Ising model to simulate this processes in discrete time steps [42]. This gave closer insight into the dynamics of the helix-coil transition and fully reproduced the local equilibrium folding and unfolding behaviour at different positions within the helix. Furthermore, the model revealed that different effects of urea on k_f and k_u at different positions are a consequence of the multi-state equilibrium. Here we extend these simulations to include the effect of temperature, but limit ourselves to the dynamics in the helix centre.

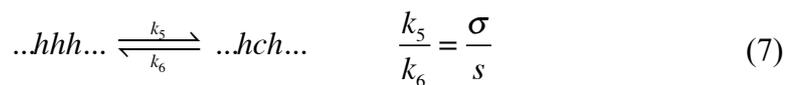
In the model the peptide is treated as a chain of 21 identical segments, which are either in the helical state h or the coil state c . Cooperativity is introduced by assigning statistical weights to individual segments, based on the state of the nearest neighbours, as it is also done by the widely used 2×2 approximation of Zimm and Bragg [8]. Helical segments having at least one neighbour in the helical conformation get a weight of s (elongation) and helical segments without helical neighbours a weight of $\sigma \cdot s$ (helix nucleation). Coil segments with two helical neighbours get a weight of σ/s (coil nucleation) and all other coil segments a statistical weight of *unity*. By this the system is divided into independent triplets. Eq. 5 – 7 summarize the possible transitions of the central segment, which have been discussed in detail by Schwarz [12]. For alanine-based model peptides $\sigma = 0.003$ and an average value of $s(0^\circ\text{C}) = 1.35$ are reasonable parameters [53]. The main processes in a helix of our length and stability are the addition (k_1), respectively the removal (k_2) of helical segments at the boundary of an existing helical segment (Eq. 5).



Next to this, helix nucleation (k_3) can occur, which is slow compared to the reverse reaction (k_4), i.e. isolated nuclei are very unstable (Eq. 6).



As third process the nucleation of coil within a helical segment (k_5) has to be considered, which is much slower than the fusion of helices (k_6) (Eq. 7).



The ratios of the forward and backward reactions are given by σ and s , for which the temperature dependencies are known. The nucleation parameter σ reflects the entropy

loss in order to form the first helical H-bond and is therefore generally assumed to be temperature independent [8,10]. Equilibrium measurements on both homopolymeric [7] and alanine-based model helices [53] are fully consistent with this assumption and hence the temperature dependence of the thermodynamic stability is solely described by $s(T)$ (Eq. 8).

$$\ln s(T) = \ln s(T^0) + \frac{\Delta H_{\text{eq}}^0(T^0)}{R} \left(\frac{1}{T} - \frac{1}{T_0} \right) + \frac{\Delta C_{p,\text{eq}}^0}{R} \left(1 - \frac{T_0}{T} - \ln \frac{T}{T_0} \right) \quad (8)$$

We chose $\sigma = 0.003$ and an average value for $s(0^\circ\text{C}) = 1.35$ [53], but the outcome of our simulations does not critically depend on the exact values. ΔH_{eq}^0 of helix unfolding is consistently reported to be around 4 kJ mol^{-1} [7,51-53]. We used the value of Scholtz *et al.* ($\Delta H_{\text{eq}}(25^\circ\text{C}) = 4.2 \text{ kJ mol}^{-1}$ [53]). Regarding the heat capacity change some controversy exists [51,52,54], but all studies indicate that $\Delta C_{p,\text{eq}}^0$ of helix unfolding is small, likely below $\pm 0.032 \text{ kJ mol}^{-1} \text{ K}^{-1}$. For the simulations we set in general $\Delta C_{p,\text{eq}}^0 = 0$, but later tested also the effect of $\Delta C_{p,\text{eq}}^0 = \pm 0.1 \text{ kJ mol}^{-1} \text{ K}^{-1}$ (see below) to account for our experimental observations (Figure 4).

Next to the equilibrium properties the rate constants k_1, k_2, \dots, k_6 in Eq. 5 - 7 have to be defined. For the rate constant of helix propagation k_1 we write (Eq. 9):

$$k_1(T) = k_1(T^0) \cdot \exp \left[\frac{\alpha_H \cdot H_{\text{eq}}(T^0)}{R} \left(\frac{1}{T} - \frac{1}{T_0} \right) + \frac{\alpha_C \cdot \Delta C_{p,\text{eq}}^0}{R} \left(1 - \frac{T_0}{T} - \ln \frac{T}{T_0} \right) \right] \quad (9a)$$

For the rate constant of the reverse reaction k_2 of helix unwinding it follows directly:

$$\begin{aligned} k_2(T) &= \frac{k_1(T)}{s(T)} \\ &= \frac{k_1(T^0)}{s_0} \cdot \exp \left[-\frac{(1 - \alpha_H) \cdot H_{\text{eq}}(T^0)}{R} \left(\frac{1}{T} - \frac{1}{T_0} \right) - \frac{(1 - \alpha_C) \cdot \Delta C_{p,\text{eq}}^0}{R} \left(1 - \frac{T_0}{T} - \ln \frac{T}{T_0} \right) \right] \quad (9b) \end{aligned}$$

$k_1(T^0)$ serves as basic rate constant that is used in the expression of the other rate constants as well. This “pre-exponential” factor also includes the activation entropy $\Delta S_i^{0\ddagger}$, which does not contribute to the temperature dependence. The changes of enthalpy $\Delta H_i^{0\ddagger}$ and heat capacity $\Delta C_{p,i}^{0\ddagger}$ between the ground states and the transition state are given explicitly. The parameters $\Delta C_{p,i}^{0\ddagger}$ and $\Delta H_i^{0\ddagger}$ for the forward and reverse reaction are linked by α_C and α_H , respectively. These α -values are in the range of 0 and 1, if the heat capacity (or enthalpy) of the transition state is in between the heat capacities (or enthalpies) of the ground states [55]. For α_H this is less likely, since most reactions encounter enthalpic barriers, corresponding to $\alpha_H < 0$. For $\Delta C_{p,i}^{0\ddagger}$ the experimental results also indicate that α_C might adopt a non-classical value. Eq. 10 and Eq. 11 describe the rate constants for helix nucleation and coil nucleation, respectively.

$$k_3 = \gamma_h \cdot \sigma \cdot k_1 \quad (10a)$$

$$k_4 = \gamma_h \cdot k_2 \quad (10b)$$

$$k_5 = \gamma_c \cdot \sigma \cdot k_2 \quad (11a)$$

$$k_6 = \gamma_c \cdot k_1 \quad (11b)$$

The scaling factors γ_h and γ_c account for the fact that in principle these processes could have completely different kinetics compared to k_1 and k_2 . As pointed out by Schwarz, it is however reasonable to assume that γ_h and γ_c are between 1 and $\sigma^{-1/2}$ [12]. Since the results of our simulations do not depend significantly on the actual values of γ_h and γ_c [42], we set $\gamma_h = \gamma_c = 1$.

The details of the simulation have been given elsewhere [42]. Briefly, the time-dependent change of the system is simulated in discrete time steps. Each time step is much shorter than $1/k_1, 1/k_2, \dots, 1/k_6$. The transition probability for every segment is calculated according to the state of the neighbouring residues (Eq. 5 - 7 and Eq. 9 - 11). This probability is compared to randomly generated numbers in order to decide, whether the particular transition takes place or not. After evaluation of all triplets all changes resulting from the transitions within the time step are made in parallel, giving the new conformation. In order to calculate helix folding and unfolding at a certain position, the first passage times until a segment in the coil state changes to helix (folding),

respectively from helix to coil (unfolding), are measured and collected in a histogram. The first passage time distributions are approximated by single exponential functions to give k_f and k_u . The size of the probed part can be extended from single residues to larger stretches. Choosing a probing segment of five residues does not change the basic outcome of the simulations, but gives a slightly better representation of our experiment, in which the labels are attached at positions $i, i + 6$ [42]. The results of the simulations with different parameter sets are summarized in Figure 5. The most simple case is a probing segment length of one, i.e. only opening and closing of the central residue is measured, with $s(0\text{ °C}) = 1.35$, $\sigma = 0.003$, $\Delta H_{\text{eq}}(25\text{ °C}) = 4.2\text{ kJ mol}^{-1}$ and $\alpha_H = 0$. The results are given as reduced rate constants $k_f / k_1(T^0)$ and $k_u / k_1(T^0)$, since this is sufficient to discuss the temperature dependencies. Comparison of experiment and simulation has shown that $k_1(5\text{ °C})$ is $\approx 2 \cdot 10^7\text{ s}^{-1}$ [42].

Figure 5

For all parameter sets used in the simulation $k_f / k_1(T^0)$ and $k_u / k_1(T^0)$ become equal at the same temperature (Figure 5), which is expected, since $\Delta H_{\text{eq}}(25\text{ °C}) = 4.2\text{ kJ mol}^{-1}$ and $\Delta C_{p,\text{eq}}^0 = 0$ were chosen in all cases. In the experiment the midpoint of thermal unfolding in the helix centre is around 22 °C (Figure 4A), in the simulation around 30 °C . The small difference might be explained by either a somewhat too small ΔH_{eq}^0 or, more likely, a somewhat too large $s(T^0)$ chosen for the simulations. It should be noted that $s(0\text{ °C}) = 1.35$ is an average value for alanine-based peptides with a lysine in every fifth position [53], whereas our peptide is somewhat less stable due to the incorporation of the triplet donor and acceptor labels.

With increasing temperature $k_u / k_1(T^0)$ becomes faster ($E_a \approx 30\text{ kJ mol}^{-1}$), whereas $k_f / k_1(T^0)$ becomes significantly slower. In other words, helix folding is predicted to have an apparently negative activation energy of $\approx -40\text{ kJ mol}^{-1}$. Essentially similar curves are found, when the probed segment is extended to five residues. Changing α_H from 0 to 1, i.e. attributing the whole change in s to a temperature dependence of k_1 instead of k_2 , has no drastic effect. For $\alpha_H = 1$ the activation energies of $k_f / k_1(T^0)$ and $k_u / k_1(T^0)$ are simply lower by 4.2 kJ mol^{-1} (Figure 5). The apparent negative activation energy of around -40 kJ mol^{-1} found for $k_f / k_1(T^0)$ does hence not depend on the parameters chosen to model the temperature dependence, but is clearly a consequence of

the multi-state character of the system. This can be rationalized as with increasing temperature the helix is destabilised in general, slowing down local refolding. This prediction is in contrast to the experiment, in which k_f gets accelerated with temperature, at least at low temperatures (Figure 4A). In order to reconcile this difference, an additional enthalpic barrier has to be assumed for both, helix propagation k_1 and helix unwinding k_2 (Eq. 9). The additional enthalpy barrier is estimated to be in the range of 30 - 50 kJ mol⁻¹, which would roughly correspond to $\alpha_H \approx -8$ in the simulation (Figure 5). A full agreement between experiment and simulation cannot be obtained, since the activation energies of $k_f / k_1(T^0)$ and $k_u / k_1(T^0)$ always differ by ~ 65 kJ mol⁻¹, whereas the difference in the experiment is significantly lower. A similar discrepancy was observed, when the effect of urea on local helix folding and unfolding was simulated [42]. In that case the predicted difference in the m -values of k_f and k_u was about 30 % larger than the experimentally found difference. As already discussed for the two-state description, the observed activation enthalpies can be attributed in part to a change of the water viscosity with temperature, up to 17.7 kJ mol⁻¹ for fully viscosity dependent reactions. A T -jump study on the viscosity dependence of helix dynamics suggests $\beta = -0.6$ (Eq. 3) [27], corresponding to 10.6 kJ mol⁻¹.

A convex curvature is observed in the Arrhenius plot of $k_f / k_1(T^0)$ (Figure 5), similar to the experimental observation (Figure 4A). Indeed, it is conceivable that the curvatures arise from the multi-state character of the system. However, the kinetics obtained for helix folding are approximated, but not well described by a single rate constant. The curvature might therefore be an artefact of the fitting procedure, as no deviation from linear behaviour is observed for k_u , for which the first passage time distributions are better represented by single exponential functions. In the experiments we also observe a slight curvature for k_u .

Other causes for the curvatures of k_f and k_u might have to be considered as well, such as heat capacity changes as discussed above (Eq. 9). Negative heat capacities for helix unfolding, tested with $\Delta C_{p,\text{eq}}^0 = -0.1$ kJ mol⁻¹ K⁻¹ and $0 < \alpha_C < 1$, introduce a convex curvature in $k_u / k_1(T^0)$ of similar strength as experimentally observed, but lead to an apparently sigmoidal behaviour of $k_f / k_1(T^0)$ (data not shown). Positive heat capacities, tested with $\Delta C_{p,\text{eq}}^0 = 0.1$ kJ mol⁻¹ K⁻¹, lead to concave curvatures in the Arrhenius plot of

$k_u / k_1(T^0)$, irrespective of whether α_C is chosen to be zero, one or intermediate (data not shown). With $\alpha_C = 2.2$, corresponding to a transition state with a lower heat capacity than both ground states, as suggested by the two-state evaluation, a significant concave curvature of $k_u / k_1(T^0)$ remains, whereas k_f appears to become more convex. However, the assessment of k_f remains problematic, as mentioned above. The simulations do not allow to draw sound conclusions on whether the observed curvature is due to the behaviour of the multi-state system, the heat capacity of the transition state or other factors. The deviation of the water viscosity from Arrhenius-like behaviour [47] is also within the multi-state model insufficient to explain the observed curvatures.

Enthalpic barriers in helix folding and unfolding.

Since many studies on the kinetics of the helix-coil transition used T -jump measurements, a wealth of data on activation enthalpies exists. The measured relaxation rate constants show in most cases Arrhenius behaviour with high activation enthalpies [25,28,30,31]. This is in some contrast to the theory of Schwarz [12], predicting a minimum of the mean relaxation rate constant at the midpoint of denaturation, i.e. concave Arrhenius behaviour might be expected for the studied temperature ranges. Similar to our experiment, high activation enthalpies of the basic processes such as helix elongation (k_1) and unwinding (k_2) would explain why this is not observed experimentally. The interpretation of relaxation rate constants in terms of a two-state model should be avoided, especially for data that were obtained with global probes. “Folding” and “unfolding” do not correspond to any defined events in this case. However, it is not surprising that negative activation enthalpies for “folding” are observed [24], as in our case helix formation (k_f) is predicted to have an apparently negative activation enthalpy due to the linked equilibria at each helical residue, which shift towards the helical state with decreasing temperature and thus favour helix formation. For more local probes the use of a two-state approximation might be justified to extract data on the local stability and dynamics. It is however important to note that for a comparison of results, e.g. obtained for different positions or experimental conditions, the full system should be taken into account. In our case, a classical two-state interpretation of the temperature-dependence of central helix folding and unfolding indicates that folding does not encounter any enthalpic barrier (Table 1). The comparison with predictions from a multi-state model reveals that the apparent absence of an activation enthalpy is very likely the result of a substantial temperature

dependence of both helix elongation and unwinding (Figure 5). Similarly, we earlier found that the dynamics observed at different positions within a helix can be understood, when the complete system is taken into account [42]. It would be interesting to see, whether also the data from a recent *T*-jump experiment, in which deuterated amino acids were used to probe the dynamics in different regions [33], could be explained in the framework of a multi-state model. Negative activation enthalpies were reported for relaxation at the helix ends and strongly positive activation enthalpies for the helix centre. The authors take this as evidence for alternative conformations, being in equilibrium with the canonical α -helix conformation but showing different melting behaviour.

A more elaborate analysis was used by Thompson *et al.* who measured the relaxation in the N-terminal region of a helical peptide using fluorescence techniques [16,35]. In these studies a minimum, i.e. concave Arrhenius behaviour of the relaxation rate constant was found near the midpoint of unfolding. The relaxation kinetics were analysed with a kinetic zipper model and high activation energies were found for helix elongation and unwinding [35]. The authors give an upper limit of 24 kJ mol⁻¹ for the activation energy after correcting for the change in solvent viscosity, similar to our observations for the dynamics in the central region of a helix.

High activation enthalpies for dynamics in helical peptides have also been observed with other experimental techniques. A NMR relaxation study on a ¹³C-labelled peptide revealed a microsecond exchange process with an activation energy of 25 kJ mol⁻¹ [37]. Activation energies of 31 kJ mol⁻¹ for folding and 37 kJ mol⁻¹ for unfolding were also found for conformational transitions in a helical peptide induced by a photoswitchable side chain cross-link [56].

We find that folding and unfolding in the centre of a helical peptide deviate from simple Arrhenius behaviour, showing a convex curvature (Figure 4). This is not readily explained by a change in the heat capacity, as it is often the case in protein folding reactions, but is likely a consequence of the multi-state character of the system. In the context of diffusional processes it was also suggested that non-Arrhenius behaviour might originate from complex free-energy surfaces [57,58].

Conclusions

In summary, our findings show that TTET coupled to the helix-coil equilibrium can be used to measure the temperature dependence of local helix folding and unfolding, as demonstrated for the centre of an alanine-based model peptide. Analysis of the temperature-dependence with a two-state model suggests an enthalpic barrier for unfolding, whereas folding appears to be limited by a purely entropic barrier. However, simulations based on a linear Ising model show that the whole system should be taken into account for the interpretation of the local dynamics, especially when different positions or the effect of a perturbation (e.g. temperature) are investigated. If we combine the experimental data with predictions from this multi-state model, we find that both, the addition and removal of single segments to an existing helix encounter a significant enthalpic barrier of around 40 kJ mol^{-1} . The temperature dependence of the solvent viscosity probably contributes to this activation enthalpy (up to 17.7 kJ mol^{-1}), but cannot fully account for it.

Material and Methods

Peptide synthesis, labelling and purification.

The peptide was synthesized using standard 9-fluorenylmethyloxycarbonyl (Fmoc) chemistry on an Applied Biosystems 433A synthesizer. Couplings were performed with 2-(1H-benzotriazole-1-yl)-*N,N,N',N'*-tetramethyl-uronium hexafluorophosphate and *N*-hydroxy-benzotriazole (HBTU/HOBt) on a Tentagel S RAM resin (Rapp Polymere). Nal was incorporated via Fmoc-protected 1-naphthylalanine (Bachem). The xanthone derivative 9-oxoxanthene-2-carboxylic acid, synthesized according to Graham and Lewis [59], was activated with benzotriazol-1-yl-oxytripyrrolidinophosphonium hexafluorophosphate (PyBOP) and coupled to the β -amino group of selectively deprotected α,β -L-diaminopropionic acid (Dpr). Methyltrityl was used as orthogonal side chain protection group, which was selectively removed with 3 % (v/v) trifluoroacetic acid (TFA) in dichloromethane.

Final cleavage from the resin and deprotection of the arginine side chains was achieved with 94/3/3 (v/v/v) TFA / triisopropylsilane / H₂O. The peptide was purified to > 97 % purity by preparative HPLC on a RP-8 column. Purity was checked by analytical HPLC and the identity verified by MALDI-TOF mass spectrometry.

Sample preparation and CD spectroscopy.

All measurements were performed in 10 mM potassium phosphate, pH 7.0. The peptide concentration was determined by the Xan absorption band at 343 nm ($\epsilon_{343} = 3900 \text{ M}^{-1} \text{ cm}^{-1}$). Urea concentrations were calculated from the refractive index [60]. CD measurements were performed on an Aviv DS62 spectropolarimeter. Spectra were recorded with $\sim 50 \mu\text{M}$ peptide in a 0.1 cm cuvette at 5.0 °C. Thermal unfolding transitions was measured at 222 nm with $\sim 5 \mu\text{M}$ peptide in a 1 cm cuvette.

TTET measurements and data evaluation.

TTET was measured with an Applied Photophysics LKS.60 Laser Flash Photolysis Reaction Analyzer. Xan triplet states were generated with a 4 ns laser pulse at 355 nm (Nd:YAG, Quantel Brilliant) and the absorbance bands of the Xan and Nal triplet states were monitored at 590 nm and 420 nm respectively. The measurements were performed in degassed solutions with $\sim 50 \mu\text{M}$ peptide. At this concentration TTET arising from intermolecular contact formation is slower than the intrinsic lifetime of the Xan triplet state, i.e. > 25 μs .

For every experiment two Xan triplet decays were averaged and the signal intensity was corrected for small differences in the peptide concentration. A sum of two exponential functions was used to describe the data ($A_1, \lambda_1; A_2, \lambda_2$).

The double exponential TTET decays and the change of the amplitudes upon urea addition agree with a three-state system as depicted in Eq. 1. The observed rate constants $\lambda_{1,2}$ are a function of the three microscopic rate constants k_f, k_u and k_c (Eq. 12).

$$\lambda_{1,2} = \frac{k_u + k_f + k_c \pm \sqrt{(k_u + k_f + k_c)^2 - 4k_u \cdot k_c}}{2} \quad (12)$$

The amplitudes $A_{1,2}$ of $\lambda_{1,2}$ are as well functions of k_f, k_u and k_c and furthermore depend on the starting concentrations, in our case given by the thermodynamic equilibrium $k_u / k_f = [C^*]/[H^*]$. The full formula for the amplitudes and their derivation is given elsewhere [42,61]. With this analytic solution and Eq. 2 the microscopic rate constants can be obtained in a global fit of all Xan triplet decays at a single temperature. All fitting procedures were performed with ProFit (QuantumSoft) using non-linear least-square methods.

Monte-Carlo simulations with a kinetic Ising model.

The basic model, the possible transitions (Eq. 5 - 7), the definition of the elementary rate constants k_1, k_2, \dots, k_6 (Eq. 9 - 11) and the choice of parameters are discussed above. Further details are also given in our previous work [42]. Similar to the method described by Poland [15] the simulation is performed in discrete time steps $\Delta t'$. $\Delta t'$ was chosen to be 0.05 times the fastest time constant of $1/k_1, 1/k_2, \dots, 1/k_6$.

The first and the last segment of the 21 residue long homopolymer are constantly set to be c . The actual starting conformation is not relevant, since the simulations run much longer than equilibrium fluctuations last. We start from a single helix (continuous stretch of h s in a row) that satisfies the expected average helix content [8]. All 19 triplets of the starting conformation are read and the probabilities $p_i = k_i \cdot \Delta t'$ for transitions of the central residues calculated (Eq. 5 - 7). These probabilities are compared to uniformly distributed random numbers r ($0 < r < 1$). If $k_i \cdot \Delta t' > r$ the particular transition occurs, otherwise the segment remains unchanged. After evaluation of all triplets, the changes

resulting from the transitions within the time step $\Delta t'$ are made in parallel to obtain the new conformation and the procedure is repeated.

The data shown in Figure 5 were obtained in three steps. First, we simulated for each condition 10^7 time steps to have a good representation of the helix in equilibrium. Second, we randomly picked starting conformations out of these simulations, started new simulations from the chosen conformations and measured the first passage times until the defined segment closed (folded) or opened (unfolded). The observations were stopped at these points and the first passage times stored in histograms (150 bins). This was repeated until 8000 points were collected for the first passage time histograms of folding and unfolding under all conditions. Probing segments were either defined as single residues (“segment length of one”) or as stretches of five residues, which were counted as helical, if four or more residues were h at a certain time (“segment length of five”). Third, we approximated the obtained histograms with single exponential functions in order to obtain the reduced rate constants $k_f / k_1(T^0)$ and $k_u / k_1(T^0)$.

The described operations were performed using the framework of Matlab 7 R14 (MathWorks Inc.) in combination with compiled C code for time consuming iterations. Random numbers were generated using a Lehmer type algorithm [62]. Fitting of single exponential functions to the first passage time distributions was performed using the non-linear least-square fitting Matlab algorithm *lsqnonlin*. The data were further analyzed using ProFit (QuantumSoft).

Acknowledgment

We thank Josef Wey for the synthesis of 9-oxoxanthene-2-carboxylic acid. This work was supported by grants from the Swiss National Science Foundation, the VolkswagenStiftung and by the Munich Center for Integrated Protein Science.

Abbreviations

CD: circular dichroism; TTET: triplet-triplet energy transfer; Xan: xanthonic acid attached to the β -amino group of α,β -L-diaminopropionic acid; Nal: naphthylalanine. $\lambda_{1,2}$ denote the experimentally observed apparent rate constants, k_f , k_u , k_c the underlying microscopic rate constants (Eq.1). The rate constants k_1 , k_2, \dots, k_6 describe the elementary steps in the kinetic Ising model (Eq. 5 - 7). All equilibrium parameters are given in the direction of helix unfolding.

References

1. Brown, J. E. & Klee, W. A. (1971). Helix-coil transition of the isolated amino terminus of ribonuclease. *Biochemistry* **10**, 470-476.
2. Marqusee, S., Robbins, V. H. & Baldwin, R. L. (1989). Unusually stable helix formation in short alanine-based peptides. *Proc. Natl. Acad. Sci. USA* **86**, 5286-5290.
3. Lyu, P. C., Marky, L. A. & Kallenbach, N. R. (1989). The role of ion pairs in α -helix stability: two new designed helical peptides. *J. Am. Chem. Soc.* **111**, 2733-2734.
4. Doig, A. J., Errington, N. & Iqbalsyah, T. M. (2005). Stability and design of α -helices. In *Protein Folding Handbook* (Buchner, J. & Kiefhaber, T., eds.), pp. 247-313. Wiley-VCH.
5. 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.
6. Schellman, J. A. (1955). The stability of hydrogen-bonded peptide structures in aqueous solution. *C.R. Trav. Lab Carlsberg Ser. Chim.* **29**, 230-259.
7. Zimm, B. H., Doty, P. & Iso, K. (1959). Determination of the parameters for helix formation in poly(γ -benzyl L-glutamate). *Proc. Natl. Acad. Sci. USA* **45**, 1601-1607.
8. Zimm, B. H. & Bragg, J. K. (1959). Theory of the phase transition between helix and random coil in polypeptide chains. *J. Chem. Phys.* **31**, 526-535.
9. Lifson, S. & Roig, A. (1961). On the theory of helix-coil transition in polypeptides. *J. Chem. Phys.* **34**, 1963-1974.
10. Poland, D. & Scheraga, H. A. (1970). *Theory of Helix-Coil Transitions in Biopolymers*, Academic Press, New York.
11. Chakrabarty, A., Schellman, J. A. & Baldwin, R. L. (1991). Large differences in the helix propensities of alanine and glycine. *Nature* **351**, 586-588.
12. Schwarz, G. (1965). On the kinetics of the helix-coil transition of polypeptides in solution. *J. Mol. Biol.* **11**, 64-77.
13. Poland, D. & Scheraga, H. A. (1966). Kinetics of the helix-coil transition in polyamino acids. *J. Chem. Phys.* **45**, 2071-2090.
14. Craig, M. E. & Crothers, D. M. (1968). Calculation of kinetic curves for the helix-coil transition of polypeptides. *Biopolymers* **6**, 385-399.
15. Poland, D. (1996). Discrete step model of helix-coil kinetics: distribution of fluctuation times. *J. Chem. Phys.* **105**, 1242-1269.

16. Thompson, P. A., Eaton, W. A. & 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.
17. Brooks III, C. L. (1996). Helix-coil kinetics: folding time scales for helical peptides from a sequential kinetic model. *J. Phys. Chem.* **100**, 2546-2549.
18. Bertsch, R. A., Vaidehi, N., Chan, S. I. & Goddard III, W. A. (1998). Kinetic steps for α -helix formation. *Proteins* **33**, 343-357.
19. Chowdhury, S., Zhang, W., Wu, C., Xiong, G. & Duan, Y. (2002). Breaking non-native hydrophobic clusters is the rate-limiting step in folding of an alanine-based peptide. *Biopolymers* **68**, 63-75.
20. Sorin, E. J. & Pande, V. S. (2005). Exploring the helix-coil transition via all-atom equilibrium ensemble simulations. *Biophys. J.* **88**, 2472-2493.
21. 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.
22. Gruenewald, B., Nicola, C. U., Lustig, A. & Schwarz, G. (1979). Kinetics of the helix-coil transition of a polypeptide with non-ionic side groups, derived from ultrasonic relaxation experiments. *Biophys. Chem.* **9**, 137-147.
23. Williams, S., Causgrove, T. P., Gilmanishin, R., Fang, K. S., Callender, R. H., Woodruff, W. H. & Dyer, R. B. (1996). Fast events in protein folding: helix melting and formation in a small peptide. *Biochemistry* **35**, 691-697.
24. Lednev, I. K., Karnoup, A. S., Sparrow, M. C. & Asher, S. A. (1999). Nanosecond UV resonance Raman examination of initial steps in α -helix secondary structure evolution. *J. Am. Chem. Soc.* **121**, 4076-4077.
25. Huang, C. Y., Klemke, J. W., Getahun, Z., DeGrado, W. F. & Gai, F. (2001). Temperature-dependent helix-coil transition of an alanine based peptide. *J. Am. Chem. Soc.* **123**, 9235-9238.
26. Huang, C. Y., Getahun, Z., Wang, T., DeGrado, W. F. & Gai, F. (2001). Time-resolved infrared study of the helix-coil transition using ^{13}C -labeled helical peptides. *J. Am. Chem. Soc.* **123**, 12111-12112.
27. 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.
28. Werner, J. H., Dyer, R. B., Fesinmeyer, R. M. & Andersen, N. H. (2002). Dynamics of the primary processes of protein folding: helix nucleation. *J. Phys. Chem. B* **106**, 487-494.
29. Huang, C. Y., Getahun, Z., Zhu, Y., Klemke, J. W., DeGrado, W. F. & Gai, F. (2002). Helix formation via conformation diffusion search. *Proc. Natl. Acad. Sci. USA* **99**, 2788-2793.
30. Wang, T., Du, D. & Gai, F. (2003). Helix-coil kinetics of two 14-residue peptides. *Chem. Phys. Lett.* **370**, 842-848.
31. Wang, T., Zhu, Y., Getahun, Z., Du, D., Huang, C. Y., DeGrado, W. F. & Gai, F. (2004). Length dependent helix-coil transition kinetics of nine alanine-based peptides. *J. Phys. Chem. B* **108**, 15301-15310.
32. Pozo Ramajo, A., Petty, S. A., Starzyk, A., Decatur, S. M. & Volk, M. (2005). The α -helix folds more rapidly at the C-terminus than at the N-terminus. *J. Am. Chem. Soc.* **127**, 13784-13785.
33. Mikhonin, A. V., Asher, S. A., Bykov, S. V. & Murza, A. (2007). UV Raman spatially resolved melting dynamics of isotopically labeled polyalanyl peptide: slow α -helix melting follows 3_{10} -helices and π -bulges premelting. *J. Phys. Chem. B* **111**, 3280-3292.

34. Balakrishnan, G., Hu, Y., Bender, G. M., Getahun, Z., DeGrado, W. F. & Spiro, T. G. (2007). Enthalpic and entropic stages in α -helical peptide unfolding, from laser T-jump/UV Raman spectroscopy. *J. Am. Chem. Soc.* **129**, 12801-12808.
35. Thompson, P. A., Munoz, V., Jas, G. S., Henry, E. R., Eaton, W. A. & Hofrichter, J. (2000). The helix-coil kinetics of a heteropeptide. *J. Phys. Chem. B* **104**, 378-389.
36. Rohl, C. A. & Baldwin, R. L. (1994). Exchange kinetics of individual amide protons in ^{15}N -labeled helical peptides measured by isotope-edited NMR. *Biochemistry* **33**, 7760-7767.
37. Nesmelova, I., Krushelnitsky, A., Idiyatullin, D., Blanco, F., Ramirez-Alvarado, M., Daragan, V. A., Serrano, L. & Mayo, K. H. (2001). Conformational exchange on the microsecond time scale in α -helix and β -hairpin peptides measured by ^{13}C NMR transverse relaxation. *Biochemistry* **40**, 2844-2853.
38. Lapidus, L. J., Eaton, W. A. & Hofrichter, J. (2002). Measuring dynamic flexibility of the coil state of a helix-forming peptide. *J. Mol. Biol.* **319**, 19-25.
39. 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.
40. 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.
41. Satzger, H., Schmidt, B., Root, C., Zinth, W., Fierz, B., Krieger, F., Kiefhaber, T. & Gilch, P. (2004). Ultrafast quenching of the xanthone triplet by energy transfer: new insight into the intersystem crossing kinetics. *J. Phys. Chem. A* **108**, 10072-10079.
42. Fierz, B., Reiner, A. & Kiefhaber, T. (2007). Local conformational dynamics in α -helices measured by fast triplet transfer. *To be submitted*.
43. 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.
44. Krieger, F., Möglich, A. & Kiefhaber, T. (2005). Effect of proline and glycine residues on dynamics and barriers of loop formation in polypeptide chains. *J. Am. Chem. Soc.* **127**, 3346-3352.
45. Möglich, A., Krieger, F. & Kiefhaber, T. (2005). Molecular basis for the effect of urea and guanidinium chloride on the dynamics of unfolded polypeptide chains. *J. Mol. Biol.* **345**, 153-162.
46. Kramers, H. A. (1940). Brownian motion in a field of force and the diffusion model of chemical reactions. *Physica* **4**, 284-304.
47. Chemical Rubber Company. (1977). *CRC Handbook of Chemistry and Physics*. 58 ed., CRC Press.
48. Pohl, F. M. (1976). Temperature-dependence of the folding kinetics of folding of chymotrypsinogen A. *FEBS Lett.* **65**, 293-296.
49. Oliveberg, M., Tan, Y. J. & Fersht, A. R. (1995). Negative activation enthalpies in the kinetics of protein folding. *Proc. Natl. Acad. Sci. USA* **92**, 8926-8929.
50. Schönbrunner, N., Pappenberger, G., Scharf, M., Engels, J. & Kiefhaber, T. (1997). Effect of preformed correct tertiary interactions on rapid two-state tendamistat folding: evidence for hairpins as initiation sites for beta-sheet formation. *Biochemistry* **36**, 9057-9065.
51. Scholtz, J. M., Marqusee, S., Baldwin, R. L., York, E. J., Stewart, J. M., Santoro, M. & Bolen, D. W. (1991). Calorimetric determination of the enthalpy change

- for the α -helix to coil transition of an alanine peptide in water. *Proc. Natl. Acad. Sci. USA* **88**, 2854-2858.
52. Lopez, M. M., Chin, D. H., Baldwin, R. L. & Makhatadze, G. I. (2002). The enthalpy of the alanine peptide helix measured by isothermal titration calorimetry using metal-binding to induce helix formation. *Proc. Natl. Acad. Sci. USA* **99**, 1298-1302.
 53. Scholtz, J. M., Qian, H., York, E. J., Stewart, J. M. & Baldwin, R. L. (1991). Parameters of helix-coil transition theory for alanine-based peptides of varying chain lengths in water. *Biopolymers* **31**, 1463-1470.
 54. Richardson, J. M. & Makhatadze, G. I. (2004). Temperature dependence of the thermodynamics of helix-coil transition. *J. Mol. Biol.* **335**, 1029-1037.
 55. Leffler, J. E. (1953). Parameters for the description of transition states. *Science* **117**, 340-341.
 56. Ihalainen, J. A., Bredenbeck, J., Pfister, R., Helbing, J., Chi, L., van Stokkum, I. H., Woolley, G. A. & Hamm, P. (2007). Folding and unfolding of a photoswitchable peptide from picoseconds to microseconds. *Proc. Natl. Acad. Sci. USA* **104**, 5383-5388.
 57. Zwanzig, R. (1988). Diffusion in a rough potential. *Proc. Natl. Acad. Sci. USA* **85**, 2029-2030.
 58. Bryngelson, J. D., Onuchic, J. N., Socci, N. D. & Wolynes, P. G. (1995). Funnels, pathways, and the energy landscape of protein folding: a synthesis. *Proteins* **21**, 167-195.
 59. Graham, R. & Lewis, J. R. (1978). Synthesis of 9-oxoxanthen-2-carboxylic acids. *J. Chem. Soc. Perkin Trans. 1*, 876-881.
 60. Pace, C. N. (1986). Determination and analysis of urea and guanidine hydrochloride denaturation curves. *Methods Enzymol.* **131**, 266-280.
 61. Kiefhaber, T., Kohler, H. H. & Schmid, F. X. (1992). Kinetic coupling between protein folding and prolyl isomerization. I. Theoretical models. *J. Mol. Biol.* **224**, 217-229.
 62. Park, S. K. & Miller, K. W. (1988). Random Number Generators. *Comm. ACM* **31**, 1192-1201.

Figure Captions

Figure 1. (A) Far-UV CD spectrum at 5 °C. (B) Thermal melting of the α -helical structure monitored at 222 nm. The data were recorded in 10 mM potassium phosphate, pH 7.0.

Figure 2. TTET measurements of the dynamics in the helix centre at different temperatures. (A) Decays of the Xan triplet absorption band (590 nm) at temperatures from 0 °C to 30 °C (violet to red). The solid lines represent fits of double exponential functions. (B) Apparent rate constants λ_1 (black) and λ_2 (grey) obtained from the double exponential fits. (C) Reduced amplitudes $A_1/(A_1+A_2)$ (black) and $A_2/(A_1+A_2)$ (grey) corresponding to λ_1 and λ_2 . All measurements were performed in 10 mM potassium phosphate, pH 7.0.

Figure 3. Results from TTET measurements with increasing concentrations of urea at 0 °C (A, C) and 30 °C (B, D). (A, B) Apparent rate constants λ_1 (black) and λ_2 (grey) obtained from double exponential fits to the decays of the Xan triplet state absorption band. (C, D) Reduced amplitudes of λ_1 and λ_2 . All data at a single temperature were globally fitted to the analytic solution of the three-state system depicted in Eq. 1, assuming linear dependencies of $\ln k_i$ on the urea concentration (Eq. 2). The global fits are shown as black and grey lines, the resulting microscopic rate constants as coloured lines (k_f green, k_u red and k_c blue).

Figure 4. Results from the double perturbation experiment in the centre of the 21 residue α -helical peptide. (A) Arrhenius plot of the rate constants for helix folding k_f^0 (◆), helix unfolding k_u^0 (■) and contact formation k_c^0 (●) in 0 M urea. The data were obtained from global fits as shown in Figure 3. The solid lines represent fits of Eq. 4 with the parameters given in Table 1, dotted lines symbolize simple Arrhenius behavior. (B) Temperature dependence of the corresponding m_i -values (Eq. 2), coloured as in (A).

Figure 5. Results obtained from the simulations on the dynamics of the helix centre. Details on the simulation and the kinetic Ising model are given in the text. The reduced rate constants of helix folding (green) and unfolding (red) for different parameter sets

are shown. The results for a segment length of one (*) are compared to results for a segment length of five (○) obtained with $\Delta H_{\text{eq}}(25\text{ °C}) = 4.2\text{ kJ mol}^{-1}$, $\alpha_H = 0$ and no heat capacity changes. Furthermore the results for the same parameters, but $\alpha_H = 1$ (□) and $\alpha_H = -8$ (●) are shown. The latter corresponds to enthalpic barriers of $\Delta H^{0\ddagger} = 33.6\text{ kJ mol}^{-1}$ for k_1 and $\Delta H^{0\ddagger} = 37.8\text{ kJ mol}^{-1}$ for k_2 . It should be noted that all points are associated with a significant error, since the first passage time distributions are only approximated by a single rate constant. The solid lines are fits of Eq. 4.

Tables

Table 1. Two-State Evaluation of the Dynamics in the Helix Centre.

Parameter	H \rightarrow TS ‡ (k_u)	C \rightarrow TS ‡ (k_f)	H \rightarrow C (K_{eq})
ΔH^0 (kJ mol $^{-1}$)	43 ± 4.7	-3 ± 11	46 ± 16
ΔS^0 (kJ mol $^{-1}$ K $^{-1}$)	0.1 ± 0.02	-0.05 ± 0.06	0.15 ± 0.08
$-T \cdot \Delta S^0$ (kJ mol $^{-1}$)	-29 ± 6.0	15 ± 18	-45 ± 24
ΔC_p^0 (kJ mol $^{-1}$ K $^{-1}$)	-1.1 ± 0.4	-2.0 ± 0.8	0.9 ± 1.2
ΔG^0 (kJ mol $^{-1}$)	13 ± 11	12 ± 30	1.3 ± 41

Helix folding k_f and unfolding k_u are described by assuming a transition state TS between the ground states helix H and coil C. The activation enthalpies $\Delta H^{0\ddagger}_{u,f}$, entropies $\Delta S^{0\ddagger}_{u,f}$ and heat capacities $\Delta C_p^{0\ddagger}_{p,u,f}$ were determined by non-linear least-square fitting of Eq. 4 with $k_i^0 = 5 \cdot 10^7\text{ s}^{-1}$ to the data shown in Figure 4A. From these values the activation free energy $\Delta G^{0\ddagger}_{u,f}$ and the equilibrium parameters were calculated. All parameters are given with their standard deviation for a reference temperature of 25 °C, pH 7.0 and 0 M urea.

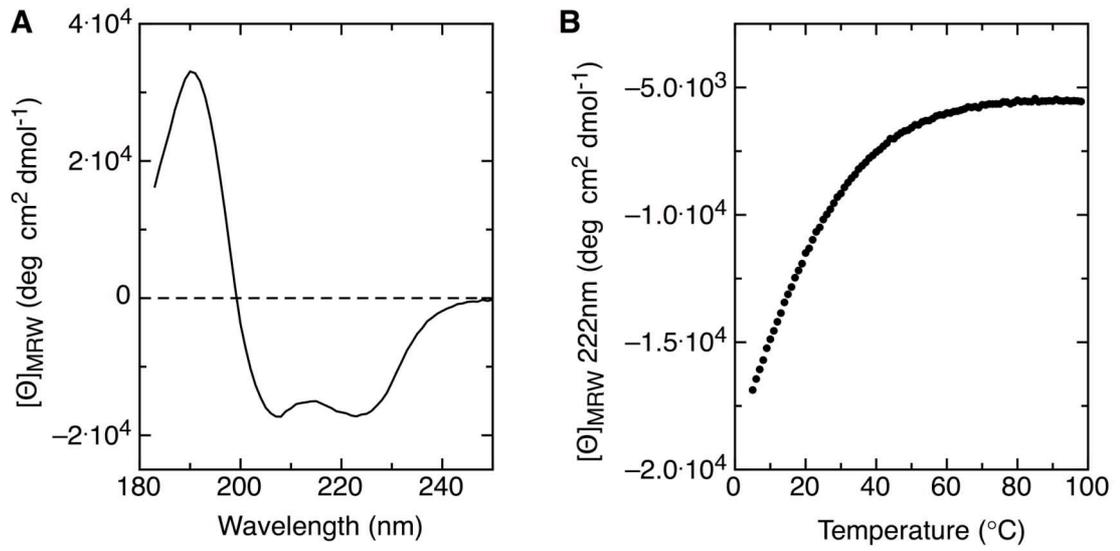


Figure 1

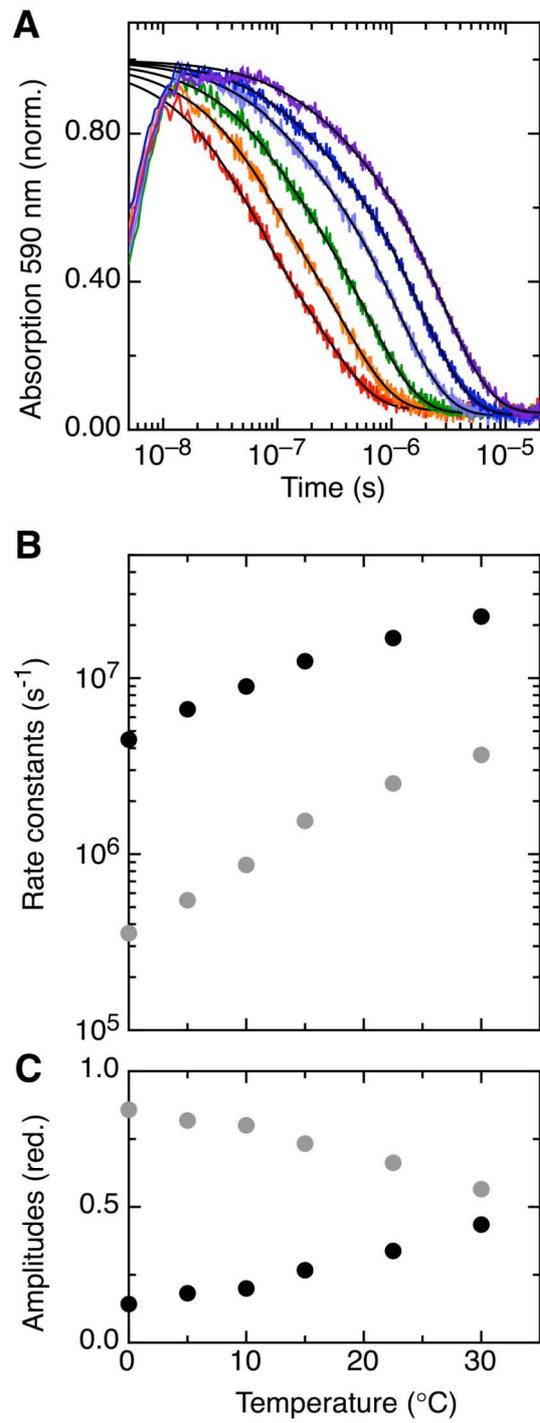


Figure 2

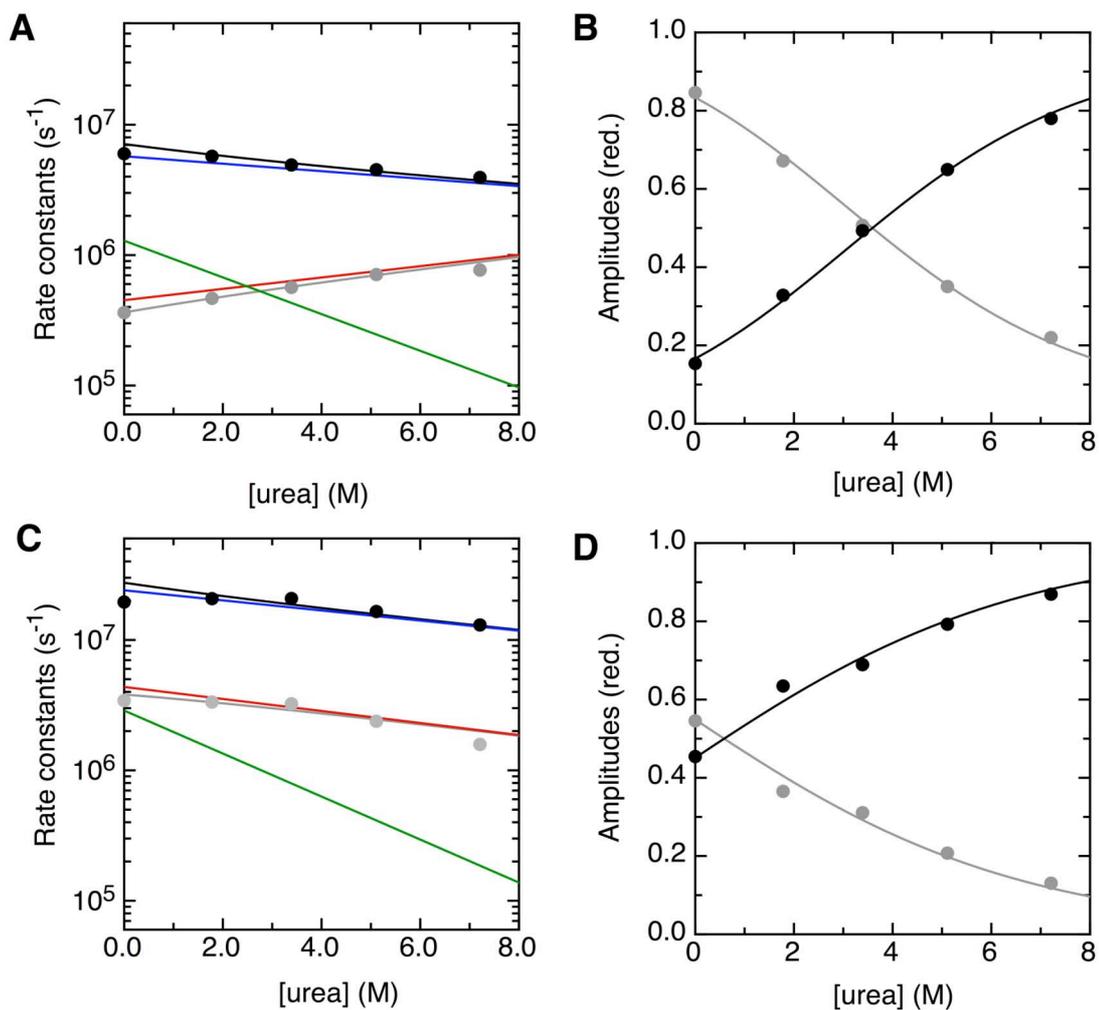


Figure 3

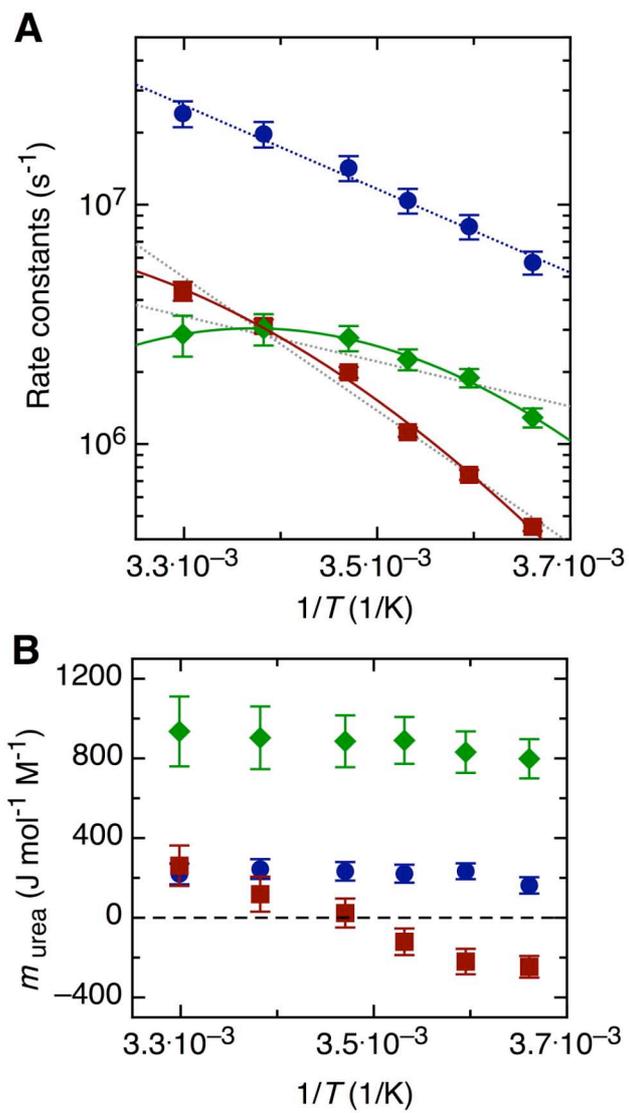


Figure 4

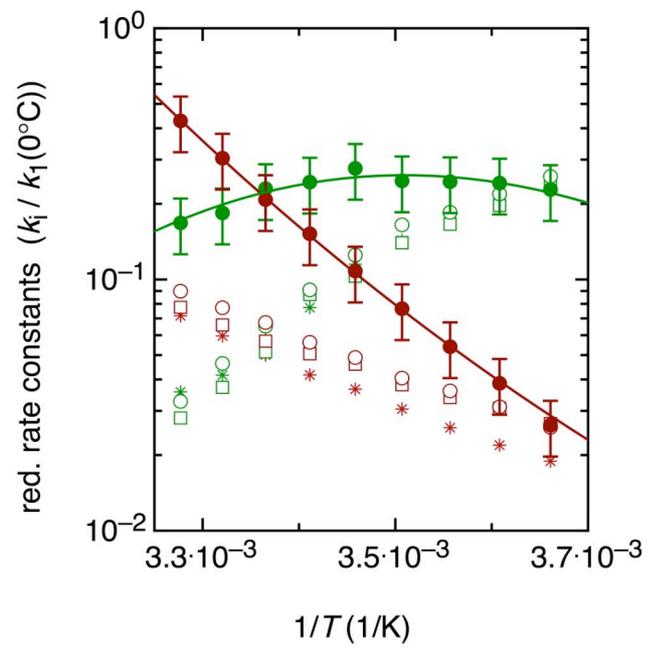


Figure 5

Effect of Thiopeptide Bonds on α -Helix Structure and Stability

Andreas Reiner,[†] Dirk Wildemann,[‡] Gunter Fischer,[‡] and Thomas Kiefhaber^{†,*}

Chemistry Department, Institute for Biophysical Chemistry, Technische Universität München and Munich Center for Integrated Protein Science, Lichtenbergstrasse 4, D-85747 Garching, Germany and the Max Planck Research Unit for Enzymology of Protein Folding, Weinbergweg 22, D-06120 Halle (Saale), Germany

Received February 28, 2008; E-mail: t.kiefhaber@tum.de

Abstract: Thioamide (thioamide) bonds are nearly isosteric substitutions for amides but have altered hydrogen-bonding and photophysical properties. They are thus well-suited backbone modifications for physicochemical studies on peptides and proteins. The effect of thioamides on protein structure and stability has not been subject to detailed experimental investigations up to date. We used alanine-based model peptides to test the influence of single thioamide bonds on α -helix structure and stability. The results from circular dichroism measurements show that thioamides are strongly helix-destabilizing. The effect of an oxo-to-thioamide backbone substitution is of similar magnitude as an alanine-to-glycine substitution resulting in a helix destabilization of about 7 kJ/mol. NMR characterization of a helical peptide with a thiopeptide bond near the N-terminus indicates that the thiopeptide moiety is tolerated in helical structures. The thioamide group is engaged in an $i, i+4$ hydrogen bond, arguing against the formation of a 3_10 -helical structure as suggested for the N-termini of α -helices in general and for thiopeptides in particular.

Introduction

Proteins are built from a small set of amino acids linked by peptide bonds. They can form systems of high complexity and versatility despite the simplicity of their building blocks. The introduction of non-natural side chains or backbone linkages allows the study, alteration, or expansion of the basic properties of polypeptide chains. The polypeptide backbone plays an important role in protein structure, stability, and dynamics, since it mediates solvation,^{1,2} restricts conformational space,³ and forms hydrogen bonds. Drastic changes in the backbone properties as introduced by glycine and proline residues strongly alter the structural and dynamic properties of polypeptide chains.^{4,5} Numerous synthetic backbone modifications have been introduced, encompassing β -peptides, peptides with reversed stereochemistry or direction, and several kinds of amide substitutions, e.g., exchanges to tertiary amides, esters, aldehydes, olefins, or thioesters. The replacement of the amide bond by a thioamide (thioamide) bond ($\Psi[\text{CS}-\text{NH}]$) represents a conservative modification of the peptide backbone (Scheme 1), which has been envisaged to improve the bioactivity and stability

Scheme 1. Oxo-to-Thioamide Substitution in a Peptide Bond



of peptidomimetics (for examples, see refs 6 and 9). Thioamide groups are nearly but not completely isosteric to amides. The C=S bond is 37% longer than the C=O bond (1.65 Å vs 1.20 Å), and sulfur has a 32% larger van der Waals radius than oxygen (1.85 Å vs 1.40 Å). Theoretical studies indicate that the conformational space of the residues preceding and following a thioamide is more restricted compared to an amide bond.^{10–13} Although the energetic estimates differ, it was predicted that the typical ϕ, ψ -regions for the peptide backbone should remain accessible, including those corresponding to β -sheets, β -turns, and right-handed α -helices. Up to date a single

[†] Technische Universität München.

[‡] Max Planck Research Unit for Enzymology of Protein Folding.

- (1) Baldwin, R. L. *Biophys. Chem.* **2002**, *101* (102), 203–210.
- (2) Baldwin, R. L. Weak interactions in protein folding: hydrophobic free energy, van der Waals interactions, peptide hydrogen bonds and peptide solvation. In *Protein Folding Handbook*; Buchner, J., Kiefhaber, T., Eds.; Wiley/VCH: Weinheim, 2005; Vol. 1, pp 127–162.
- (3) Ramachandran, G. N.; Sasisekharan, V. *Adv. Protein Chem.* **1968**, *23*, 283–437.
- (4) Flory, P. J. *Statistical Mechanics of Chain Molecules*; Hanser Publishers: Munich, 1969.
- (5) Krieger, F.; Möglich, A.; Kiefhaber, T. *J. Am. Chem. Soc.* **2005**, *127*, 3346–3352.

- (6) Seebach, D.; Ko, S. Y.; Kessler, H.; Köck, M.; Reggelin, M.; Schmieder, P.; Walkinshaw, M. D.; Bölsterli, J. J.; Bevec, D. *Helv. Chim. Acta* **1991**, *74*, 1953–1990.
- (7) Kessler, H.; Matter, H.; Geyer, A.; Diehl, H.-J.; Köck, M.; Kurz, G.; Oppendoerf, F. R.; Callens, M.; Wierenga, R. K. *Angew. Chem., Int. Ed. Engl.* **1992**, *31*, 328–330.
- (8) Yao, S.; Zutshi, R.; Chmielewski, J. *Bioorg. Med. Chem. Lett.* **1998**, *8*, 699–704.
- (9) Frank, R.; Jakob, M.; Thuncke, F.; Fischer, G.; Schutkowski, M. *Angew. Chem., Int. Ed.* **2000**, *39*, 1120–1122.
- (10) Balaji, V. N.; Profeta, S.; Dietrich, S. W. *Biochem. Biophys. Res. Commun.* **1987**, *145*, 834–841.
- (11) Cour, T. F. M. *Int. J. Peptide Protein Res.* **1987**, *30*, 564–571.
- (12) Artis, D. R.; Lipton, M. A. *J. Am. Chem. Soc.* **1998**, *120*, 12200–12206.
- (13) Tran, T. T.; Treutlein, H.; Burgess, A. W. *J. Comput. Chem.* **2001**, *22*, 1026–1037.

example of a naturally occurring protein was reported, which contains a backbone thiopeptide bond.¹⁴ Thioamides can serve as both hydrogen bond donors and acceptors. In mixed hydrogen bonds between thioamides and amides, the thioamide is predicted to be a weaker hydrogen bond acceptor and a better hydrogen bond donor compared to an amide.¹⁵ Therefore thiopeptide groups may be suitable to evaluate the contribution of single hydrogen bonds to the stability, folding, and function of proteins. A special feature of the thiopeptide bond arises from the red-shifted absorption bands and the higher rotational barrier for the *cis*–*trans* isomerization.^{16,17} This property allows a selective shift of the equilibrium either to the *cis* or to the *trans* isomerization state by irradiation with UV light of different wavelengths. The photoinduced isomerization is efficient and fast (<600 ps),^{18–20} whereas thermal relaxation back to equilibrium is slow (>10 min).^{9,16} Thiopeptides are therefore good candidates for fast photoswitches in proteins, either to regulate biological or enzymatic activity²¹ or to initiate conformational transitions for time-resolved studies.

In order to use the potential of thioamide substitutions to study protein structure, folding, and dynamics, it is important to evaluate their effects on protein stability. The introduction of thioamides at hydrogen-bonded positions in β -sheets and α -helices should be highly unfavorable, since these structures are optimally packed to accommodate the shorter amide–amide hydrogen bonds (2.1 Å for an amide–amide hydrogen bond vs 2.7 Å for a thioamide–amide hydrogen bond).¹¹ In an experimental study on a thioamide substituted β -hairpin peptide, steric clashes were avoided by placing the sulfur in a non-hydrogen-bonded turn position.²² The results suggested hairpin formation in the thioylated peptide but did not address the effect on hairpin stability. In α -helices all carbonyl groups except for those of the four C-terminal residues are involved in *i*, *i*+4 hydrogen bonds. Thioamides are unlikely to fit into the tight groove between two helical turns, which might cause helix disruption or kink formation.¹¹ It was further proposed that thiopeptide groups favor the 3_{10} -helix conformation, since it is better suited to accommodate the larger sulfur.²³ Surprisingly, an experimental study on a dimeric α -helical coiled-coil structure (GCN4) reported that thiopeptide groups are structurally tolerated and even a slight stabilizing effect on the coiled-coil was observed.²⁴ The contradiction to theoretical considerations was attributed to favorable contributions from hydrogen bonding of the thioamide group. However, the two positions

Table 1. Amino Acid Sequences of the Peptides Used in This Study

	central substitution
cΨ ^a	Ac-KAAAA KAA Ψ AAA KAAAA K-NH ₂
cA	Ac-KAAAA KAA AA KAAAA KGY-NH ₂
cG	Ac-KAAAA KAA GA KAAAA KGY-NH ₂
cP	Ac-KAAAA KAA PA KAAAA KGY-NH ₂
	N-terminal substitution
nΨ	Ac-DF Ψ AAA KAAAA KAAAA K-NH ₂
nA	Ac-DF AAA KAAAA KAAAA K-NH ₂
nG	Ac-DF GAA KAAAA KAAAA K-NH ₂
nP	Ac-DF PAA KAAAA KAAAA K-NH ₂

^a Ψ denotes the substitution of the amide bond [CO–NH] by a thioamide bond [CS–NH].

were chosen²⁴ to place the larger sulfur atom in a widened part of the slightly distorted helix structure present in coiled-coils.²⁵

To investigate the effect of thioamides on α -helix structure and stability we introduced single amide-to-thioamide substitutions at a central and a N-terminal position in alanine-based helical peptides. Using circular dichroism (CD) and nuclear magnetic resonance (NMR) spectroscopy we tested the effect of thiopeptide bonds on α -helix structure and stability.

Results and Discussions

Thioamide Substitution at a Central Helix Position. We chose a well-characterized 16 amino acid alanine-based model helix^{26–31} to investigate the effect of single thiopeptide substitutions on α -helix structure and stability (Table 1). The peptides contain lysine residues at every fifth position to increase solubility and were shown to be monomeric.²⁶ Capping of the termini prevents unfavorable electrostatic interactions with the helix dipole. This peptide was shown to display a helix content of ~70%.^{27,30} We synthesized a peptide with the central amide bond between Ala8 and Ala9 substituted by a thioamide bond (Ψ [CS–NH]), indicated by the short notation **c Ψ** (Table 1). The helix content is expected to reach a maximum in the central region of the peptide,³² rendering this part especially sensitive to perturbations.²⁷ To characterize the effect of the thiopeptide substitution we synthesized three reference peptides. One peptide has the same amino acid sequence as **c Ψ** but contains the regular oxopeptide bond between Ala8 and Ala9 (**cA**). The other two peptides also have an oxopeptide bond between residues 8 and 9 but have Ala9 replaced by the helix-destabilizing amino acids glycine (**cG**) or proline (**cP**).

To test the effect of the thiopeptide bond on helix structure and stability we recorded far-UV CD spectra of all peptides. Figure 1 compares the spectrum of **c Ψ** to the spectra of the reference peptides. The **c Ψ** peptide is significantly less structured than the **cA** peptide, which shows a characteristic α -helical

(14) Grabarse, W.; Mahlert, F.; Shima, S.; Thauer, R. K.; Ermler, U. *J. Mol. Biol.* **2000**, *303*, 329–344.

(15) Aleman, C. *J. Phys. Chem. A* **2001**, *105*, 6717–6723.

(16) Zhao, J.; Wildemann, D.; Jakob, M.; Vargas, C.; Schiene-Fischer, C. *Chem. Commun.* **2003**, 2810–2811.

(17) Wiberg, K. B.; Rush, D. J. *J. Am. Chem. Soc.* **2001**, *123*, 2038–2046.

(18) Satzger, H.; Root, C.; Gilch, P.; Zinth, W.; Wildemann, D.; Fischer, G. *J. Phys. Chem. B* **2005**, *109*, 4770–4775.

(19) Cervetto, V.; Bregy, H.; Hamm, P.; Helbing, J. *J. Phys. Chem. A* **2006**, *110*, 11473–11478.

(20) Helbing, J.; Bregy, H.; Bredenbeck, J.; Pfister, R.; Hamm, P.; Huber, R.; Wachtveitl, J.; DeVico, L.; Olivucci, M. *J. Am. Chem. Soc.* **2004**, *126*, 8823–8834.

(21) Wildemann, D.; Schiene-Fischer, C.; Aumüller, T.; Bachmann, A.; Kiefhaber, T.; Lücke, C.; Fischer, G. *J. Am. Chem. Soc.* **2007**, *129*, 4910–4918.

(22) Miwa, J. H.; Patel, A. K.; Vivatrat, N.; Popek, S. M.; Meyer, A. M. *Org. Lett.* **2001**, *3*, 3373–3375.

(23) Tran, T. T.; Zeng, J.; Treutlein, H.; Burgess, A. W. *J. Am. Chem. Soc.* **2002**, *124*, 5222–5230.

(24) Miwa, J. H.; Pallivathucal, L.; Gowda, S.; Lee, K. E. *Org. Lett.* **2002**, *4*, 4655–4657.

(25) Goodman, E. M.; Kim, P. S. *Biochemistry* **1991**, *30*, 11615–11620.

(26) Marqusee, S.; Robbins, V. H.; Baldwin, R. L. *Proc. Natl. Acad. Sci. U.S.A.* **1989**, *86*, 5286–5290.

(27) Chakrabarty, A.; Schellman, J. A.; Baldwin, R. L. *Nature* **1991**, *351*, 586–588.

(28) Scholtz, J. M.; Qian, H.; York, E. J.; Stewart, J. M.; Baldwin, R. L. *Biopolymers* **1991**, *31*, 1463–1470.

(29) Scholtz, J. M.; Baldwin, R. L. *Annu. Rev. Biophys. Biomol. Struct.* **1992**, *21*, 95–118.

(30) Chakrabarty, A.; Kortemme, T.; Baldwin, R. L. *Protein Sci.* **1994**, *3*, 843–852.

(31) Rohl, C. A.; Chakrabarty, A.; Baldwin, R. L. *Protein Sci.* **1996**, *5*, 2623–2637.

(32) Lifson, S.; Roig, A. *J. Chem. Phys.* **1961**, *34*, 1963–1974.

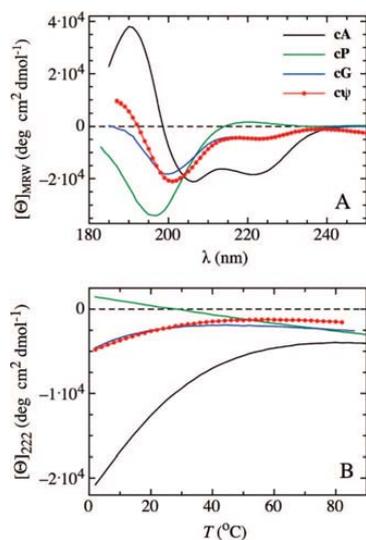


Figure 1. (A) Far-UV CD spectra of the peptides with central substitutions at 5 °C. (B) Thermal melting of the peptides monitored at 222 nm. The peptide sequences are given in Table 1. The data were recorded in 10 mM potassium phosphate, pH 7.0.

spectrum with minima at 222 and 208 nm and a maximum at 190 nm. However, **cPsi** retains some helical structure as indicated by minima around 222 and 201 nm. In the far-UV region the CD spectrum of **cPsi** is comparable to the spectrum of the peptide with the glycine substitution (**cG**), whereas the reference peptide with a proline residue (**cP**) shows a typical random coil spectrum with a maximum around 218 nm and a minimum around 197 nm. The small differences between the spectra of **cPsi** and **cG** below 215 nm are probably due to contributions from the thioamide bond, which has an additional absorbance band centered around 210 nm.³³ The near-UV CD spectrum of **cPsi** differs from the spectra of all reference peptides by an additional negative band around 266 nm and a very weak positive band around 350 nm (Figure S1). These bands originate from the red-shifted $\pi_0\pi^*$ transition and the $n\pi^*$ transition, respectively, of the thioamide group.³³

In order to quantify the destabilizing effect of the thiopeptide bond, we calculated the helicity of the peptides based on the mean residual ellipticity at 222 nm ($[\Theta]_{222}$) in the presence of 1 M NaCl (Table 2). The presence of NaCl was shown to allow a more reliable determination of the helical content in alanine-based model peptides and is commonly used for a quantitative analysis.^{30,31} For the calculation of the helical content we used values of $[\Theta]_{\text{H}} = -44\,000 \cdot (1-3/N)$ deg cm² dmol⁻¹, where N is the number of residues and $[\Theta]_{\text{C}} = 2220$ deg cm² dmol⁻¹ for $[\Theta]_{222}$ of the 100% helical and 100% unfolded state, respectively.³⁴ At 222 nm the influence of the thiopeptide group on the CD spectrum of the unfolded state is small, since thiopeptide groups in unstructured alanine peptides give only a weak CD signal, which is, however, slightly larger than the CD signal of an oxopeptide bond.³³ We do not know the CD signal of a thioamide bond in the helical conformation, but the contribution of a single thioamide bond

Table 2. Helix Content Obtained from CD Spectroscopy and Helix Propagation Parameters Obtained from Lifson–Roig Theory

peptide	$[\Theta]_{222}^a$ (deg cm ² dmol ⁻¹)	f_{H}^b (%)	w^c	$\Delta\Delta G^{\text{od}}$ (kJ mol ⁻¹)
cPsi	-6900	24 ± 8	0.044	7.2
cA	-24 200	68 ± 5	1.70 ^e	0
cG	-7900	26 ± 5	0.050	7.0
cP	-290	6 ± 5	<0.001	>15.8
nPsi	-16 000	48 ± 8	n.d. ^f	n.d.
nA	-23 900	69 ± 5	n.d.	n.d.
nG	-17 000	51 ± 5	n.d.	n.d.
nP	-14 200	43 ± 5	n.d.	n.d.

^a $[\Theta]_{\text{MRW}}$ at 222 nm measured in 10 mM potassium phosphate, pH 7.0 at 0 °C. For the centrally substituted peptides 1 M NaCl was added.
^b Mean helix content. The error is estimated to be <5%, for **cPsi** < ±8%.
^c Propagation parameter for the single substitutions, obtained using Lifson–Roig theory with the parameters given in ref 31.
^d Free enthalpy difference for helix formation compared to alanine at 0 °C. The values were calculated using $\Delta G^{\circ} = -RT \ln(w/(1+v))$.
^e Not fitted, since it is part of the parameter set used for analysis.
^f Using this value Lifson–Roig theory predicts a helix content of 69% for **cA**.
^g The N-terminal position is not very sensitive to perturbations. Thus, no attempt was made to calculate Lifson–Roig parameters from the CD signals of the N-terminally substituted peptides.

to the CD signal in the 16 amino acid peptide is expected to be small. Based on these considerations, we assume an upper limit for the error of the calculated helix content in our single thioamide bond containing peptides of ~8%. For **cPsi** this results in a helix content of ~24 ± 8% compared to helix contents of 68 ± 5%, 26 ± 5%, and 6 ± 5% (Table 2) for **cA**, **cG**, and **cP**, respectively. The values for the peptides with all-oxoamide bonds are in good agreement with reported values for similar peptides.³⁰

The helical content of the different peptides can further be obtained from the thermal melting behavior monitored by the change in CD at 222 nm (Figure 1B). The unfolding transitions of the **cPsi** and **cG** peptides show similar cooperativity and a similar change in ellipticity. At low temperature the $[\Theta]_{222}$ -values of the two peptides and the change in $[\Theta]_{222}$ with temperature are almost identical. These findings provide further evidence for similar stability and helicity of the **cPsi** and **cG** peptides. The $[\Theta]_{222}$ -value of the unfolded baseline is increased by about 1000 deg cm² dmol⁻¹ in **cPsi** as expected from the contributions of the thioamide to the CD signal of the unfolded state, but shows the same linear decrease with increasing temperature as in the **cG** peptide. The **cA** peptide shows a much larger signal change upon thermal unfolding due to its higher helix content. The **cP** peptide exhibits a linear decrease in CD signal with increasing temperature, in accordance with the presence of a predominately unfolded conformation at all temperatures.³⁴ Temperature-induced structural changes in the **cPsi** peptide are also observed when the additional CD bands of the thioamide bond at 266 and 350 nm are monitored (Figure S1). However, the CD changes upon helix unfolding in this region are much smaller than those at 222 nm.

A quantitative evaluation of the effect of side chain or backbone modifications on helix stability can be obtained using Lifson–Roig theory for helix–coil transition.³² In Lifson–Roig theory the propagation parameter, w^{Xaa} , is a statistical weight that describes the tendency of a single amino acid residue to adopt the helical conformation. It can be interpreted as an equilibrium constant between a residue in the coiled and in the helical state. A w^{Ala} -value of 1.70 has been determined for alanine, the amino acid with the highest tendency to adopt the helical conformation.³¹ To evaluate the effect of a single

(33) Hollósi, M.; Kollát, E.; Kajtár, J.; Kajtár, M.; Fasman, G. D. *Biopolymers* **1990**, *30*, 1061–1072.

(34) Luo, P.; Baldwin, R. L. *Biochemistry* **1997**, *36*, 8413–8421.

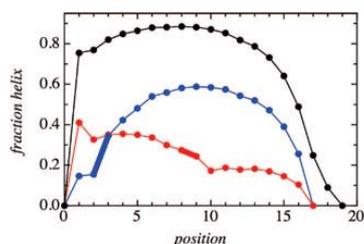


Figure 2. Position-dependent helix content of the peptides **cA** (black), **cΨ** (red), and **nΨ** (blue) as predicted by Lifson–Roig theory with the parameters given in Table 2. The bold line indicates the location of the thiooxopeptide bond.

thiooxamide bond on helix stability we ascribe a special propagation parameter, w^{thio} , to the residue C-terminal of the thiooxamide bond (Table 2). To determine the w^{thio} -value, we used the w^{aa} -values for the standard amino acids and a weight for the boundary positions, v , of 0.036. Additional weights were introduced to account for capping effects at the helix ends.^{31,35} Analysis of the **cΨ** peptide shows that the highly destabilizing effect of the thiooxopeptide group is reflected in a low propagation parameter, w^{thio} , of 0.044, which corresponds to a helix destabilizing effect of ~ 7 kJ/mol relative to an oxopeptide bond in the **cA** reference peptide (Table 2). This leads to a major decrease of helical content in the central part of the α -helix (Figure 2). The w^{thio} -value is almost identical to the value of 0.050 determined for w^{Gly} in the **cG** peptide (Table 2). It should be noted that a thiooxopeptide bond in a central helix position involves hydrogen bonds to amides four residues apart in both directions and therefore might have a less local effect than the glycine and proline substitutions in the reference peptides.

Thiooxamide Substitution in the N-Terminal Region of an α -Helix. The results on **cΨ** have shown that a thiooxopeptide group in the central region of an α -helix has a strongly helix-destabilizing effect (Figure 1). The terminal regions of helical peptides exhibit less helical content compared to the central parts, and thus perturbations in this region have less effect on helix content.²⁷ To study the effect of a thiooxopeptide bond on helix structure and stability in a less destabilized environment, we placed an oxo-to-thioxo substitution between residues 2 and 3 in an alanine-based peptide (**nΨ**; Table 1). In addition, Lys1 was replaced by a stabilizing aspartate, adding a favorable interaction with the helix dipole,³⁶ and a phenylalanine was introduced at position 2 to facilitate structural analysis of this region of the helix using NMR spectroscopy.³⁷ The corresponding reference peptides **nA**, **nG**, and **nP** have an alanine, glycine, or proline at position 3, respectively, and regular oxo-peptide bonds between Phe2 and Ala3 (Table 1).

The CD spectrum of **nΨ** indicates a significant fraction of α -helix (Figure 3A). Comparison with the spectrum of the unsubstituted reference peptide **nA** reveals again a destabilizing effect of the thiooxamide bond. However, the decrease in helical content caused by thioxylation is less pronounced compared to the **cΨ** peptide (cf. Figure 1). As for the central substitutions, the CD spectra of **nΨ** and **nG** peptide are similar in the region

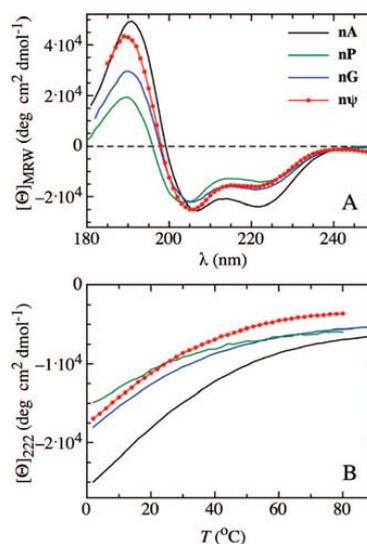


Figure 3. (A) Far-UV CD spectra of the peptides with N-terminal substitutions at 5 °C. (B) Thermal melting of the peptides monitored at 222 nm. Measurements were performed in 10 mM potassium phosphate, pH 7.0.

between 215 and 240 nm, whereas **nP** exhibits significantly less helicity. Also the thermal melting behaviors of **nΨ** and **nG** are similar (Figure 3B) confirming comparable helix contents of these peptides. Based on the ellipticity at 222 nm the helix content of **nΨ** is $\sim 45\%$, compared to 69% for **nA** (Table 1). Using a w^{thio} -value of 0.044 determined in **cΨ** (Table 1), Lifson–Roig theory predicts 46% helix content for **nΨ**. This agreement indicates that the destabilizing effect of the thiooxamide bond is similar at different positions within the helix. An earlier experimental study on a thiooxamide substituted dimeric GCN4 fragment placed the thiooxopeptide group in a widened position in the N-terminal region of a helical coiled-coil.²⁴ In this case no destabilization was observed, which led us to the conclusion that α -helices can tolerate thiooxamide substitutions without major changes in stability. The discrepancy to our results may arise from the special features of oligomeric coiled-coil structures, which have twisted helices with altered hydrogen-bonding geometry.²⁵

NMR Characterization of **nΨ.** Lifson–Roig theory predicts that the N-terminal three amino acids of **nΨ** have a very low helix propensity ($\leq 16\%$), whereas the alanine residue following the thiooxopeptide linkage (Ala4) populates the helical conformation to $\sim 35\%$ (Figure 2). A plateau of 60% helicity is reached at the residues following the thioxylated peptide bond. The substantial amount of helix formation in the N-terminal part of **nΨ** allows structural investigations around the thiooxopeptide bond by recording one- and two-dimensional ^1H NMR spectra (Figure 4). The amide protons of **nΨ** give sharp signals in the region from 7.3 to 8.4 ppm, whereas the thiooxamide proton appears as a broadened peak at 9.83 ppm. A downfield shift of this magnitude is generally observed for thiooxamide protons.^{6,38} Many of the other resonances overlap, as expected for such a uniform sequence, which prevents a complete assignment based

(35) Doig, A. J.; Chakrabarty, A.; Klingler, T. M.; Baldwin, R. L. *Biochemistry* **1994**, *33*, 3396–3403.

(36) Shoemaker, K. R.; Kim, P. S.; York, E. J.; Stewart, J. M.; Baldwin, R. L. *Nature* **1987**, *326*, 563–567.

(37) Scherer, G.; Kramer, M. L.; Schutkowski, M.; Reimer, U.; Fischer, G. *J. Am. Chem. Soc.* **1998**, *120*, 5568–5574.

(38) Jensen, O. E.; Lawesson, S.-O.; Baadi, R.; Piazzesi, A. N.; Toniolo, C. *Tetrahedron* **1985**, *41*, 5595–5606.

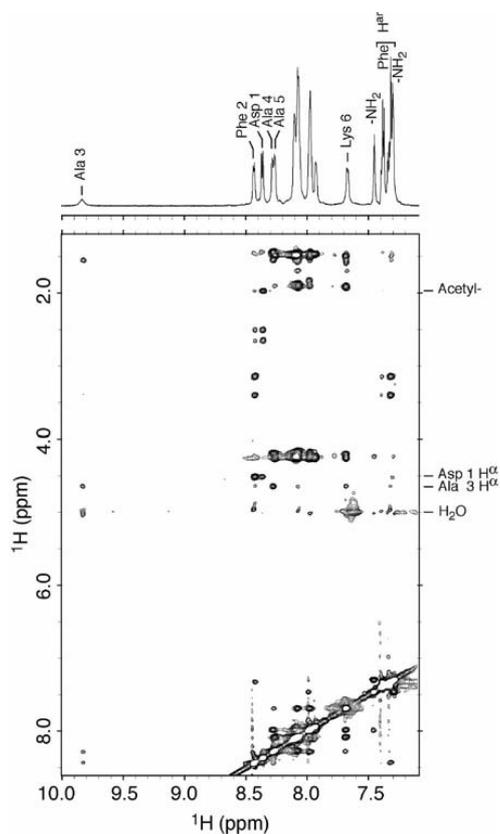


Figure 4. NMR characterization of $n\Psi$. The upper panel shows the amide region of the 1D spectrum at pH 7.0. The lower panel displays the corresponding region of the NOESY spectrum recorded with a mixing time of 460 ms at pH 5.0. Some resonances are labelled; the other assignments are listed in Table S1. Measurements were performed in 10 mM potassium phosphate at 2 °C.

on proton spectra, which was achieved for a similar peptide without a destabilizing substitution.³⁹ However, in the N-terminal region a number of resonances can be assigned using TOCSY and NOESY experiments (Figure 4), since the sequence is less uniform and some resonances are additionally influenced by the sulfur atom and by ring current effects of Phe2 (see Table S2). A key observation in the NMR spectrum is the upfield shift of the Lys6 amide proton ($\delta = 7.68$ ppm). This is due to the influence of the thioamide sulfur atom, since a similar Lys NH upfield shift is absent in the spectra of the all-oxoamide bond containing reference peptides (data not shown). This points to the close vicinity of the sulfur atom and the Lys6 amide proton and makes Lys6 a likely hydrogen bonding partner of the thiocarbonyl group of Phe2. This is supported by the positive temperature coefficient of the Lys6 amide proton resonance ($\Delta\delta/\Delta T = 2.4$ ppb/K) indicating that this proton is involved in a hydrogen bond (Figure S2 and Table S1). In contrast, the amide protons of Asp1, Phe2, Ala3, and Ala4/Ala5 have negative

(39) Millhauser, G. L.; Stenland, C. J.; Hanson, P.; Bolin, K. A.; van de Ven, F. J. *J. Mol. Biol.* **1997**, *267*, 963–974.

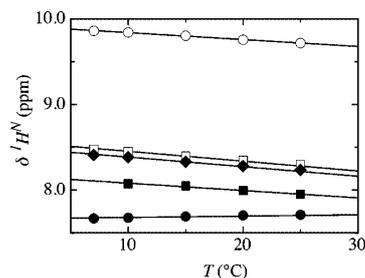


Figure 5. Temperature-dependence of the chemical shifts of the amide and thioxoamide protons in the N-terminal region of $n\Psi$. The amide protons of Asp1 (◆), Phe2 (□), Ala4/Ala5 (■), and Lys6 (●) are displayed in addition to the thioxoamide proton of Ala3 (○). Experiments were carried out in 10 mM potassium phosphate, pH 7.0.

temperature coefficients with $\Delta\delta/\Delta T \leq -8$ ppb/K (Figure 5 and Table S1), which is typical for non-hydrogen-bonded amide protons.^{40,41}

NOE cross-peaks from the thioxoamide proton become detectable upon lowering the pH from 7.0 to 5.0, which slows down thioxoamide proton exchange with water. The chemical shifts of all resonances (Figure 4) as well as the CD spectrum (data not shown) remain unchanged in the range from pH 5 to pH 7. The pattern of NOE cross-peaks in the NOESY spectrum reveals interactions between residues in the region from Phe2 to Lys6 pointing toward helical structure in this region.⁴² The well-resolved C^α -proton of Ala3 gives particularly useful structural information on the N-terminal region, since it shows cross-peaks to the amide protons of residues Ala3 (*i*), Ala4 (*i*+1), Lys6 (*i*+3), and likely also to the unresolved Ala7 (*i*, *i*+4) (Figure 4). NOEs from Ala3 to Ala5 (*i*+2) are absent, as expected for an α -helical structure, since these residues are on opposite sides of the helix. These observations indicate the formation of an α -helical structure in the region between Phe2 and Lys6. No NOE information can be obtained for the C-terminal region of Lys6 due to spectral overlap of the resonances (Figure 4).

The observed α -helical NOE pattern in the N-terminal region is in agreement with the strong upfield shift of the Lys6 amide proton, caused by the vicinity of the Phe2 thioxoamide group, and with the temperature-dependence of the Lys6 amide proton chemical shift (Figure 5). These findings indicate partial α -helix formation in the N-terminal region of $n\Psi$ including an *i*, *i*+4 hydrogen bond between the thiocarbonyl group of Phe2 and the amide proton of Lys6.

No evidence for a 3_{10} -helical structure was found for $n\Psi$, which was proposed to be generally present at the N-termini of helical peptides^{39,43} and was also predicted to be well suited to incorporate the larger sulfur of thioxoamides.²³ Formation of a 3_{10} -helix would result in an *i*, *i*+3 hydrogen bond, i.e., from the thiocarbonyl to the amide proton of Phe2 to Ala5, which can be excluded from the NMR spectra, since the NH proton of Ala5 experiences a similar chemical shift (Figure 4) and a similar temperature-dependence (Figure 5) as the neighboring Ala4.

(40) Merutka, G.; Dyson, H. J.; Wright, P. E. *J. Biomol. NMR* **1995**, *5*, 14–24.

(41) Cierpicki, T.; Otlewski, J. *J. Biomol. NMR* **2001**, *21*, 249–261.

(42) Wüthrich, K.; Billeter, M.; Braun, W. *J. Mol. Biol.* **1984**, *180*, 715–740.

(43) Bolin, K. A.; Millhauser, G. L. *Acc. Chem. Res.* **1999**, *32*, 1027–1033.

The finding that thioamide bonds can be incorporated into α -helical structures confirms the results from measurements on a thioamide substituted S-peptide, which is mainly unstructured in solution but forms an α -helical structure in complex with S-protein (RNase S).²¹ A thiopeptide bond introduced in the N-terminal part of the S-helix had no effect on the catalytic activity of the RNaseS complex. This was interpreted as evidence for helix formation around the thiopeptide bond.

Conclusions

Introducing thiopeptide groups at two different positions in alanine-based α -helical peptides reveals a highly helix-destabilizing effect of ~ 7 kJ/mol, independent of the position within the helix. This can be explained by the longer C=S bond and the larger sulfur atom, which are difficult to integrate into the densely packed α -helix.²³ Additional contributions are likely to arise from the altered hydrogen bonding properties of thiopeptide groups. The destabilizing effect of a thiopeptide bond may be different in native protein structures due to tertiary interactions, as observed for alanine-to-glycine substitutions in protein helices compared to peptide models.⁴⁴ The NMR data obtained on the peptide with a thioamide bond in the N-terminal region of the helix point to the formation of an $i, i+4$ hydrogen bond involving the thiocarbonyl acceptor group, whereas no evidence for an $i, i+3$ hydrogen bond was observed, which would indicate a 3_{10} -helical structure. This contrasts with the idea that a 3_{10} -helical conformation prevails at the N-terminus of helical peptides,^{39,43} which should be even more favored since it is better-suited to accommodate the thiopeptide bond.²³

Our results indicate that it is possible to incorporate thiopeptide groups into α -helical structures despite their strong helix-destabilizing effect. This shows that α -helical regions of proteins can be targets for thioamide substitutions in order to test the role of individual backbone hydrogen bonds for protein dynamics and stability. The observed destabilizing effect of an oxo- to thioamide bond in the range of 7 kJ/mol is large enough for a reliable determination of protein folding φ -values to test the formation of individual hydrogen bonds in the transition state of a folding reaction or in folding intermediates.⁴⁵ In addition, the possibility of fast, light-induced *cis*–*trans* isomerization reactions around thiopeptide bonds^{18–20} should allow the use of thioamide groups incorporated into helical structures to trigger fast time-resolved measurement of protein folding reactions.

Materials and Methods

Peptide Synthesis. Peptides were synthesized using standard 9-fluorenylmethyloxycarbonyl (Fmoc) chemistry on an Applied Biosystems 433A synthesizer. Couplings were performed with

HBTU/HOBt on Tentagel S RAM resin (Rapp Polymere). Cleavage from the resin and side chain deprotection were achieved with 96/2/2 TFA/TIPS/H₂O (*v/v/v*). For the synthesis of the thioamide substituted peptides a special protocol was used.⁴⁶ All peptides were purified to >95% purity by HPLC on an RP C8 column. Purity was checked by analytical HPLC, and the identity was verified by MALDI mass spectrometry.

Concentration Determination. Concentrations of peptides with a thioamide bond or a phenylalanine were determined spectroscopically using the extinction coefficients $\epsilon_{266} = 12\,000\text{ M}^{-1}\text{ cm}^{-1}$ ²¹ and $\epsilon_{257} = 197\text{ M}^{-1}\text{ cm}^{-1}$,⁴⁷ respectively. A Gly-Tyr sequence was added to the C-terminus of peptides containing neither a phenylalanine nor a thio group. The addition of the tyrosine residue facilitates spectrometric concentration determination ($\epsilon_{275} = 1450\text{ M}^{-1}\text{ cm}^{-1}$) and does not significantly affect the far-UV CD spectrum, when introduced at the C-terminus and separated by a glycine.³⁰ Concentrations were additionally checked by one-dimensional NMR measurements (see Supporting Information) and differed by less than 5% from the spectroscopically determined concentrations.

CD Measurements. CD measurements were performed on an Aviv DS62 spectropolarimeter. The spectra were independent of the irradiation length, indicating that no significant *trans* to *cis* isomerization of the thioamide bond occurred during the measurements. Thermal unfolding transitions were monitored at 222 nm and were fully reversible. Peptide concentrations were between 50 and 100 μM .

NMR Measurements. Proton NMR spectra were recorded on a Bruker DRX 600 spectrometer at 2 °C and referenced to 2,2-dimethyl-2-silapentane-5-sulfonate (DSS) at 0.0 ppm. A detailed description of the one- and two-dimensional experiments is given in the Supporting Information.

Determination of Helix Stability. Lifson and Roig theory³² was applied as described by Doig et al.³⁵ to determine the effect of a thio substitution on helix stability. The Lifson–Roig parameters for the alanine-based helices were taken from ref 31. The matrix products were calculated without any further approximations using MATLAB 7.2 (MathWorks). To estimate w^{thio} , the w of the alanine residue following the thioamide bond was adjusted to give the observed helix content.

Acknowledgment. We thank Horst Kessler and Judith Habazettl for discussion and Annett Bachmann for comments on the manuscript. This work was supported by a grant from the Schweizerischen Nationalfonds and by the Munich Center for Integrative Protein Science (CIPSM).

Supporting Information Available: Experimental details of the NMR measurements, CD spectra of the thioylated peptides in the near-UV CD region, and a table with chemical shift assignments and temperature coefficients. This material is available free of charge via the Internet at <http://pubs.acs.org>.

JA8015044

(44) Blaber, M.; Zhang, X. J.; Matthews, B. W. *Science* **1993**, *260*, 1637–1640.

(45) Sánchez, I. E.; Kiefhaber, T. J. *Mol. Biol.* **2003**, *334*, 1077–1085.

(46) Wildemann, D.; Drewelow, M.; Fischer, G.; Schutkowski, M. *Chem. Commun.* **1999**, 1809–1810.

(47) Fasman, G. D. *Handbook of Biochemistry and Molecular Biology*, 3rd ed.; CRC Press: 1976; p 186.

Supporting Information

Experimental Details of NMR Experiments

Proton NMR spectra were recorded on a Bruker DRX 600 spectrometer with peptide concentrations of about 300 μM in buffered solution containing 5 % (v/v) D_2O . The spectra were processed with XWINNMR (Bruker) and evaluated with Sparky (T. D. Goddard and D. G. Kneller, UCSF). One-dimensional spectra were recorded with solvent suppression by excitation sculpting using 180° water-selective pulses. Typically 8192 complex data points were recorded and multiplied with a squared cosine bell window function before Fourier transformation.

One-dimensional spectra were acquired between 2 $^\circ\text{C}$ and 35 $^\circ\text{C}$ at pH 7.0 to determine the temperature coefficients. Additional spectra with a jump-and-return water suppression were recorded to observe the thioamide proton, which is in exchange with solvent. 2048 complex data points were recorded and processed as above. Temperature was calibrated with methanol and ethylene glycol reference samples.

The TOCSY spectrum was acquired with a pulse program incorporating a WATERGATE pulse sequence for solvent signal suppression and the MLEV-17 sequence was applied during the TOCSY mixing-time of 110 ms. 4096 and 450 complex data points were recorded in direct and indirect dimension, respectively. Data in the indirect dimension data were zero-filled. Data were multiplied with squared cosine bell window functions in both dimensions.

NOESY spectra at pH 7.0 and pH 5.0 were acquired with a WATERGATE pulse sequence for solvent signal suppression and a mixing time of 460 ms. 4096 and 328 complex data points were recorded in direct and indirect dimension, respectively. Data in the indirect dimension data were zero-filled. Data were multiplied with squared cosine bell window functions in both dimensions.

Peptide concentrations were determined using standard one-dimensional experiments at 30 $^\circ\text{C}$. A calibration curve was measured, which relates the integrated signal of the acetyl protons of the peptide **cG** (concentration determined by UV absorption) to the integrated signal of the methyl protons of 100 μM trimethylsilyl propionate (TMSP). The linear calibration curve was used to determine peptide concentrations by evaluating the ratio of the integrals of the well resolved N-terminal acetyl groups and the one of 100 μM TMSP that was added as internal reference.

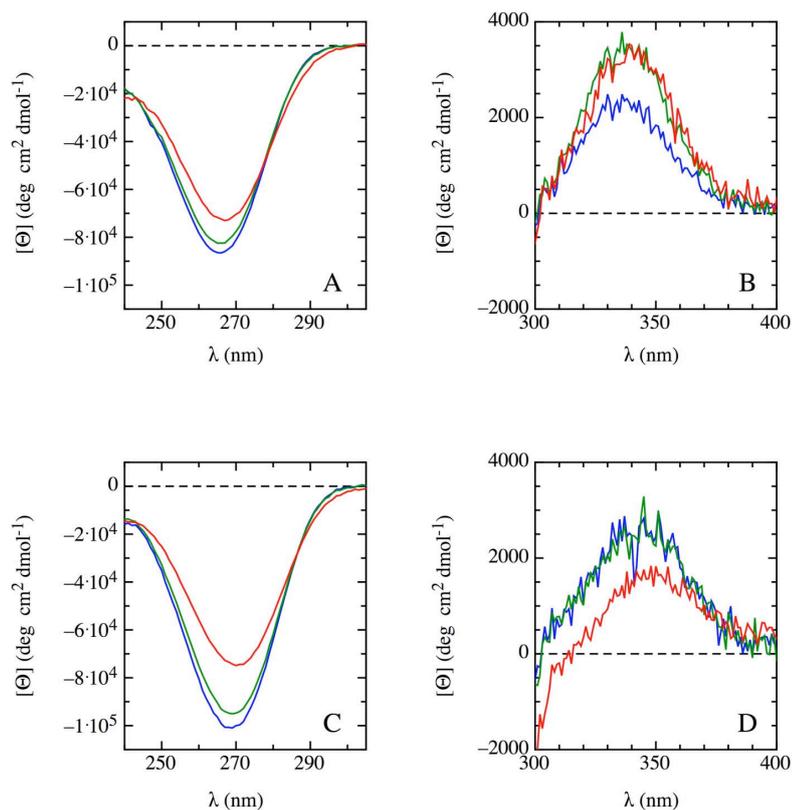


Figure S1: Near-UV CD spectra of $c\psi$ (A, B) and $n\psi$ (C, D) at 5 °C (blue), 22.5 °C (green) and 90 °C (red). Panels A and C display the bands of the $\pi_0\pi^*$ -transition. Panels B and D show the $n\pi^*$ - transition. The CD signal was normalized to the molar concentration of thioamide bond, i.e. to the peptide concentration. Spectra were recorded in 10 mM potassium phosphate at pH 7.0.

Table S1: NMR characterization of $\mathbf{n\psi}$.

residue	H ^N	$\Delta\delta/\Delta T^a$	H ^{α}	H ^{β}	other
Ac-			1.966		
Asp 1	8.365	-11.2	4.509	2.651, 2.499	
Phe 2	8.433	-11.5	5.037	3.399, 3.142	H2,6 7.321 H3,5 7.391 H4 7.342
Ala 3	9.826 ^b	-8.1	4.637	1.547	
Ala 4	8.282	-8.6 ^c	4.184	1.432	
Ala 5	8.268	-8.6 ^c	4.216	1.478	
Lys 6	7.682	2.4	4.199	1.913	HC 1.568, 1.469 HD 1.690 HE 2.963
...					
Lys 11	8.071		4.159	1.895	HC ^d HD 1.690 HE 2.978
...					
Lys 16	7.988		4.219	1.898, 1.829	HC 1.549 HD 1.699 HE 2.999
-NH ₂	7.462 7.303				

Chemical shifts in ppm relative to internal DSS in 10 mM potassium phosphate, pH 7.0 at 2 °C.
^aTemperature coefficient of amide proton in ppb/°C measured between 2 - 25 °C.
^bBroad resonance due to exchange.
^cThe resonances of the Ala4 and Ala5 amide protons are not separated in the 1D ¹H spectrum over the complete temperature-range.
^dNot resolved.

NMR Structure of a Monomeric Intermediate on the Evolutionarily Optimized Assembly Pathway of the Trimeric Foldon Domain

Judith Habazettl¹, Andreas Reiner¹ & Thomas Kiefhaber^{2*}

¹Biozentrum der Universität Basel, Division of Biophysical Chemistry, Klingelbergstrasse 70, CH-4056 Basel, Switzerland

²Munich Center for Integrated Protein Science & Chemistry Department, Technische Universität München, Lichtenbergstrasse 4, D-85747 Garching, Germany

Classification: Biophysics, Chemistry

Running title: Structure of an Assembly Intermediate

*corresponding author. Phone: ++49 89 289 13420; fax: ++49 89 289 13416;

e-mail: t.kiefhaber@tum.de

Abstract

The foldon domain is an evolutionarily optimized trimerization module that is essential for assembly of the large, trimeric protein fibrin. Monomers consisting of the 27 amino acids comprising a single foldon domain subunit spontaneously assemble into a native trimer. During the folding/association process a monomeric intermediate is formed on the submillisecond time scale, which is the starting point for two consecutive very fast association reactions. Disruption of an intermolecular salt bridge by mutation leads to a monomeric state that structurally resembles the kinetic intermediate. NMR spectroscopy revealed native topology of the monomeric intermediate with defined hydrogen bonds and side chain interactions, but largely reduced stability compared to the native trimer. The native topology in the intermediate is mainly determined by a hydrophobic cluster formed by five amino acids from different parts of the sequence. This structural preorganization leads to an asymmetric charge distribution on the surface, which can direct rapid subunit recognition. All interactions within the monomer are more flexible compared to the trimer and the intermediate is in equilibrium with the unfolded state on the 100 μ s time scale. The low stability of the intermediate leads to a large gain in energy upon formation of the intermolecular subunit interactions, which serves as driving force for rapid assembly. These results indicate different free energy landscapes for folding of small oligomeric and monomeric proteins, which typically avoid the transient population of intermediates.

Introduction

Fibrin is a trimeric rod-like structural protein from bacteriophage T4 which consists of an N-terminal anchor domain (residues 1-46), a large central coiled-coil part (residues 47-456), and a small C-terminal globular domain (residues 457-486) (1). The C-terminal domain is essential for fibrin folding and assembly *in vivo* and *in vitro* and was termed foldon domain (1-3). The foldon domain was proposed to be an evolutionarily optimized trimerization/folding motif, since its only known function is to promote folding and association of fibrin (3). This idea is supported by the observation that a covalently attached foldon domain can substitute for the effect of molecular chaperones during assembly of other trimeric proteins from phage T4 (4) and induces trimerization of other proteins (5-7). Each subunit of the foldon domain consists of an N-terminal extended structure, a β -hairpin, a 3_{10} -helix near the C-terminus and three unstructured C-terminal residues (Fig. 1A,B). Three identical subunits assemble into a β -propeller structure, which is stabilized by hydrophobic interactions involving Trp20 of each subunit, intermolecular salt bridges between Glu5 and Arg15, and intermolecular backbone hydrogen bonds between Tyr13 and Arg15 (Fig. 1A) (1, 8). The first 27 amino acid residues of the foldon domain (residues 457-483 in fibrin, in the following denoted as residues 1-27) were chemically synthesized and shown to form the same trimeric β -propeller structure as in full-length fibrin (8). The 3×27 amino acid foldon trimer undergoes a two-state equilibrium unfolding transition from folded trimer to unfolded monomers and is very stable with $\Delta G^0 = -89.2 \pm 0.6$ kJ/mol at 20 °C, which corresponds to $\Delta G = -29.7$ kJ/mol at 5 μ M monomer concentration (8).

The small size of the foldon domain, which makes it accessible to chemical synthesis, its simple fold and its high stability make it an ideal model for detailed mechanistic studies

on a folding reaction linked to an intermolecular assembly processes. Kinetic studies showed that 19 % of the unfolded molecules assemble slowly, limited by proline isomerization reactions (8). The remaining unfolded molecules have all Xaa-Pro bonds in the native isomerization state and assemble via a sequential pathway comprising several rapid folding and association reactions (Scheme 1).



Scheme 1

Within the dead-time of stopped-flow mixing (~1 ms) a monomeric intermediate (I) is formed which reacts in two consecutive bimolecular steps first to a dimeric (D) and then to a trimeric (T) intermediate with rate constants of $1.9 \times 10^6 \text{ M}^{-1} \text{ s}^{-1}$ and $5.4 \times 10^6 \text{ M}^{-1} \text{ s}^{-1}$, respectively (8). Both reactions are about as fast as the fastest known bimolecular folding reaction of a natural dimeric protein reported for dimerization of arc repressor ($k=6.2 \times 10^6 \text{ M}^{-1} \text{ s}^{-1}$) (9). The trimeric intermediate is rapidly converted to the final native state in a concentration-independent reaction with a rate constant of 210 s^{-1} . This reaction becomes rate-limiting at high protein concentrations. This sequence of very fast unimolecular and bimolecular reactions allows the foldon domain to assemble very rapidly with a half-life of about 100 ms at $10 \mu\text{M}$ monomer concentration and of about 3 ms at high protein concentrations ($\geq 100 \mu\text{M}$). An especially intriguing feature of the folding pathway is the rapid formation of a monomeric intermediate, which is surprising considering the small size of a subunit. This observation is in contrast to the folding behavior of small monomeric proteins, which typically avoid to transiently populate intermediates (10), although they seem to be important for folding (11, 12). Therefore, it is very likely that

the monomeric intermediate is the key to understanding the rapid and efficient assembly process of the foldon domain. Due to the transient nature of the intermediate it is not directly accessible to structural investigations. Earlier NMR studies on a monomeric state populated at pH 2 (A-state) revealed a native β -hairpin and unstructured N- and C-terminal regions (13). However, the fluorescence spectra of the A-state and the folding intermediate differ significantly indicating major structural differences (8).

To obtain a better model for the monomeric folding intermediate we destabilized intersubunit interactions by mutation in order to populate a monomeric state of the foldon domain at neutral pH. The results show that replacement of Glu5 by arginine (E5R variant) leads to a strong destabilization of the native trimer by removing an intermolecular salt bridge. The E5R variant forms a monomeric state at neutral pH with essentially the same fluorescence and CD properties as the transient folding intermediate. This allowed us to determine the NMR structure of the monomeric state at pH 7 in order to characterize the structural and dynamic properties of the rapidly formed transient folding intermediate.

Results & Discussion

Comparison of foldon E5R with the transient folding intermediate. The native foldon trimer is stabilized by three intermolecular salt bridges between Glu5 and Arg15 of adjacent subunits (Fig.1A). NMR studies revealed a pK_a -value of 1.7 for Glu5 in the native trimer indicating that each salt bridge contributes about 16 kJ/mol to trimer stability (13). To obtain a monomeric state of the foldon domain at pH 7 we replaced Glu5 by arginine (foldon E5R), which removes the salt bridge and additionally introduces an electrostatic repulsion between the positive charges at Arg5 and Arg15. The

fluorescence and CD properties of the E5R variant reveal major structural differences compared to the native wild-type trimer (Fig. 2). Fluorescence of the single tryptophan residue at position 20 of each subunit is a good probe to test for trimer formation, since Trp20 is located at the subunit interface and is completely solvent-shielded in the native state resulting in a largely blue-shifted fluorescence emission maximum at 318 nm. The emission maximum is significantly red-shifted to 336 nm in the E5R variant and the fluorescence intensity is decreased (Fig. 2A), indicating increased solvent exposure of Trp20 and arguing for the disruption of the trimer interface. Fluorescence contributions from tyrosine residues around 303 nm are not detectable in the E5R variant, which points at a folded structure with efficient fluorescence resonance energy transfer from the two tyrosine residues at position 2 and 13 to Trp20. The fluorescence properties of the E5R variant differ from those of the GdmCl-unfolded state, which has weaker intensity, a further red-shifted maximum ($\lambda_{\text{max}}=342$ nm) and a shoulder around 303 nm arising from contributions of the tyrosine residues. The E5R variant gives virtually the same fluorescence spectrum as the monomeric intermediate that transiently populates during the assembly process in the wild-type protein (Fig. 2A) indicating that the E5R variant has the same structural properties as the transient folding intermediate. This finding is supported by the far-UV CD spectrum of the E5R variant, which is identical to the spectrum of the folding intermediate but differs from both the spectra of the native trimer and from the unfolded state (Fig. 2B). These results show that foldon E5R is better suited to elucidate the structural properties of the monomeric intermediate than the A-state which shows reduced fluorescence intensity (Fig. 2A) (8).

To obtain information on the stability of the E5R variant we monitored thermal unfolding at different protein concentrations by the change in CD at 228 nm (Fig. 2C). The

unfolding transition is broad and concentration-independent as expected for a monomeric protein consisting of only 27 amino acids. The unfolding transition of the wild-type trimer, in contrast, is much steeper and shifted to higher temperatures (Fig. 2C). In addition, unfolding of the wild-type protein shows a strong concentration-dependence in accordance with an unfolding transition from a trimer to three monomers. The ill-defined baselines of the unfolding curves of the E5R variant make it difficult to obtain accurate thermodynamic data on the stability of foldon E5R.

These results indicate that E5R variant forms a monomeric folded state, which resembles the transient folding intermediate of the wild-type protein in its spectroscopic properties. Foldon E5R thus represents a good model to characterize the structural, dynamic and energetic properties of the monomeric folding intermediate.

NMR structure of foldon E5R. We used NMR spectroscopy to determine the solution structure of the foldon E5R variant. Thermal unfolding transitions suggested a monomeric state of foldon E5R at protein concentrations up to 30 μM (Fig. 2C). Two-dimensional TOCSY experiments gave identical spectra at 30 μM and 200 μM protein concentration indicating the absence of trimers even at higher protein concentrations. The chemical shifts of foldon E5R in the HSQC spectrum are different from those of the wild-type trimer (Fig. 3) but similar to those of the previously characterized monomeric A-state (13), which is in slow-exchange with the native trimer at pH 2. In contrast to the A-state, we did not observe a second set of cross peaks in the HSQC spectrum of foldon E5R at any investigated protein concentration and all NOEs could be fulfilled with a monomeric structure. These results indicate the population of less than 5 % of any trimeric form in

the foldon E5R variant, which shows that the E5R substitution has the expected stronger destabilizing effect on the trimer interface compared to the protonation of Glu5 at low pH.

The NMR structure of the monomeric E5R variant was determined based on ^1H TOCSY and NOESY spectra (pdb accession code: ???). Some cross-peaks were assigned with the help of ^{15}N -edited spectra. All NMR spectra contain only a single resonance line per ^1H or ^{15}N nucleus, indicating that putative alternative protein conformations are in rapid exchange on the chemical shift time scale. Ensembles of structures were calculated using 139 medium and long-range, 141 sequential and 80 intra-residue NOE restraints. Details of the structure calculation and of the refinement procedure are given in the Supporting Information (SI). The structure of the E5R variant has the same topology and very similar intramolecular interactions as a subunit in the native trimer (Fig. 1), although the chemical shift dispersion of foldon E5R is smaller compared to the wild-type trimer. The β -hairpin region from Ala12 to Leu23 is well-defined in the E5R variant, with the expected β -sheet NOE pattern (14), except for some cross peaks involving the weak amide resonances of Val14, Asp17 and Val21. The hairpin of foldon E5R has essentially the same conformation as in the native foldon trimer (Figs. 1B,C). This is reflected by the pair-wise root mean square deviation (RMSD) of 0.7 Å for the backbone atoms and 1.5 Å for all heavy atoms in the hairpin between a subunit in the trimer and the E5R variant. Also the hydrophobic residues Ala12, Tyr13, Val14, Val21, Leu22 and Leu23 at the ends of the β -hairpin have native side chain orientations. In addition, hydrophobic contacts between Tyr13, Trp20 and Leu22 across the two strands indicate a similar, but slightly weaker, hairpin twist compared to the native trimer.

Long-range NOEs are further observed from mostly weak contacts between the side chains of Pro4/Trp20, Ala6/Tyr13, Pro7/Tyr13 and Pro7/Trp20. These interactions show that the extended N-terminal region in the E5R variant packs against one side of the hairpin and participates in a hydrophobic cluster involving residues of the N-terminal region (Pro4, Ala6, Pro7), both strands of the hairpin (Tyr13, Trp20) and the C-terminal region (Leu22; Fig. 1B,C). The side chains of Tyr2 and Ile3 that are involved in hydrophobic contacts along the trimer symmetry axes have different orientations in foldon E5R. The ring of Tyr2 is flipped back to Ile3 and Pro4 and points towards the hairpin, which is reflected by NOE contacts from the side chain of Ile3 to Trp20 and Lys16. The complete region from Gly1 to Gln11 of foldon E5R is less well defined than the hairpin region, with RMSD-values of 1.9 Å for the backbone atoms and 2.4 Å for all heavy atoms (Table S1). In the native trimer the peptide bonds involving Pro4 and Pro7 are exclusively in the *trans* conformation. The Ile3-Pro4 bond shows 10–14 % *cis* isomer in the E5R variant, in agreement with structural flexibility in this region. The spectra did not show any signature of the *cis* isomer at the Ala6-Pro7 bond suggesting the importance of a *trans* peptide bond at this position for the interactions of the N-terminal region with the hairpin.

We did not observe any backbone NOEs indicative for a native helical turn in the C-terminal region. However, interactions involving Val14, Val21 as well as the side chains of Leu23, Thr25 and Phe26 are detected indicating the presence of a defined structure in this region, similar to the 3_{10} -helical turn in the trimer. The side chains of Phe26 and Leu27, which are part of the hydrophobic interactions along the trimer symmetry axis, have different orientations in foldon E5R compared to the native trimer. The C-terminal region appears to be less structured than the N-terminal region and the

hairpin with RMSD-values of 1.7 Å for the backbone atoms and 2.2 Å for all heavy atoms. The RMSD-values are much smaller (0.3 Å and 1.2 Å, respectively) when the backbone atoms of the C-terminal region are fitted onto each other. A similar effect is observed for the N-terminal region indicating that both terminal regions form native-like structure but their orientation relative to the hairpin is less well defined than in the native trimer.

Characterization of intramolecular interactions by chemical shift analysis. A protein structure determined by NMR typically includes many redundant distance constraints. Medium or long range NOEs are usually only absent for unstructured loops and terminal ends. The distance constraints of foldon E5R define a native-like topology with specific side chain and backbone interactions. The structure of the hairpin region is particularly well defined, whereas distance constraints in the N- and C-terminal regions are more scarce. To gain further information on the presence of specific interactions we studied the chemical shifts, which are usually more sensitive to residual structure than NOEs (15). Ring current effects in the vicinity of aromatic side chains are the major source for chemical shift deviations relative to the random coil values, $\Delta\delta(^1\text{H}^{\text{N}})$ (16, 17). Hydrogen bonds and exchange with water represent an additional source for chemical shift changes for amide protons. Almost all $\Delta\delta(^1\text{H}^{\text{N}})$ -values in foldon E5R are closer to their random coil values compared to the trimer but the $\Delta\delta(^1\text{H}^{\text{N}})$ -values of amide protons in the β -sheet show a similar chemical shift pattern as in the trimer (Fig. 4A). Especially resonances from Val14 and Val21, which are involved in hydrogen bonds, show strong downfield shifts. With increasing temperature the chemical shift deviations decrease and approach their random coil values at high temperature. Only minor chemical shift perturbations are observed for the amide protons from Gly1 to Ile3 and Leu24 to Leu27 indicating the lack

of stable structures in the terminal regions and the absence of ring current effects, which are only observed near aromatic residues in tightly packed structures. In the native foldon trimer, in contrast, the C-terminal 3_{10} -helices of the three monomers are packed against each other via hydrophobic contacts involving an aromatic ring (residues Thr25, Phe26 and Leu27), which causes large $\Delta\delta(^1\text{H}^N)$ values for Thr25 (Fig. 4A).

Several side chain protons in the E5R variant show chemical shift deviations. The protons of Ala6 CH_3 , Pro7 $\text{H}\delta^3/\text{H}\gamma^{1,2}$ and Leu22 CH_3 are significantly shifted from their random coil values indicating their vicinity to the ring systems of Trp20 and Tyr13 (Fig. S3), in agreement with their involvement in the hydrophobic cluster formed between residues of the N-terminal region and both strands of the hairpin (Fig. 1C).

Temperature-dependence of the chemical shifts. The temperature dependence of chemical shifts, $\Delta\delta/\Delta T$ gives further insight into the presence of weak interactions (18). In general, proton chemical shifts display linear temperature dependencies in the native and unfolded state. Deviations from typical $\Delta\delta/\Delta T$ -values can be caused by altered rotation rates, by specific orientations of aromatic side chains, by formation of hydrogen bonds and by formation of defined interactions. These effects may also lead to a sigmoidal temperature dependence in the region of a structural transition. Solvent-exposed amide protons typically have a much stronger temperature dependence compared to non-exchangeable protons with $\Delta\delta/\Delta T$ -values ranging from -5.8 to -9.1 ppb/K. Formation of hydrogen bonds leads to typical $\Delta\delta/\Delta T$ -values >-4.5 ppb/K (19). The $\Delta\delta/\Delta T$ -values of amide protons in foldon E5R fall in different categories (Fig. 4B). Although the NOEs revealed hydrogen bonds between Val14 and Val21 and between Lys16 and Glu19 in the

β -hairpin region, only the Glu19 NH shows a typical $\Delta\delta/\Delta T$ -value of +4.7 ppb/K. The $^1\text{H}^{\text{N}}$ chemical shifts of Val14 and Val21 show identical strong and non-linear temperature-dependencies, which indicates unfolding of the central β -hairpin structure with increasing temperature (Fig. S2). These findings suggest the presence of weak hydrogen bonds within the β -hairpin that break upon heating.

All amide resonances in the C- and N-terminal regions, except for Ala6 and Asp9, depend linearly on temperature with typical random coil $\Delta\delta/\Delta T$ -values (Fig. S2). With increasing temperature Ala6 and Asp9 show almost identical non-linear shifts towards their random coil values suggesting that they are also involved in residual structure that unfolds at high temperature. Similar temperature dependencies were previously observed for residues in loop regions of proteins containing a hydrogen bonding donor (19). Pro7 and Pro4 are possible hydrogen bonding acceptors for Ala6 and Asp9, respectively.

The temperature coefficients of side chain resonances of Ala6 (CH_3) and Pro7 ($\delta^3\text{H}$, $\delta^2\text{H}$, $\gamma^{1,2}\text{H}$) are very large and non-linear, which indicates that their vicinity to the ring systems of Trp20 and Tyr13, respectively, is lost with increasing temperature (Fig. 5). Above 320 K the resonances of Ala6 and Pro7 reach their unperturbed values as seen from the comparison with the temperature dependencies of the side chains of Pro4 and Ala12 as reference (Fig. 5). The methyl groups of Leu22, which are located above the ring of Tyr13, exhibit temperature coefficients of 2.3 ppb/K and 5.3 ppb/K, indicating that they are also forming specific interactions. These observations confirm the presence of a well-defined hydrophobic cluster involving residues from the N-terminal region (Ala6, Pro7), from the central part (Tyr13, Trp20) of the β -hairpin and from the C-terminal end (Leu22)

of the hairpin. Our results indicate that the cluster is stabilized by backbone hydrogen bonds and unfolds upon increasing temperature. Analysis of the NOEs has suggested that Pro4 is also part of this cluster. The lack of perturbed chemical shifts of the Pro4 protons is probably due to the larger distance and/or different orientation between Pro4 and the aromatic ring systems of Tyr13 and Trp20. Formation of the hydrophobic cluster involving residues from all regions of the foldon sequence seems to be the key to the formation of the native topology in the monomeric intermediate (Fig. 1C).

The amide protons of Thr25 (-3.2 ppb/K) and Phe26 (-3.7 ppb/K) have $\Delta\delta/\Delta T$ -values typical for weakly hydrogen-bonded amides. In combination with the observation of a few NOEs in the C-terminal region these results suggest some helical structure in the C-terminal region without stable hydrogen bonds. This interpretation is confirmed by the R_1 and R_2 values, which indicate that the $^1\text{H}^{\text{N}}\text{-}^{15}\text{N}$ vectors of residues 24 to 26 are less flexible than expected for unstructured terminal residues in proteins but more flexible than those of residues in the β -hairpin and in stable secondary structures.

pH-dependence of chemical shifts. The NOEs together with the chemical shift analysis show that all parts of foldon E5R have native-like structure. The presence of a hydrophobic cluster involving residues from the N-terminal region and from different parts of the β -hairpin distinguishes the structure of the E5R variant from the previously characterized monomeric A-state, which is formed at low pH and has only the β -hairpin region formed (13). This difference is confirmed by TOCSY experiments on the E5R variant at different pH. Upon lowering the pH from 7 to pH 2 the chemical shifts of the side chain resonances of Ala6, Pro7, Tyr13 and Trp20 experience a similar shift towards their random coil values as observed when the temperature is increased from 290 K to

320 K (Fig. S4). This result demonstrates that the interactions between the N-terminal region and the hairpin are lost both at low pH and at high temperature, which explains the major difference in the fluorescence spectrum of the A-state and the E5R variant. No obvious electrostatic interactions between the hairpin and the N-terminal region can be inferred from the structure and none of the pK_a values of the acidic side chains coincides with the pK_a value for the structural transition (data not shown) indicating that the loss of the hydrophobic cluster may be due to a general electrostatic repulsion caused by net positive charge at low pH. A similar effect was observed for the structure of the acid molten globule state of cytochrome *c* (20).

Backbone dynamics measured by ^{15}N relaxation experiments. The atomic positions of residues in the N- and C-terminal regions and of residues in the hairpin turn are significantly less defined by distance and dihedral angle constraints than the rest of the protein. To determine whether the lack of consensus of the calculated structures in these regions is due to conformational flexibility, we performed ^{15}N relaxation experiments (Fig.4 C,D). The R_1 and R_2 values of residues in the N-terminal (Tyr2, Ile3, Arg5) and C-terminal (Thr25, Phe26, Leu27) regions are decreased, typical for less structured or unstructured protein ends (21). For the rest of the protein R_1 displays remarkable little variation along the sequence, suggesting that these backbone atoms undergo similar degrees of motion on the picoseconds to nanoseconds time scale. Asp9, Val14, Lys16, Asp17, Gly18, Glu19, Val21, and Leu22 show significantly increased R_2/R_1 values, which are too large to be accounted for by internal motions. This indicates chemical exchange processes with motions in the nanosecond to microsecond time scale. Applying different spin-lock powers during transverse relaxation of ^{15}N gave a time constant below 40 μs for

the conformational exchange process for residues Gly18 and Glu19 and time constants of $98 \pm 10 \mu\text{s}$ for Asp9, $88 \pm 12 \mu\text{s}$ for Val14 and $78 \pm 17 \mu\text{s}$ for Leu22.

Residues Gly10 and Ala12 could be used to estimate a global molecular rotational correlation time, τ_c , due to their small temperature coefficients and their pH-independence. The results gave τ_c of 2.7 ns consistent with isotropic rotational correlation times for a monomer of this size (22).

Structural basis for rapid assembly of the foldon trimer. The CD and fluorescence properties of the foldon E5R variant reveal that it resembles the transiently formed monomeric folding intermediate in its structural properties and thus provides a good model to elucidate the structural basis for the fast association reactions during assembly of the foldon domain. NMR characterization showed essentially native topology and intramolecular side chain interactions in the intermediate, which demonstrates that the intermolecular interactions of the native trimer are not required to define the monomer structure. The main effect of trimerization is a large increase in stability and a decrease in conformational flexibility. These findings suggest that the native topology serves as template for the fast bimolecular association reactions, which may be especially important for assembly processes with consecutive association reactions. During trimer formation two monomeric intermediates react to form a dimer but a monomer also reacts with a dimer to form the trimer (Scheme 1). Thus, the structure of the monomer has to be optimized to efficiently interact with different partners. Natively pre-organized recognition sites for intermolecular interactions most likely increase the chance for productive monomer-monomer and monomer-dimer encounters. The structure of the monomeric intermediate reveals the nature of the recognition sites for rapid

intermolecular interactions. The extended hydrophobic cluster formed by residues in the N-terminal region (Pro4, Ala6 and Pro7), in the central part of the β -hairpin (Tyr13, Trp20) and in the C-terminal region (Leu22) (Fig. 1C) defines the native topology which leads to the formation of well-defined positively and negatively charged regions on opposite sides of the foldon monomer (Fig. 1E,F). This asymmetric charge distribution is well suited to provide recognition sites for rapid assembly into a dimer. In addition, the two oppositely charged sides of the foldon monomer have hydrophobic regions in their C-terminal regions around Leu22 and Phe26, which may support the correct orientation of the monomers. This arrangement of opposite charges and hydrophobic regions should further be able to guide the third monomer to the correct orientation relative to the dimer during formation of the native trimer. A similar recognition between positively and negatively charged surfaces was identified as crucial for efficient interaction between the folded states of barnase and its inhibitor barstar (23).

The fast consecutive bimolecular reactions during foldon assembly with rate constants around $5 \times 10^6 \text{ M}^{-1}\text{s}^{-1}$ in the presence of 0.6 M GdmCl are in the same range as the fastest reported rate constants for folding of the small homodimeric arc repressor (2x56 amino acids) (9) and of a designed heterodimeric leucine zipper (2x29 amino acids) (24). In contrast to the foldon domain, no evidence for the formation of partially folded monomers was found for these proteins. Rapidly formed monomeric intermediates have, however, been observed during folding of larger homo- and heterodimeric proteins (25).

Despite the presence of the native topology and defined intrachain interactions, the monomer is only marginally stable and is in rapid equilibrium with the unfolded state. Formation of the native intermolecular interactions upon assembly into the trimer leads to

a large gain in free energy, which serves as a driving force for the rapid assembly reactions. This large increase in stability is expected, since the loss in conformational entropy has already occurred upon formation of the native topology in the monomer. Thus, formation of the intermolecular interactions and stabilization of the intramolecular interactions mainly results in a gain in enthalpy during assembly. These considerations indicate a perfect balance between native-like topology, required for efficient recognition of the monomers at the different stages during assembly, and low stability, required to provide a driving force for rapid assembly.

Comparison of the folding/assembly process of the foldon domain to folding of small monomeric proteins reveals major differences in the free energy landscapes. Small monomeric proteins typically fold in the absence of detectable intermediate states. However, the frequent observation of non linear rate-equilibrium free energy relationships revealed the importance of obligatory high-energy intermediates in apparent two-state folding (11, 12). It is obviously important for monomeric proteins to keep the stability of partially folded states low in order to optimize the free energy landscape for fast folding (26) and to prevent aggregation of partially folded states. Optimized folding landscapes for oligomeric proteins seem to include highly populated monomeric intermediates. These intermediates can provide specific recognition sites for intersubunit interactions, which probably facilitates formation and stabilization of specific intermolecular interactions upon encounter of the subunits. Comparison with the folding reactions of monomeric proteins further reveals that the monomeric foldon intermediate has similar structural properties as transition states for folding of small single domain proteins, which often have native-like topology, but loosely formed side chain interactions with the major gain in enthalpy typically occurring between the transition state and the native state.

The foldon monomer represents the smallest known natural sequence that folds into a well-defined monomeric tertiary structure with specific side chain and backbone interactions. It thus provides a good model for experiments and simulations on the structure and dynamics of the monomer and on the assembly reactions from the monomer to the native trimer.

Materials and Methods

Chemical Synthesis and Expression of Foldon E5R. The foldon variant E5R with the sequence GYIPRAPRDG QAYVRKDGEW VLLSTFL was either synthesized using 9-fluorenylmethyloxycarbonyl (Fmoc) chemistry or expressed in *E. coli* and purified as described (8, 13). Site-directed mutagenesis was used to change Glu5 to Arg, and to replace the thrombin cleavage site by an enterokinase cleavage site (DDDDKIG⁻¹S⁰). Protein concentrations were determined using $\epsilon_{280} = 8250 \text{ M}^{-1} \text{ cm}^{-1}$ (27).

Fluorescence and CD measurements. All measurements were performed in 10 mM potassium phosphate, pH 7.0 at 20 °C. Fluorescence measurements were performed with an Aminco Bowman Series2 spectrofluorimeter (SLM Aminco) with 5 μM protein. CD measurements were performed on an Aviv DS62 spectropolarimeter with 50 μM protein. The CD spectrum of the monomeric intermediate of wild-type foldon was obtained in refolding experiments in an Applied Photophysics π^* CD stopped-flow instrument with a final monomer concentration of 20 μM in 0.58 M GdmCl, 10 mM potassium phosphate, pH 7.0.

NMR measurements. NMR measurements were performed with 200 μM foldon E5R and 210 μM ^{15}N -labelled foldon E5R in 10 mM potassium phosphate, pH 7.0, containing 0.02 % (w/v) NaN_3 and 5 % (v/v) D_2O .

2,2-Dimethyl-2-silapentane-5-sulfonic acid (DSS) was added as internal standard. NMR experiments were carried out on a Bruker DRX 600 spectrometer at a temperature of 296.7 K. Two-dimensional homonuclear TOCSY and NOESY experiments and a three-dimensional ^{15}N -edited TOCSY experiment were used for ^1H and ^{15}N assignments. Further details of the NMR experiments and of the structure calculations are given in the SI. Calculated structures were visualized using MOLMOL (28). The Bruker software XWINNMR and NMRPipe were used to process the NMR data (29). Spectra were displayed and analyzed with the program SPARKY (30).

Acknowledgements. This work was supported by grants from the Swiss National Science Foundation, the DFG (SFB 749) and the Munich Center for Integrated Protein Science (CIPSM).

References

1. Tao, Y., Strelkov, S. V., Mesyanzhinov, V. V. & Rossmann, M. G. (1997) Structure of bacteriophage T4 fibrin: a segmented coiled coil and the role of the C-terminal domain. *Structure* 5: 789-798.
2. Letarov, A. V., Londer, Y. Y., Boudko, S. P. & Mesyanzhinov, V. V. (1999) The carboxy-terminal domain initiates trimerization of bacteriophage T4 fibrin. *Biochemistry (Mosc.)* 64: 817-823.
3. Boudko, S. P., *et al.* (2002) Domain organization, folding and stability of bacteriophage T4 fibrin, a segmented coiled-coil protein. *Eur. J. Biochem.* 269: 833-841.
4. Miroshnikov, K. A., *et al.* (1998) Engineering trimeric fibrous proteins based on bacteriophage T4 adhesins. *Protein Eng.* 11: 329-332.
5. Frank, S., *et al.* (2001) Stabilization of short collagen-like triple helices by protein engineering. *J. Mol. Biol.* 308: 1081-1089.
6. Yang, X., *et al.* (2002) Highly stable trimers formed by human immunodeficiency virus type 1 envelope glycoproteins fused with the trimeric motif of T4 bacteriophage fibrin. *J. Virol.* 76: 4634-4642.
7. Papanikolopoulou, K., *et al.* (2004) Adenovirus fibre shaft sequences fold into the native triple beta-spiral fold when N-terminally fused to the bacteriophage T4 fibrin foldon trimerisation motif. *J. Mol. Biol.* 342: 219-227.
8. Güthe, S., *et al.* (2004) Very fast folding and association of a trimerization domain from bacteriophage T4 fibrin. *J. Mol. Biol.* 337: 905-915.
9. Milla, M. E. & Sauer, R. T. (1994) P22 arc repressor: Folding kinetics of a single-domain, dimeric protein. *Biochemistry* 33: 1125-1133.
10. Jackson, S. E. (1998) How do small single-domain proteins fold? *Folding & Design* 3: R81-R91.
11. Bachmann, A. & Kiefhaber, T. (2001) Apparent two-state tendamistat folding is a sequential process along a defined route. *J. Mol. Biol.* 306: 375-386.
12. 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.
13. Meier, S., Güthe, S., Kiefhaber, T. & Grzesiek, S. (2004) Foldon - the natural trimerization domain of T4 fibrin - dissociates into a monomeric A-state form containing a stable beta-hairpin. Atomic details of trimer dissociation and local beta-hairpin stability from residual dipolar couplings. *J. Mol. Biol.* 344: 1051-1069.
14. Wüthrich, K., Billeter, M. & Braun, W. (1984) Polypeptide secondary structure determination by nuclear magnetic resonance observation of short proton-proton distances. *J. Mol. Biol.* 180: 715-740.
15. Evans, P. A., Topping, K. D., Woolfson, D. N. & Dobson, C. M. (1991) Hydrophobic clustering in nonnative states of a protein: interpretation of chemical shifts in NMR spectra of denatured states of lysozyme. *Proteins* 9: 248-266.
16. Bundi, A., Wüthrich, K. (1979) Proton NMR parameters of the common amino acid residues in aqueous solutions of the linear tetrapeptides H-Gly-Gly-X-L-Ala-OH. *Biopolymers* 18: 285-296.
17. Perkins, S. J. & Wüthrich, K. (1979) Ring current effects in the conformation dependent NMR chemical shifts of aliphatic protons in the basic pancreatic trypsin inhibitor. *Biochim. Biophys. Acta* 576: 409-423.

18. Pardi, A., Wagner, G. & Wüthrich, K. (1983) Protein conformation and proton nuclear magnetic resonance chemical shifts. *Eur. J. Biochem.* 137: 445-454.
19. Baxter, N. J. & Williamson, M. P. (1997) Temperature dependence of ¹H chemical shifts in proteins. *J. Biomolec. NMR* 9: 359-369.
20. Goto, Y. & Nishikiori, S. (1991) Role of electrostatic repulsion in the acidic molten globule of cytochrome c. *J. Mol. Biol.* 222: 679-686.
21. Klein-Seetharaman, J., *et al.* (2002) Long-range interactions within a nonnative protein. *Science* 295: 1719-1722.
22. Peng, J. W. & Wagner, G. (1992) Mapping of the spectral densities of N-H bond motions in eglin c using heteronuclear relaxation experiments. *Biochemistry* 31: 8571-8586.
23. Schreiber, G. & Fersht, A. R. (1996) Rapid, electrostatically assisted association of proteins. *Nat. Struc. Biol.* 3: 427-431.
24. Wendt, H., Berger, C., Baici, A., Thomas, R. M. & Bosshard, H. R. (1995) Kinetics of folding of leucine zipper domains. *Biochemistry* 34: 4097-4107.
25. Bosshard, H. R. (2005) in *Protein Folding Handbook*, eds. Buchner, J. & Kiefhaber, T. (Wiley/VCH, Weinheim), Vol. 2, pp. 965-997.
26. Wagner, C. & Kiefhaber, T. (1999) Intermediates can accelerate protein folding. *Proc. Natl. Acad. Sci. USA* 96: 6716-6721.
27. Gill, S. C. & von Hippel, P. H. (1989) Calculation of protein extinction coefficients from amino acid sequence data. *Anal. Biochem.* 182: 319-326.
28. Koradi, R., Billeter, M. & Wüthrich, K. (1996) MOLMOL: a program for display and analysis of macromolecular structures. *J. Mol. Graphics* 14: 51-55.
29. Delaglio, F., *et al.* (1995) NMRPipe: a multidimensional spectral processing system based on UNIX pipes. *J. Biomol. NMR* 6: 277-293.
30. Goddard, T. D. & Kneller, D. G., University of California, San Francisco.
31. Schwarzing, S., *et al.* (2001) Sequence-dependent correction of random coil NMR chemical shifts. *J. Am. Chem. Soc.* 123: 2970-2978.

Figure Legends

Fig. 1: Structural comparison of the monomeric foldon E5R variant with the wild-type protein. The structure of the wild-type trimer (A; 1RFO) and of a single subunit in the native trimer (B) are compared to the ten lowest energy structures of the E5R monomer (C; ???). Trp20 is shown in green in panel A. The side chains of the amino acids forming the hydrophobic cluster between the hairpin and the N-terminal region are shown in red. (D): Overlay of the ten lowest energy structures in the E5R monomer showing all side chains. (E, F): Charge distribution on the different sides of a foldon monomer based on the E5R structure. Negative charges are shown in red, positive charges in blue.

Fig. 2: (A) Comparison of the fluorescence and (B) far UV CD spectra of foldon E5R (–) with the spectra of the monomeric transient folding intermediate (I, ●), the native trimer (–), the GdmCl-unfolded state (–, U) and the A-state (–, A) of the wild-type protein. The spectrum of the A-state was taken from Güthe *et al.* (8). (C) Temperature-induced equilibrium unfolding transitions of 5 μM (–) and 30 μM (–) foldon E5R and foldon wild-type at 5 μM (–) and 30 μM (–) monomer concentration monitored by the change in CD at 228 nm.

Fig. 3. Comparison of the ^{15}N -edited HSQC spectra of foldon E5R (red) and native foldon trimer (blue). The wild-type trimer resonances are indicated by asterisks (*). Spectra were taken in 10 mM potassium phosphate, pH 7.0 at 297 K.

Fig. 4: (A) Secondary amide proton chemical shifts $^1\text{H}^{\text{N}}$ of foldon E5R (●) and native foldon trimer (●) at 297 K, pH 7. The values correspond to the difference between the experimental shifts and typical random coil shifts of the protons, corrected for

neighboring residue effects (31). (B) Temperature coefficients for the backbone $^1\text{H}^{\text{N}}$ chemical shifts (●) in foldon E5R. Ala6 and Asp9 show non-linear temperature-dependencies with an average slope around 0 (see SI). (C) ^{15}N longitudinal, R_1 , and (D) transversal, R_2 , relaxation rates for the backbone amide groups in foldon E5R at 297 K, pH 7.

Fig. 5: Effect of temperature on the chemical shift for selected side chain protons of located in the N-terminal region of foldon E5R. Temperature coefficients for additional protons are shown in Fig. S4.

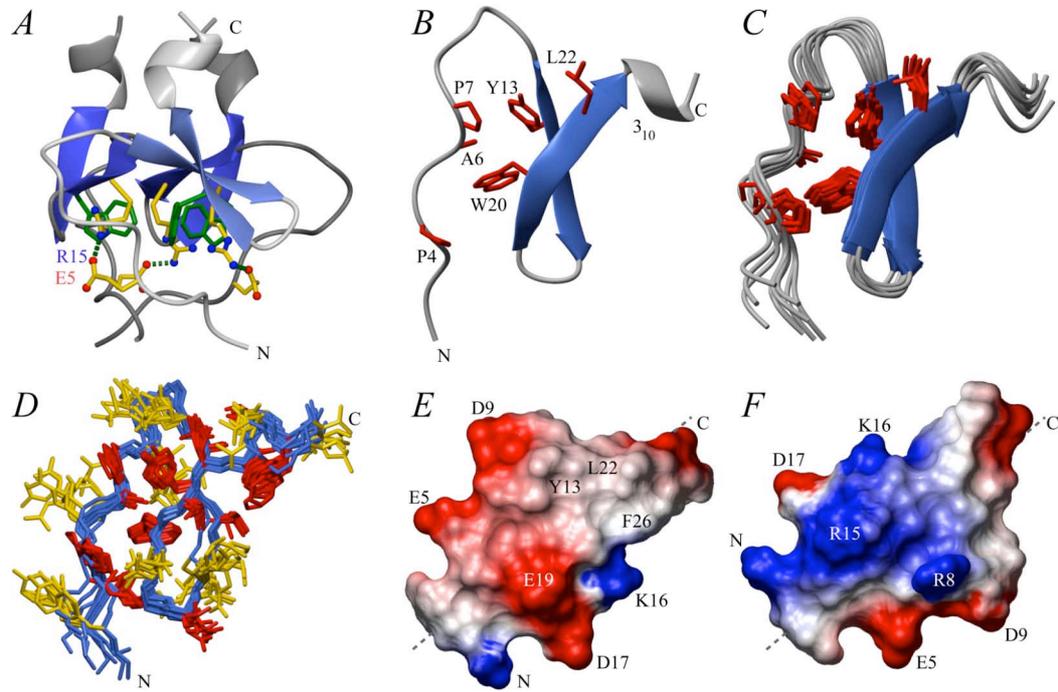


Figure 1

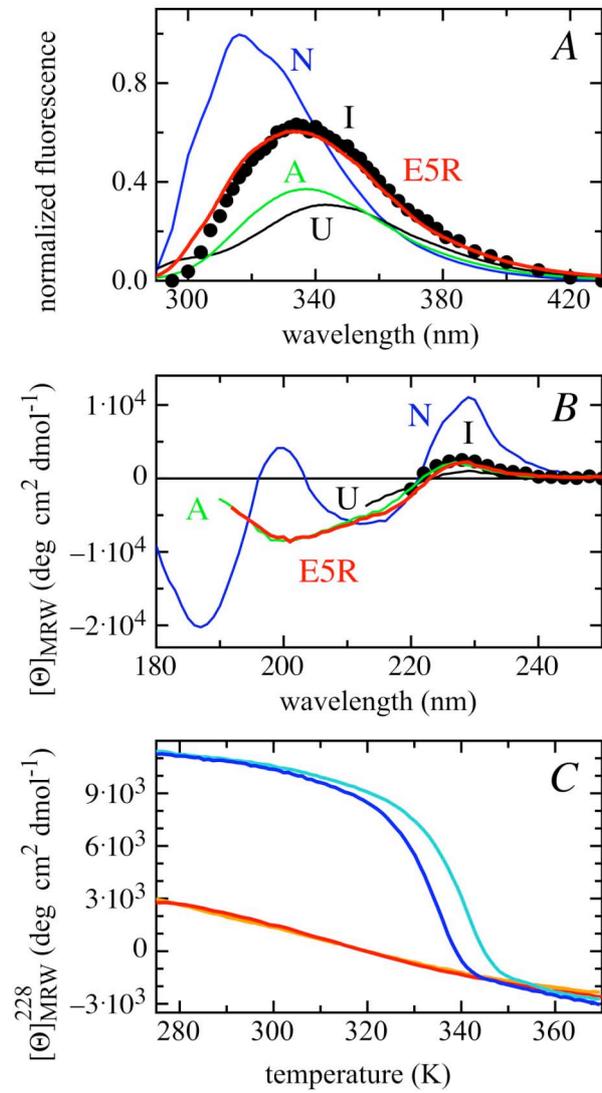


Figure 2

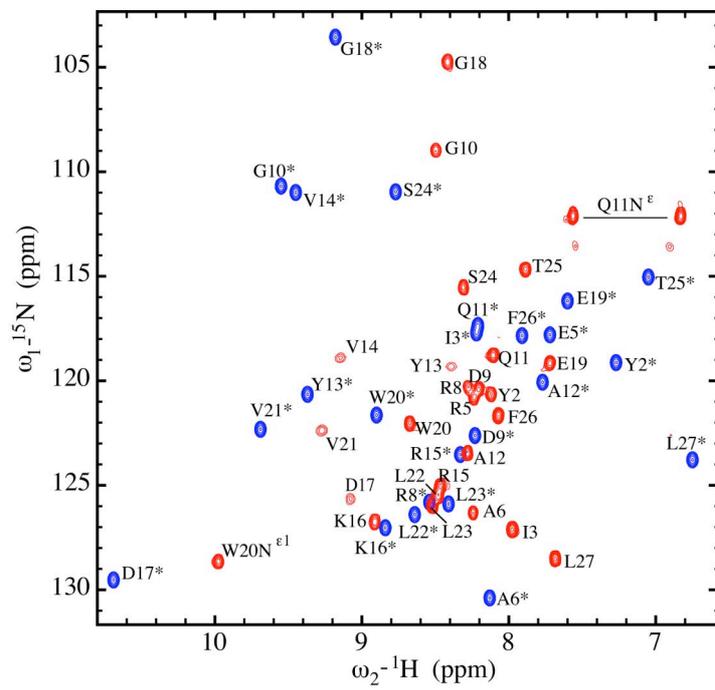


Figure 3

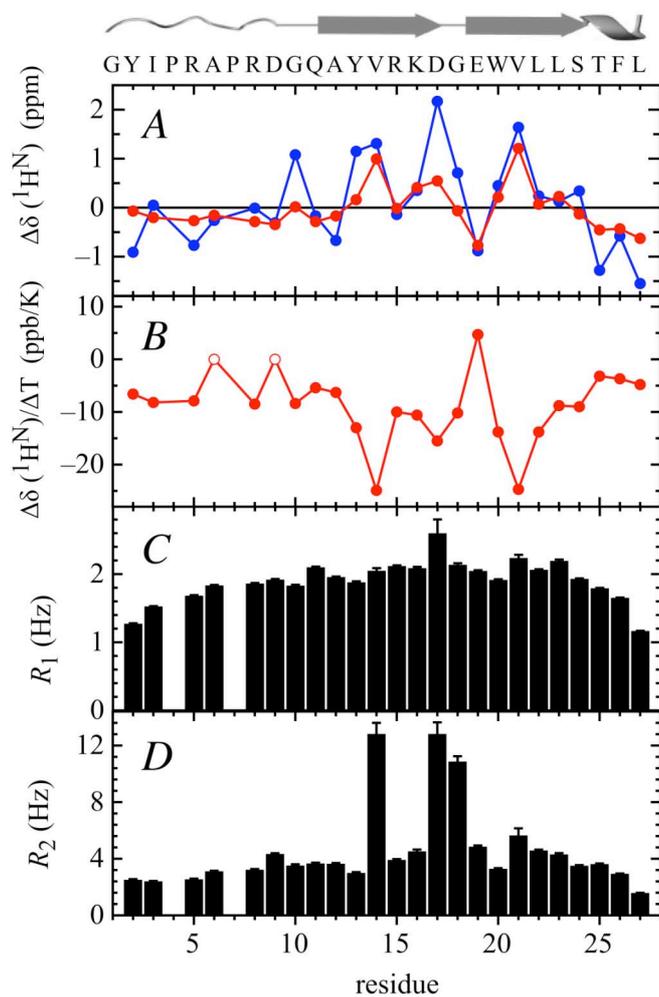


Figure 4

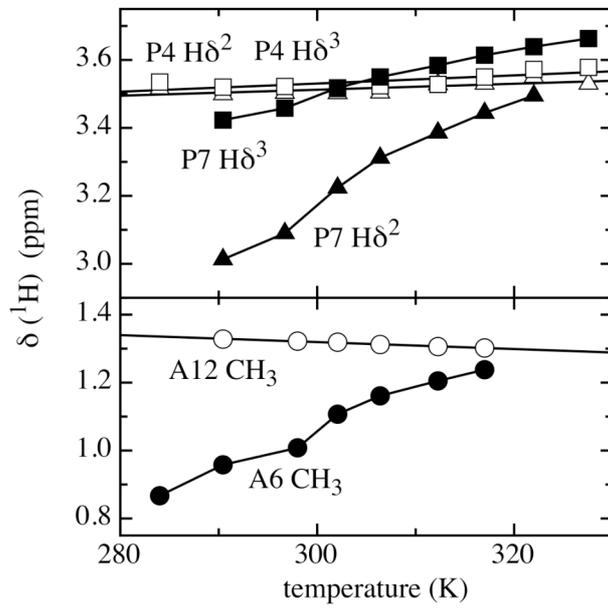


Figure 5

Supporting Information

containing a detailed description of the NMR methods and NMR data analysis, one table containing the statistics of the NMR structure determination and five “Supporting Figures” for the manuscript.

NMR experiments.

In homonuclear two-dimensional and ^{15}N -edited three-dimensional TOCSY experiments the MLEV17 sequence with gradients was used (1). In NOESY and HSQC pulse programs the WATERGATE pulse sequence for solvent signal suppression was incorporated (2). To determine NOE intensities for structure calculations, an additional two-dimensional NOESY spectrum was recorded with a mixing time of 300 ms on a Bruker DRX 800 MHz spectrometer equipped with a cryo-probe. Coupling constants $^3J_{\text{HN-HA}}$, determined from homonuclear TOCSY experiments, were used to establish constraints on ϕ dihedral angles. HSQC experiments at 13 temperatures in the range from 275 K to 337 K and 9 homonuclear TOCSY experiments from 284 K to 327 K were measured for the temperature dependence of the chemical shifts of the ^{15}N and $^1\text{H}^{\text{N}}$ resonances and the carbon bound ^1H resonances, respectively. The probe temperature was calibrated using 100 % methanol for below 305 K and 100 % ethylene glycol for higher temperatures (3). DSS was used as reference for ^1H chemical shift. The ^{15}N chemical shift was referenced relative to the ^2H lock resonance of water. Longitudinal (R_1) and transverse (R_2) ^{15}N relaxation rates were recorded interleaved at 296.7 K as described (4, 5). R_2 data were measured with Carr-Purcell-Maiboom-Gill (CPMG) spin-echo experiments using a delay of 0.5 ms between the 180° pulses. The contributions from chemical exchange to the transverse ^{15}N relaxation rates were investigated with different delays between the π pulses in the CPMG sequence, assuming two state exchange in the fast exchange limit (6).

Structure calculation.

NOE intensities were determined by integrating the cross-peaks in the 800 MHz two-dimensional NOESY spectrum and were converted to distance constraints by classifying them into strong ($d_{ij} \leq 2.7 \text{ \AA}$), medium ($d_{ij} \leq 3.5 \text{ \AA}$) and weak ($d_{ij} \leq 5 \text{ \AA}$). An uniform averaging model for NOE distance constraints involving peripheral amino acid side chain protons was applied to account for time variations of the proton-proton distances due to internal mobility of the protein (7). In this procedure an upper limit of

5 Å was attributed to all observable medium and long-range NOEs. As lower distance limit the van der Waals radius of 2 Å was applied. A more quantitative evaluation was not justified, considering the fact that the mixing time in the NOESY experiment was 300 ms and that some residues show line broadening. The common corrections were added to the upper distance limits when pseudo atoms were used to represent methyl, not stereospecifically assigned methyl or aromatic ring protons (8). For residues with two β -protons, the relative values of the $J_{\text{HA-HB}}$ coupling constants were estimated from a homonuclear TOCSY spectrum collected with a short mixing time (20 ms). Together with the HN-HB and HA-HB distances stereospecific assignments were made for the β protons of 9 residues. The methyl groups of Val21 were stereospecifically assigned during preliminary structure calculations applying random exchange of valine methyl groups.

The protein structures were calculated with the molecular dynamics program XPLOR (9) using a standard simulated annealing protocol. Distance constraints consisted of 80 intra-residue NOEs, 141 sequential NOEs, and 139 medium and long range NOEs. The dihedral constraints comprised 13 ϕ angles. A total of 40 structures were calculated starting from 20 different randomized structures. For each structure, a total of 6400 time steps were performed, 2400 at high “temperature” and 3000 at decreasing “temperature”, followed by 1000 steps of conjugated gradient minimization. From the initial structures NOE constraint violations were used to re-evaluate the assignments in the two dimensional NOESY spectrum. After this process four hydrogen bonds (Val14-Val21 and Lys16-Glu19) between the two hairpin strands were uniquely identified and introduced as additional distance constraints for a final set of calculations. The results of the simulated annealing calculations *in vacuo* show that the medium and long-range NOEs at the N- and C-terminal end are not sufficient to define a closely packed protein structure. Furthermore, many hydrophobic contacts along the trimer symmetry axis in the native wild type are on the surface and in contact with water in the E5R variant. The structures were submitted to further refinement in a thin layer of explicit water molecules using the TIP3P water model (10) to account for the most prominent effects of water, to provide polar and no polar interactions on the surface of the protein and to prevent unrealistic packing of the side-chains (11). The non-bonded interactions included Lennard-Jones, van der Waals and electrostatic interactions from the OPLS force field (12). Coulomb interactions were set to zero energy at a distance of 8.5 Å. The refinement was performed

by a simulated annealing protocol with a first heating stage from 100 K to 500 K, a refinement stage with 2500 steps of molecular dynamics and cooling stage from 500 K to 100 K. This procedure was followed by one hundred steps of conjugate gradient minimization. Ten structures were chosen with lowest target energy values and all NOE constraint violations less than 0.5 Å and dihedral angle violations less than 5°.

SI References

1. Bax, A. & Drobny, G. (1985) Optimization of two-dimensional homonuclear relayed coherence transfer NMR spectroscopy. *J. Mag. Res.* 61: 306-320.
2. Piotto, M., Saudek, V. & Sklenář, V. (1992) Gradient-tailored excitation for single-quantum NMR spectroscopy of aqueous solutions. *J. Biomol. NMR* 2: 661-666.
3. Raiford, D. S., Fisk, C. L. & Becker, E. D. (1979) Calibration of methanol and ethylene glycol nuclear magnetic resonance thermometers. *Anal. Chem.* 51: 2050-2051.
4. Grzesiek, S., *et al.* (1997) Refined solution structure and backbone dynamics of HIV-1 Nef. *Protein Sci.* 6: 1248-1263.
5. Holland, D. R., *et al.* (1992) Structural comparison suggests that thermolysin and related neutral proteases undergo hinge-bending motion during catalysis. *Biochemistry* 31: 11310-11316.
6. Ishima, R. & Torchia, D. A. (1999) Estimating the time scale of chemical exchange of proteins from measurements of transverse relaxation rates in solution. *J. Biomol. NMR* 14: 369-372.
7. Güntert, P., Braun, W. & Wüthrich, K. (1991) Efficient computation of 3D protein structures. *J. Mol. Biol.* 217: 517-530.
8. Wüthrich, K., Billeter, M. & Braun, W. (1983) Pseudo-structures for the 20 common amino acids for use in studies of protein conformations by measurements of intramolecular proton-proton distance constraints with nuclear magnetic resonance. *J. Mol. Biol.* 169: 949-961.
9. Schwieters, C. D., Kuszewski, J. J., Tjandra, N. & Clore, G. M. (2003) The Xplor-NIH NMR molecular structure determination package. *J. Magn. Res.* 160: 66-74.
10. Linge, J. P., Williams, M. A., Spronk, C. A., Bonvin, A. M. & Nilges, M. (2003) Refinement of protein structures in explicit solvent. *Proteins* 50: 496-506.
11. Beglov, D. & Roux, B. (1995) Dominant solvation effects from the primary shell of hydration: approximation for molecular dynamics simulations. *Biopolymers* 35: 171-178.
12. Jorgensen, W. L. & Tirado-Rives, J. (1988) The OPLS [optimized potentials for liquid simulations] potential functions for proteins, energy minimizations for crystals of cyclic peptides and crambin. *J. Am. Chem. Soc.* 110: 1657-1666.
13. Güthe, S., *et al.* (2004) Very fast folding and association of a trimerization domain from bacteriophage T4 fibritin. *J. Mol. Biol.* 337: 905-915.

Table S1. Statistics of the foldon E5R NMR structure^a

RMSDs from experimental distance constraints (Å) of all 360 NOEs	0.062 ± 0.002
RMSDs from experimental dihedral constraints (°) of all 15 angles	0.95 ± 0.19
Deviation from idealized covalent geometry	
bonds (Å)	0.018 ± 0.0001
angles (°)	0.68 ± 0.02
impropers (°)	0.77 ± 0.03
Coordinate precision of the whole protein ^a (Å)	
pairwise RMSD of all heavy backbone atoms	1.08
pairwise RMSD of all heavy atoms	1.77
Coordinate precision of the β-sheet ^a (residues 12 to 23) (Å)	
pairwise RMSD of all heavy backbone atoms	0.51
pairwise RMSD of all heavy atoms	1.34
Coordinate precision of the N-terminal region (residues 1 to 11) (Å)	
pairwise RMSD of all heavy backbone atoms ^a	1.91
pairwise RMSD of all heavy atoms ^a	2.45
pairwise RMSD of all heavy backbone atoms ^b	0.85
pairwise RMSD of all heavy atoms ^b	1.81
Coordinate precision of the C-terminal end (residues 24 to 27) (Å)	
pairwise RMSD of all heavy backbone atoms ^a	1.72
pairwise RMSD of all heavy atoms ^a	2.18
pairwise RMSD of all heavy backbone atoms ^c	0.28
pairwise RMSD of all heavy atoms ^c	1.18
^a All heavy backbone atoms of the hairpin (Ala12-Leu23) were used in the fit.	
^b All heavy backbone atoms of the N-terminal region (Gly1-Gln11) were fitted.	
^c All heavy backbone atoms of the C-terminal region (Ser24-Leu27) were fitted.	

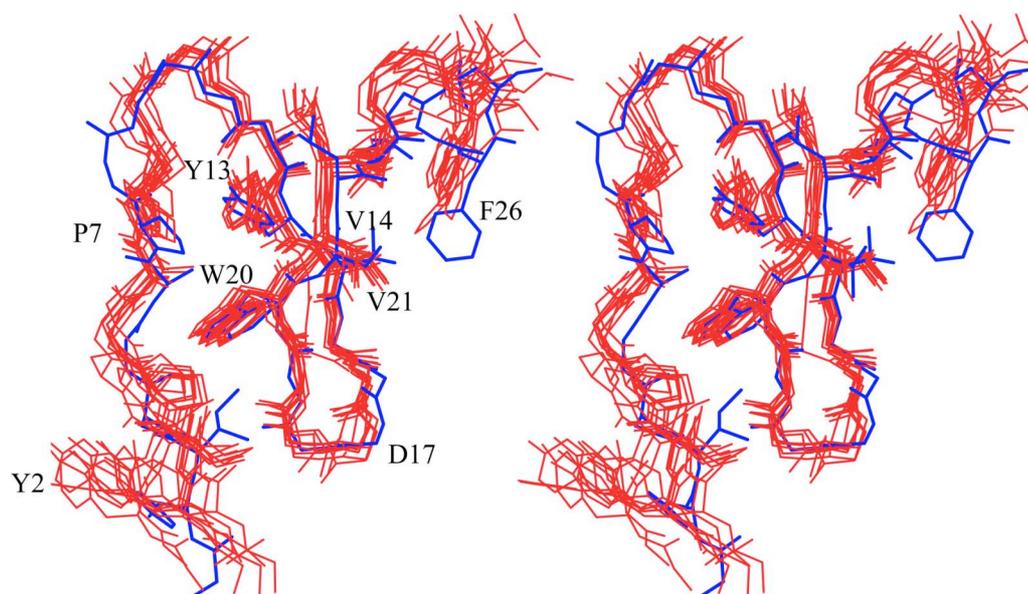


Fig. S1: Stereo view of the ten foldon E5R structures with the lowest energy (red) and a single monomer structure from the foldon trimer (blue) (13). The backbone and the heavy side chain atoms of Tyr2, Ile3, Pro4, Ala6, Pro7, Ala12, Tyr13, Val14, Trp20, Val21, Leu22, Leu23, Phe26 and Leu27 are shown.

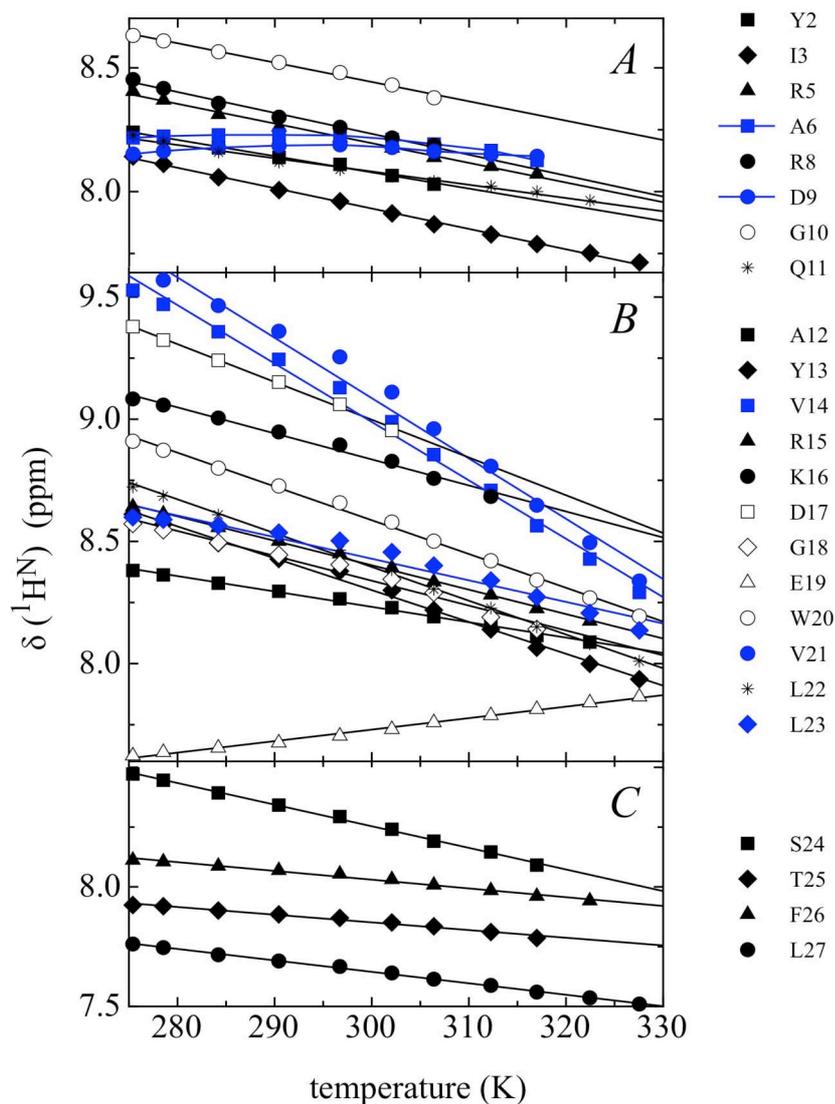


Fig. S2: Temperature dependence of the amide proton resonances in (A) the N-terminal, (B) the central β -hairpin and (C) the C-terminal part of foldon E5R. The results of the linear fits shown as lines are given in Fig. 4A. Resonances displaying non-linear behavior are colored in blue. The measurements were performed in 10 mM potassium phosphate, pH 7.0.

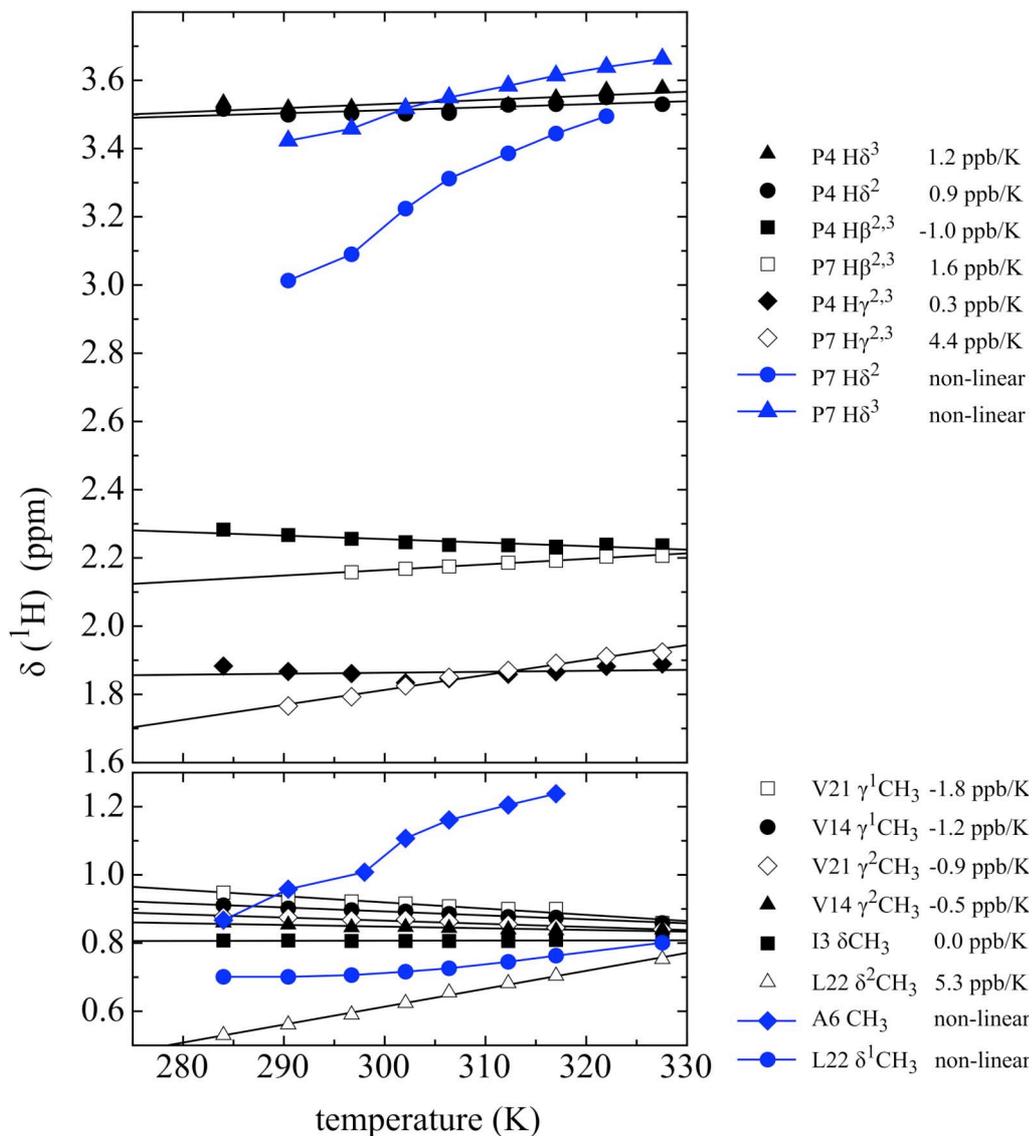


Fig. S3: Temperature dependence of selected side chain proton resonances. The results of the linear fits are shown as lines. Resonances displaying non-linear behavior are colored in blue. The measurements were performed in 10 mM potassium phosphate, pH 7.0.

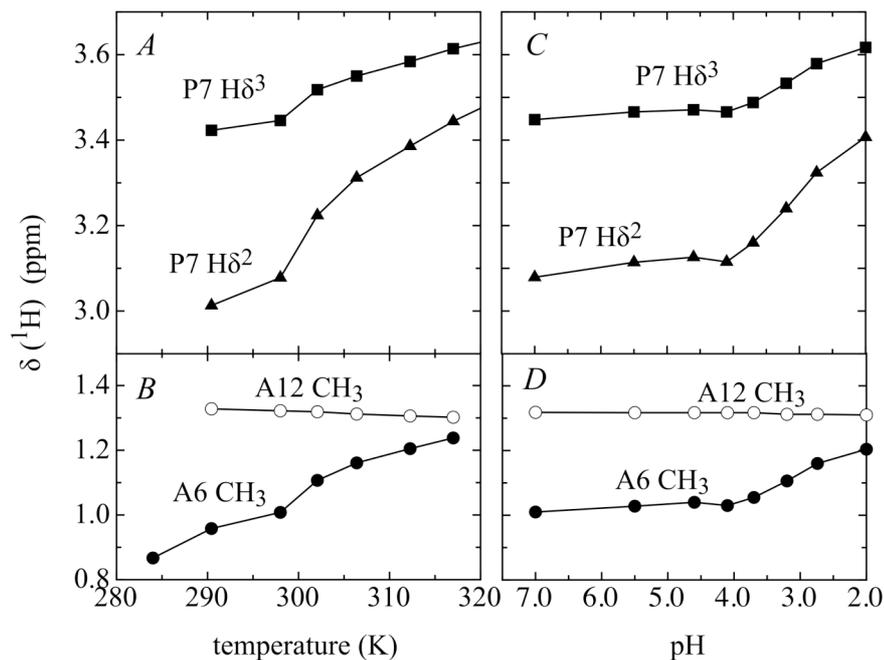


Fig. S4: Comparison of (A,B) the temperature dependencies at pH 7 and (C,D) the pH dependencies of the proton resonances of Ala6, Pro7 and Ala12 at 297 K. The hydrophobic interaction between the N-terminal strand and the central β -hairpin, reported by the shift deviations of P7 and A6 protons, is lost: their chemical shifts move toward typical random coil values, both with increasing temperature and with decreasing pH.

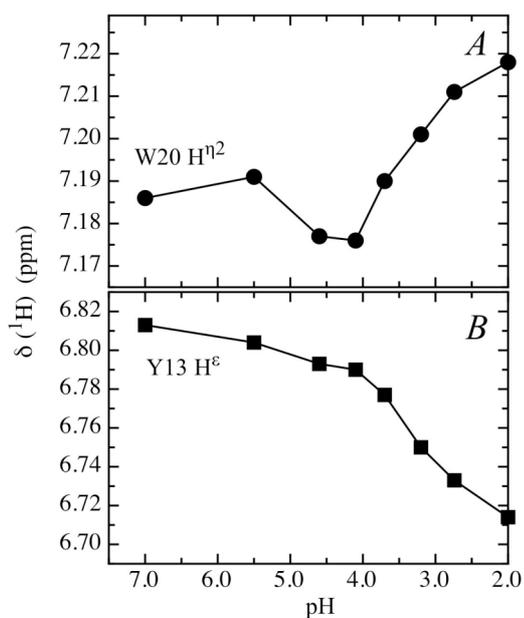


Fig. S5: pH-Dependencies of the side chain proton resonances of Tyr13 and Trp20.

**Conformational Dynamics in the Native and Denatured State
of the Villin Headpiece Subdomain
Measured by Triplet-Triplet Energy Transfer**

Andreas Reiner¹, Peter Henklein² and Thomas Kiefhaber^{3*}

¹Biozentrum der Universität Basel, Division of Biophysical Chemistry,
Klingelbergstrasse 70, CH-4056 Basel, Switzerland

²Universitätsmedizin Berlin, Institut für Biochemie, Humboldt Universität Berlin
Monbijoustrasse 2, 10117 Berlin.

³Munich Center for Integrated Protein Science & Chemistry Department, Technische
Universität München, Lichtenbergstrasse 4, D-85747 Garching, Germany

*corresponding author. Phone: ++49 89 289 13420; fax: ++49 89 289 13416;

e-mail: t.kiefhaber@tum.de

Abstract

Conformational dynamics are an inherent feature of folded proteins and are the prerequisite for the conformational search within the unfolded state ensemble. We investigated dynamics in the native and denatured states of the villin headpiece subdomain using triplet-triplet energy transfer. This technique allows the measurement of absolute rate constants for van der Waals contact formation between a triplet donor and a triplet acceptor group. Labelling villin headpiece with donor and acceptor groups at different positions enabled us to study the dynamics of loop formation in the unfolded state and to probe for local unfolding events in the native protein. The experiments revealed efficient contact formation in the native state if the labels are attached at the opposite ends of the C-terminal α -helix, indicating local unfolding of this helix on the microsecond time scale. The results further point to conformational heterogeneity at the C-terminus giving rise to fast and slow electron transfer reactions between labels attached at the N- and C-terminus, which are in close proximity in the native state. Single exponential kinetics for loop formation on the timescale of hundreds of nanoseconds were observed for all donor acceptor positions in the unfolded state. Comparison with data on unstructured model sequences suggests that the chain dynamics in the denatured state are not affected by residual structure, even under conditions strongly favouring the native state. The dynamics of loop formation can be compared to molecular dynamic simulations and are important in context of the fast folding reaction of the villin headpiece subdomain.

Introduction

Conformational dynamics play an essential role, both for protein folding and protein function. Conformational fluctuations occur over a wide range of timescales and include all levels of structural organization. In the native state of proteins flexibility is usually necessary to facilitate processes such as catalysis, binding or signalling. In the unfolded state proteins sample a vast number of conformations in short times, eventually leading to a folding event. Although the conformational search is closely related to the folding process itself, little is known about the dynamics in the unfolded state ensemble of proteins, since many experimental techniques fail to give an accurate experimental description of their heterogeneity and dynamic nature.

The properties of unfolded state ensembles are controversially discussed, as well as the actual consequences for folding. In some proteins, residual structure is found [1-4], which might serve as nucleation site for folding. Even in the absence of any specific interactions local conformational preferences exist as a result of steric factors [5], polymer properties and hydration, which are not well understood [6]. To assess the role of the unfolded state for the folding process, it is important to know which conformations are populated to which extent and how fast they interconvert. Despite recent progress, mainly in NMR techniques, several timescales and types of motions remain unexplored. Of particular interest for folding are the chain dynamics, large amplitude motions of the polypeptide chain occurring on the nanosecond timescale. These dynamics determine, how fast two positions within a polypeptide can form contact, i.e. how fast single interactions can be established.

An experimental method well suited to measure chain dynamics is triplet-triplet energy transfer (TTET) [7-9] from xanthone (Xan) to naphthylalanine (Nal) which reports on

contact formation between the two groups attached to specific sites on a polypeptide chain. The triplet donor Xan is excited to its triplet state by a short laser pulse. Upon van der Waals contact with the triplet acceptor Nal, the triplet state is exothermically transferred within 1 ps [10]. The high efficiency of this Dexter exchange ensures that even very short encounters will result in TTET, making the system completely diffusion controlled. Therefore absolute rate constants of contact formation can be measured directly, monitoring the characteristic triplet absorption bands of the donor or acceptor. TTET provided detailed insight into the dynamics of unstructured model peptides ranging from picoseconds to microseconds [8,11]. The influence of peptide length, composition [7,8], end-extensions [12], the role of proline [13] and the effects of denaturants were investigated in detail [8,9,14].

In this study, we introduced the donor and acceptor groups at different positions in the chicken villin headpiece subdomain (HP35) to investigate contact formation in the native and denatured state. This three-helix bundle consists of 35 amino acids and is the smallest known naturally occurring protein unit which folds cooperatively without stabilization by disulfide bonds or cofactors [15]. It forms the last part of the C-terminal headpiece domain of villin, where it is involved in actin binding [16]. High resolution structures determined by NMR [17] and X-ray crystallography [18], show a well-packed hydrophobic core mainly formed by three phenylalanines. *T*-jump experiments [18-22] and NMR line-shape analysis [23] suggested that folding is fast on the timescale of microseconds. This was also confirmed by measuring unfolding followed by loop-formation [24]. Recently a variant of HP35 has been reported, that even folds on the sub-microsecond timescale [20]. Ultrafast folding of HP35 was suggested to be a

consequence of residual structure in the unfolded state, which was detected in a fragment of HP35 comprising the first two helices [25,26].

The small size and fast folding of HP35 attracts not only experimentalists, but made it to a popular subject of simulations, as it may allow to close the gap between experimental and computational approaches. Numerous MD simulations, including a 1 μ s all-atom simulation of Kollman in 1998 [27] and the distributed computing project of Pande [28], addressed both the properties of the native and denatured state ensembles, as well as the course of folding. Only recently Duan and colleagues [29] achieved complete folding to the native state and extracted a free-energy surface, using replica exchange MD.

The chain dynamics in the unfolded state of an ultrafast folding protein, such as HP35, are of particular interest, since the timescales for contact formation, representing the early stages of folding, and for the completion of the folding reaction differ only by a factor of thousand. If the folding process would be facilitated by residual structure, this would most likely also affect loop formation dynamics in the unfolded state. TTET provides absolute rate constants of contact formation, which should be also of great use for direct comparison with MD simulations. In addition, contact formation in the native state is suitable to probe local dynamics or partial unfolding events on the nanosecond to microsecond timescale [30], which are difficult to access with other techniques.

We measured TTET between four different donor-acceptor positions in HP35 to address these aspects of protein dynamics. We detected significant flexibility at the C-terminus and characterised the chain dynamics in the presence of high concentrations of denaturant, as well as under conditions strongly favouring the native state.

Results

Positions of the Triplet Donor and Acceptor Groups.

We chose four different donor–acceptor positions in HP35 to measure contact formation (Fig. 1). The double-labelled variants of HP35, which had the general sequence H-¹LSDEDFKAVF¹¹GNleTRSAFANL²¹PLFKQQNLKK³¹EKGLF-OH, were prepared by solid-phase peptide synthesis. The naturally occurring Met12 was replaced by norleucine (Nle), since methionine is prone to oxidation during peptide synthesis. In addition, methionine is a quencher of the Xan triplet state and would interfere with TTET [31]. For the same reason Trp23 was substituted by phenylalanine, unless the site was not used to incorporate the triplet acceptor. The triplet donor Xan was attached as carboxylic acid derivative to either the N-terminal amino group of Leu1 (Xan0), the ϵ -amino group of Lys7 (Xan7) or the β -amino group of α,β -diaminopropionic acid replacing Phe35 (Xan35). The non-natural amino acid Nal was used as triplet acceptor substituting either Phe23 (Nal23) or Phe35 (Nal35). For all Xan positions donor-only variants of HP35 were synthesized as references to determine the donor lifetimes in the absence of acceptor.

Figure 1

The donor-acceptor positions were chosen in order to probe different features of the native and denatured state. In the variant Xan0-Nal23 donor and acceptor are placed far apart from each other in the folded state (Fig. 1A). They are separated by helix 1 and helix 2, a region that retains helical structure even in the absence of tertiary contacts with helix 3 [25,26]. The same positions were used by Buscaglia *et al.* in a tryptophan triplet quenching experiment [24]. Moving Xan to the side chain of Lys7 (Xan7-Nal23) places the labels at the ends of helix 2 (Fig. 1B), which is likely sufficient to separate the labels

from each other and to investigate the local dynamics of this region. In the Xan0-Nal35 variant the labels are in close proximity in the native state (Fig. 1C), which should result in efficient contact formation. This should lead to fast transfer in the native state and will give information on the dynamics of formation of this native contact in the unfolded state. The variant Nal23-Xan35 was synthesized to probe the dynamics of helix 3 and the conserved C-terminal GLF-motif (Fig. 1D). Structural data indicate that helix 3 is well defined but slightly different orientations relative to the other helices were reported [17,18]. The three C-terminal amino acid residues, however, seem to remain flexible in the context of HP35 [17,32]. Contact formation experiments are likely to provide data on the extent and the timescale of these local dynamics.

Structure and Stability of the HP35 Variants.

All donor-acceptor and donor-only labelled HP35 variants show α -helical Far-UV CD spectra (Fig. S1, Supporting Information), with signal intensities comparable to the wild-type protein reported by other groups [15]. One-dimensional NMR spectra of the variants are similar to those reported for the wild-type subdomain with some well resolved resonances (Fig. 2A and Fig. S2). Especially the upfield shift of the Val9 methyl group, and the downfield shift of the amide protons of Ser2 and Asp3 are indicative for a folded structure. This shows that the overall structure is not affected by introducing the TTET labels.

Figure 2

To evaluate the thermodynamic stability of the different donor-acceptor variants we recorded GdmCl-induced unfolding transitions at 5.0 °C by measuring the change in

ellipticity at 222 nm (Fig. 2B, Fig. S1). Table 1 summarizes the data obtained from fitting the equilibrium transitions with a two state-model (see Material and Methods). The free energies for unfolding, $\Delta G_D^{5^\circ\text{C}}$, are in the range of 6 - 9 kJ mol⁻¹. The m_D values are around 3 kJ mol⁻¹ M⁻¹, which is identical to reported values [24] and in agreement with the value expected by considering the change in accessible surface area [33]. The ΔG_D values are lower than those reported for wild-type HP35, which is likely a consequence of the Met12 replacement with Nle [34]. The corresponding midpoints of the unfolding transitions vary between 2.1 M and 2.6 M GdmCl (Tab. 1).

Table 1

TTET in Xan0-Nal23 and Xan7-Nal23.

In the TTET measurements a short laser pulse is used to generate the Xan triplet state, which can be monitored via its absorption band at 590 nm. Under the experimental conditions the Xan triplet state has an intrinsic lifetime of about 10 μs , as observed in the donor-only variant Xan0 (Fig. 3A), as well as in the donor-only variants Xan7 (Fig. S3) and Xan35 (Fig. 5A). The intrinsic triplet lifetime is independent of the protein environment. It remains unchanged during the GdmCl-induced unfolding transitions of the HP35 variants (data not shown). The Xan triplet state can only be transferred to the acceptor Nal when the labels form van der Waals contact. In the native state of the Xan0-Nal23 variant this requires at least partial unfolding of the protein (Fig. 1A). In the presence of 1 M GdmCl, where HP35 is mainly folded, the Xan triplet decay resembles the decay observed in the donor-only reference (Fig. 3A) arguing against local unfolding on the ns to μs time scale in the folded state of HP35. At higher GdmCl concentrations the unfolded state of HP35 becomes significantly populated and the Xan triplet states

decay faster (Fig. 3A). Around the transition midpoint of 2.2 M GdmCl the decays show a pronounced double exponential behaviour with equal contributions of the two phases. At high GdmCl concentrations, where HP35 is mainly unfolded, the fast process on the 1 μ s timescale dominates. This fast process can be attributed to TTET, which is confirmed by a subsequent increase of the NaI triplet state absorption band at 420 nm (data not shown). Over the whole range of GdmCl concentrations the decay curves can be well described by the sum of two exponential functions. The obtained rate constants and their amplitudes are shown in Fig. 3B and Fig. 3C, respectively.

Figure 3

The amplitude of the fast phase representing TTET corresponds well to the fraction of unfolded molecules determined by CD (Fig. 3C), indicating two-state behaviour. This means that in the fraction of folded molecules no contact formation occurs within the Xan triplet lifetime of 10 μ s, whereas in the denatured state ensemble all molecules form contact equally fast. A small fraction (~10 %) of slow phase remains in all TTET experiments, which may result from molecules with bleached acceptor or from molecules which cannot form contact, e.g. in small aggregates. This behaviour has also been observed for other peptides [8].

High concentrations of GdmCl slow down the contact formation in the denatured state ensemble (Fig. 3B). The logarithm of the rate constant of contact formation, $\ln k_c$, depends linearly on the denaturant concentration [D] (Eq. 1).

$$\ln k_c = \ln k_c^0 - \frac{m_c [D]}{RT} \quad (1)$$

The slope m_c reflects the sensitivity of the contact formation processes towards denaturant. A linear reduction of $\ln k_c$ has also been observed in earlier studies on unstructured model sequences [13,14], where it could be shown to result from additive effects of solvent viscosity and denaturant binding to the polypeptide chain. Remarkably, at low concentrations of GdmCl, where the folded state is favoured, no deviation from this behaviour is observed. Extrapolating k_c to 0 M GdmCl gives $k_c^0 = 2.7 \cdot 10^6 \text{ s}^{-1}$ at 5 °C ($\tau \sim 370 \text{ ns}$; Tab. 1) for formation of this 22 amino acid loop in the unfolded state in the absence of denaturant.

TTET measurements with the HP35 variant Xan7-Nal23 reveal a similar behaviour as in Xan0-Nal23. For the native state population the triplet decay corresponds to the Xan lifetime (Fig. S3), indicating the absence of local unfolding of helix 2 on the ns to μs time scale. In the GdmCl-denatured state TTET occurs, leading to double exponential kinetics in the transition region (Fig. S3). The logarithm of the contact formation rate constant depends again linearly on the GdmCl concentration. With $k_c^0 = 4 \cdot 10^6 \text{ s}^{-1}$ (Tab. 1) this process is faster than for Xan0-Nal23, which is expected, since the labels are separated by only 13 amino acids.

TTET in Xan0-Nal35.

In the variant Xan0-Nal35 the labels are located at the termini of HP35, which should allow rapid TTET in the native state (Fig. 1C). The results show that part of the Xan triplets are transferred to Nal within the 12 ns dead-time of our nanosecond laserflash experiment, leading to a dead time amplitude change of about 45 % compared to the donor-only reference protein. The remaining Xan triplets are transferred with a rate constant of $\sim 3 \cdot 10^7 \text{ s}^{-1}$ (Fig. 4B), which is ten times faster than loop formation in the denatured state of this variant. This process becomes slower as the GdmCl concentration is increased, until it disappears together with the dead-time amplitude when the protein is completely unfolded (Fig. 4C). The chain dynamics in the denatured state are similar to those observed for Xan0-Nal23 and Xan7-Nal23.

Figure 4

TTET in Nal23-Xan35.

In the HP35 variant Nal23-Xan35 the labels are placed to probe the dynamics of the C-terminal helix. The model based on the X-ray structure (Fig. 1D) suggests that the 2.5 helical turns should avoid direct contact formation between positions 23 and 35. TTET kinetics in this variant are complex (Fig. 5). Two kinetic phases for TTET are observed at low denaturant concentrations indicating that in the native state contact formation is double exponential (Fig. 5B,C). In the absence of GdmCl the major phase (80 % amplitude) has a time constant of $\sim 1 \mu\text{s}$, which remains virtually unchanged by the addition of GdmCl. The second, faster phase ($\tau \sim 170 \text{ ns}$) becomes strongly accelerated with increasing concentrations of denaturant. At higher denaturant concentrations the

unfolded state becomes populated and a third phase appears, representing contact formation in the denatured state ensemble.

Figure 5

Discussion

The TTET data from the four donor-acceptor labelled HP35 variants agree with two-state behaviour, i.e. all contact formation processes can be attributed to either the native or the denatured state ensemble indicating the absence of partially folded states with altered dynamics. In the following, contact formation in the native state and denatured state ensemble is discussed independently from each other.

Dynamics in the Native State.

In the native state of the variants Xan0-Nal23 and Xan7-Nal23 the donor triplet state decays with its intrinsic triplet lifetime. This indicates, that within 10 μs no local or global unfolding occurs that leads to contact formation between these positions. *T*-jump data [19,24] on other HP35 variants gave unfolding time constants of more than 400 μs in the absence of GdmCl and of about 30 μs in the presence of 4 M GdmCl at 20 °C. Using the longer lived triplet state of tryptophan ($\sim 100 \mu\text{s}$) in a cysteine quenching experiment, Buscaglia *et al.* were able to extract contributions arising from global unfolding [24].

In the variant Xan0-Nal35 the labels are close to each other in the folded state and form contact much faster than in the denatured state. Contact formation occurs on two well-separated timescales, pointing to heterogeneity within the native state ensemble. Accordingly, the interconversion between the substates has to be slower or on a similar

time scale as TTET. In about 50 % of native molecules contact is formed in a sub-nanosecond process, as inferred from TTET during the dead-time of the experiment. In these molecules van der Waals contact might be established at the time point of excitation or can be accomplished by a few bond rotations. The second TTET process in the native state has a time constant of about 30 ns. In these molecules the encounter is less probable, either due to larger conformational freedom of the labels or due to an energy barrier encountered by contact formation. The process is slowed down by the addition of GdmCl ($m = (750 \pm 140) \text{ J mol}^{-1} \text{ M}^{-1}$) similar to contact formation in unfolded peptide chains ($m_c \sim 500 \text{ J mol}^{-1} \text{ M}^{-1}$, [14]). Therefore, this reaction seems to be associated with only little change in the accessible surface area. The observed heterogeneity may be attributed to different orientations of the labels or to local flexibility at the termini of HP35. Indeed, NMR experiments have shown that Leu1 and the last two amino acids, Leu34 and Phe35, are not well ordered. This was deduced from a lack of NOE constraints [17], low protection factors [15] and low order parameters [32].

Additional evidence for local flexibility in the C-terminal region of HP35 is provided by the variant Nal23-Xan35. In the native state efficient contact formation is observed, although the segment between the labels compromises helix 3 with 2.5 turns and the C-terminus. TTET in the native state is described by two exponential functions, with the main process occurring on the timescale of $\sim 1 \mu\text{s}$. While this is somewhat slower than contact formation in the denatured state, the second phase is considerably faster with a time constant of 170 ns. It makes up $\sim 20\%$ of the amplitude and is strongly accelerated by the addition of GdmCl up to a time constant of 30 ns indicating that it reflects an unfolding reaction. As already discussed, the C-terminal two amino acids, Leu34 and Phe35, seem to be flexible, although they are highly conserved and important for actin

binding [16]. Together with the inherent flexibility of Gly33 the last segment might be highly mobile, consistent with the different conformations reported [17,18]. This local flexibility at the C-terminus might be sufficient to allow eventual contact between Xan35 and Nal23. However, this does not explain the strong denaturant dependence of the process. It is more likely that the fast process originates from local unfolding of helix 3, in agreement with the acceleration of this process by addition of GdmCl. Helices in isolation show end-fraying on the hundreds of nanoseconds timescale, which results in double exponential kinetics [30].

Interestingly, in all *T*-jump experiments on HP35 a fast process on the 100 ns timescale was observed as well. This process was detected with different probes, such as fluorescence quenching of Trp23 by the engineered His27 in the middle of helix 3 [19], by the bulk IR signal [21] and by a site-specific IR measurement with an isotope-labelled carbonyl in helix 2 [22]. In all cases the fast phase was attributed to helix-coil transitions, which might originate from local unfolding events.

Dynamics in the Denatured State.

While the native state structure of HP35 is well known and the folding process has been investigated in a number of *T*-jump experiments, less is known about the denatured state ensemble. An accurate description of this heterogeneous and quickly converting ensemble is especially desirable in the case of HP35 for several reasons. Firstly, it makes up the bulk of data in molecular dynamic simulations, which are difficult to validate with experimental data. Secondly, HP35 is an ultrafast folding protein with folding time constants up to 700 ns, observed in a designed variant [20]. This raised the question,

whether this might be the consequence of residual structure in the denatured state. We measured contact formation in the GdmCl-denatured state of HP35 at 5.0 °C (Fig. 3-5) and at 22.5 °C (Fig. 6A). In all cases contact formation in the unfolded state is described by a single exponential function. This points to a rapid equilibration of chain conformations in the denatured state ensemble [35], much faster than the observed contact formation processes on the 100 ns timescale. The logarithms of the rate constants for contact formation, $\ln k_c$, depend linearly on the GdmCl concentration (Eq. 1). Even under conditions strongly favouring the native state ($\sim 80\%$ in 1 M GdmCl at 5.0 °C) no evidence for deviations from this behaviour was found. The GdmCl dependencies give m_c values of $\sim 500 \text{ J mol}^{-1} \text{ M}^{-1}$ (Tab. 1), which are identical to those obtained for unstructured model peptides [14]. Hence there is no evidence for specific residual structure present at low GdmCl concentrations. Also the rate constants can be directly compared to those obtained in model peptides by extrapolating to 0 M denaturant (k_c^0 , Eq. 1 and Tab. 1). End-to-end contact formation in poly(GlySer) and poly(Ser) depends on the number of peptide bonds between the labels N , scaling with $k_c \sim N^{-1.7}$ and $k_c \sim N^{-2.1}$, respectively [8]. End-to-end contact formation in HP35, probed with the variant Xan0-Nal35, is exactly as fast as expected from the data on poly(Ser) peptides (Fig. 6B). In the other three variants the labels are located in the interior of the chain. For model peptides it was found that N- and C-terminal extensions reduce k_c up to factor of 2.5 [12]. When this effect is taken into account, the rate constants of contact formation between all donor-acceptor positions are comparable to those observed for unstructured poly(Ser) chains (Fig. 6B).

Figure 6

Contact formation data for HP35 were also reported by Buscaglia *et al.*, who measured the quenching of the tryptophan 23 triplet state by a cysteine added at position 0 [24]. The reported rate constants of triplet quenching are about 20-fold slower than the rate constants for contact formation we obtain for the identical positions using TTET (Xan0-Nal23). This can be attributed to the fact, that triplet quenching by cysteine is not fast enough to occur upon every encounter [36,37], i.e. the system is not diffusion controlled. In contrast, TTET gives absolute rate constants for contact formation, since the triplet state is transferred within ~ 1 ps after van der Waals contact has been established [10]. The absolute rate constants for contact formation reported in Table 1 should be of great use for direct comparison with molecular dynamics simulations.

For all variants temperature dependencies of the contact formation rates were measured in 4 M GdmCl, which gave linear Arrhenius plots with activation energies of about 28 kJ mol^{-1} (Tab. 1). These are apparent activation energies, since the temperature dependence of the solvent viscosity is not taken into account, which would amount to $\sim 15 \text{ kJ mol}^{-1}$ [38] in the case of fully viscosity dependent reactions. Similar activation energies were observed for contact formation in unstructured peptides [13,39].

Contact formation in the HP35 variant Xan0-Nal35 was also investigated as a function of the urea concentration (data not shown). In contrast to GdmCl, urea is non-ionic and seems to be less efficient in disrupting hydrophobic clusters [3]. Also in the presence of urea, contact formation in the unfolded state is described by a single exponential function and $\ln k_c$ depends linearly on the urea concentration. The rate constants of contact formation extrapolate to the same k_c^0 as found with GdmCl (Tab. 1). The m_c value for urea is $190 \text{ J mol}^{-1} \text{ M}^{-1}$. Similar values were found for urea interacting with hydrophobic

host-guest peptides [40]. Increasing the ionic strength by adding up to 500 mM NaCl did not affect contact formation in 8 M urea (data not shown). This indicates that the chain dynamics in the urea-denatured state are not directed by electrostatic effects.

In summary, the dynamic properties of the denatured state ensemble of HP35 do not differ from unstructured peptides, a behaviour that is retained down to 1 M GdmCl. In contrast, Havlin and Tycko detected conformational preferences in ≥ 5 M GdmCl using solid-state NMR in frozen glycerol/water mixtures [41]. Our experiments further gave no indication for a general loss of tertiary contacts prior to loss of secondary structure, since the unfolding transitions probed by CD and contact formation agree well. Raleigh and coworkers could show that a fragment spanning the first two helices has a considerable bias towards helical structure and native side chain clustering [25,26]. These conformational propensities seem to be absent in the presence of low amounts of GdmCl or they do not significantly affect the dynamics in the denatured state, as contact formation along this region (Xan0-Nal23) and within this region (Xan7-Nal23) is comparable to the dynamics of unfolded model peptide chains. However, the observation that contact formation in HP35 does not differ significantly from contact formation in poly(Ser) does not exclude the presence of residual structure or a collapsed state. It rather indicates that the chain dynamics and hence the early stages of folding are not influenced in a specific manner. This is also demonstrated by the formation of the native-like contact between Xan0 and Nal35. These positions form very fast or instantaneous contact in the native state, but the denatured state dynamics are not faster than for the non-native contacts obtained for unstructured poly(Ser). We have therefore no evidence for biased dynamics or “mean native-like” topology in the denatured state, which was proposed from MD simulations [28].

Implications for Folding of HP35.

HP35 attracted much interest since it belongs to the fastest folding proteins known [42]. The wild-type folds within a few microseconds and a variant has been reported, which folds with a time constant of 700 ns [20]. Ultrafast folding might be a consequence of residual structure in the unfolded state, a view that was supported by the fragment study of Raleigh [25,26]. However, as discussed above, the chain dynamics in the GdmCl- and urea-denatured state do not deviate from those in unstructured model peptides, even under conditions strongly favouring the native state, where folding is still very fast [20]. It is therefore unlikely, that specific residual structure or a collapse governs the folding process. This is in agreement with the observation that a drastic change of the hydrophobic core residues, which would be likely to affect residual structure as well, had not effect on the folding rate constant [21].

The N- and C-terminal regions, which are close to each other in the native state, form contact with a time constant of ~ 180 ns at 22.5 °C. This seems to be close to the folding time constant of 700 ns and might suggest that folding in this variant is already close to the speed limit. However, it should be noted that the conformational re-equilibration in unfolded HP35 itself, i.e. sampling of the free-energy surface, has to be significantly faster, since single exponential kinetics are observed. In addition, local contacts can be formed much faster than long-range contacts [7] and folding might proceed from local folding nuclei.

The application of TTET to HP35, a small model protein, demonstrates the unique potential of the method to characterise dynamics in native states, intermediates or unfolded state ensembles in larger systems.

Material and Methods

Synthesis, Labelling and Purification of the HP35 Variants.

All HP35 variants were synthesized using standard 9-fluorenylmethyloxycarbonyl (Fmoc) chemistry on an Applied Biosystems 433A synthesizer. Couplings were performed with HBTU/HOBt or HATU on preloaded Tentagel S PHB-resin (Rapp Polymere).

The xanthone derivative 9-oxoxanthene-2-carboxylic acid was synthesized according to Graham and Lewis [43], activated with PyBOP and coupled to a selectively deprotected amino functionality. This was either the N-terminal amino group, the ϵ -amino group of Lys7 or the β -amino group of an α,β -diaminopropionic acid incorporated at position 35. In the latter cases methyltrityl was used as orthogonal side chain protection group, which was selectively removed with 3 % (v/v) TFA in dichloromethane. Nal was incorporated via Fmoc-protected 1-naphthylalanine (Bachem).

Final cleavage from the resin and deprotection of all other side chains was achieved with 94/2/2/2 TFA/TIPS/phenole/H₂O (v/v/v/v). All peptides were purified to > 95 % purity by preparative HPLC on a RP-8 column. Purity was checked by analytical HPLC and the identity verified by MALDI or ESI mass spectrometry.

Sample Preparation and Characterisation by CD and NMR Spectroscopy.

For all measurements 10 mM potassium phosphate buffer, pH 7.0 was used. The concentration of the HP35 variants was determined by the Xan absorption band at 343 nm ($\epsilon_{343} = 3900 \text{ M}^{-1} \text{ cm}^{-1}$). GdmCl and urea concentrations were calculated from the refractive index [44].

CD measurements were performed on an Aviv DS62 spectropolarimeter at 5.0 °C. Spectra were recorded with 50 μM HP35 in a 0.1 cm cuvette. Equilibrium unfolding transitions

were measured at 222 nm with 5 μM HP35 in a 1 cm cuvette and evaluated according to Santoro & Bolen [45].

Proton NMR spectra were recorded on a Bruker Avance 600 MHz spectrometer at 25.0 $^{\circ}\text{C}$ and referenced to 2,2-dimethyl-2-silapentane-5-sulfonate (DSS) at 0.0 ppm. Details on two-dimensional spectra and assignments are given as Supporting Information.

TTET Measurements and Data Evaluation.

TTET was measured with an Applied Photophysics LKS.60 Laser Flash Photolysis Reaction Analyzer. Xan triplet states were generated with a 4 ns laser pulse at 355 nm (Nd:YAG, Quantel Brilliant) and the absorbance bands of the Xan and Nal triplets states were monitored at 590 nm and 420 nm respectively. The measurements were performed in degassed solutions at 5.0 $^{\circ}\text{C}$ or 22.5 $^{\circ}\text{C}$ with 50 μM HP35. At this concentration intermolecular TTET can be excluded, since the Xan triplet decay of 50 μM donor-only HP35 is unaffected by the presence of 50 μM acceptor-only HP35.

For every experiment four Xan triplet decays were averaged and a sum of up to three exponential functions was used to describe the data. The signal intensity was corrected for small differences in peptide concentration. All fitting procedures were performed with ProFit (QuantumSoft) using a non-linear least-square method.

Acknowledgment

We thank Beat Fierz for discussion, Judith Habazettl for help with NMR measurements and Josef Wey for the synthesis of 9-oxoxanthene-2-carboxylic acid. This work was supported by grants from the Swiss National Science Foundation and by the Munich Center for Integrated Protein Science.

Abbreviations used:

HP35: chicken villin headpiece subdomain; TTET: triplet-triplet energy transfer; Xan: xanthone; Nal: naphthylalanine. For chemicals standard abbreviations were used, listed as Supporting Information.

References

1. Cho, J. H. & Raleigh, D. P. (2005). Mutational analysis demonstrates that specific electrostatic interactions can play a key role in the denatured state ensemble of proteins. *J. Mol. Biol.* **353**, 174-185.
2. Sanchez, 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.
3. 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.
4. Meier, S., Strohmeier, M., Blackledge, M. & Grzesiek, S. (2007). Direct observation of dipolar couplings and hydrogen bonds across a β -hairpin in 8 M urea. *J. Am. Chem. Soc.* **129**, 754-755.
5. 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.
6. Shi, Z., Chen, K., Liu, Z. & Kallenbach, N. R. (2006). Conformation of the backbone in unfolded proteins. *Chem. Rev.* **106**, 1877-1897.
7. 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.
8. 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.
9. Fierz, B. & Kiefhaber, T. (2005). Dynamics of unfolded polypeptide chains. In *Handbook of Protein Folding* (Buchner, J. & Kiefhaber, T., eds.), pp. 805-851. Wiley-VCH, Weinheim, Germany.
10. Satzger, H., Schmidt, B., Root, C., Zinth, W., Fierz, B., Krieger, F., Kiefhaber, T. & Gilch, P. (2004). Ultrafast quenching of the xanthone triplet by energy transfer: new insight into the intersystem crossing kinetics. *J. Phys. Chem. A* **108**, 10072-10079.
11. Fierz, B., Satzger, H., Root, C., Gilch, P., Zinth, W. & Kiefhaber, T. (2007). Loop formation in unfolded polypeptide chains on the picoseconds to microseconds time scale. *Proc. Natl. Acad. Sci. USA* **104**, 2163-2168.
12. Fierz, B. & Kiefhaber, T. (2007). End-to-end vs. interior loop formation kinetics in unfolded polypeptide chains. *J. Am. Chem. Soc.* **129**, 672-679.
13. Krieger, F., Möglich, A. & Kiefhaber, T. (2005). Effect of proline and glycine residues on dynamics and barriers of loop formation in polypeptide chains. *J. Am. Chem. Soc.* **127**, 3346-3352.
14. Möglich, A., Krieger, F. & Kiefhaber, T. (2005). Molecular basis for the effect of urea and guanidinium chloride on the dynamics of unfolded polypeptide chains. *J. Mol. Biol.* **345**, 153-162.
15. McKnight, C. J., Doering, D. S., Matsudaira, P. T. & Kim, P. S. (1996). A thermostable 35-residue subdomain within villin headpiece. *J. Mol. Biol.* **260**, 126-134.
16. Friederich, E., Vancompernelle, K., Huet, C., Goethals, M., Finidori, J., Vandekerckhove, J. & Louvard, D. (1992). An actin-binding site containing a

- conserved motif of charged amino acid residues is essential for the morphogenic effect of villin. *Cell* **70**, 81-92.
17. McKnight, C. J., Matsudaira, P. T. & Kim, P. S. (1997). NMR structure of the 35-residue villin headpiece subdomain. *Nat. Struct. Biol.* **4**, 180-184.
 18. Chiu, T. K., Kubelka, J., Herbst-Irmer, R., Eaton, W. A., Hofrichter, J. & Davies, D. R. (2005). High-resolution x-ray crystal structures of the villin headpiece subdomain, an ultrafast folding protein. *Proc. Natl. Acad. Sci. USA* **102**, 7517-7522.
 19. Kubelka, J., Eaton, W. A. & Hofrichter, J. (2003). Experimental tests of villin subdomain folding simulations. *J. Mol. Biol.* **329**, 625-630.
 20. Kubelka, J., Chiu, T. K., Davies, D. R., Eaton, W. A. & Hofrichter, J. (2006). Sub-microsecond protein folding. *J. Mol. Biol.* **359**, 546-553.
 21. Brewer, S. H., Vu, D. M., Tang, Y., Li, Y., Franzen, S., Raleigh, D. P. & Dyer, R. B. (2005). Effect of modulating unfolded state structure on the folding kinetics of the villin headpiece subdomain. *Proc. Natl. Acad. Sci. USA* **102**, 16662-16667.
 22. Brewer, S. H., Song, B., Raleigh, D. P. & Dyer, R. B. (2007). Residue specific resolution of protein folding dynamics using isotope-edited infrared temperature jump spectroscopy. *Biochemistry* **46**, 3279-3285.
 23. Wang, M., Tang, Y., Sato, S., Vugmeyster, L., McKnight, C. J. & Raleigh, D. P. (2003). Dynamic NMR line-shape analysis demonstrates that the villin headpiece subdomain folds on the microsecond time scale. *J. Am. Chem. Soc.* **125**, 6032-6033.
 24. Buscaglia, M., Kubelka, J., Eaton, W. A. & Hofrichter, J. (2005). Determination of ultrafast protein folding rates from loop formation dynamics. *J. Mol. Biol.* **347**, 657-664.
 25. Tang, Y., Rigotti, D. J., Fairman, R. & Raleigh, D. P. (2004). Peptide models provide evidence for significant structure in the denatured state of a rapidly folding protein: the villin headpiece subdomain. *Biochemistry* **43**, 3264-3272.
 26. Tang, Y., Goger, M. J. & Raleigh, D. P. (2006). NMR characterization of a peptide model provides evidence for significant structure in the unfolded state of the villin headpiece helical subdomain. *Biochemistry* **45**, 6940-6946.
 27. Duan, Y. & Kollman, P. A. (1998). Pathways to a protein folding intermediate observed in a 1-microsecond simulation in aqueous solution. *Science* **282**, 740-744.
 28. Zagrovic, B., Snow, C. D., Khaliq, S., Shirts, M. R. & Pande, V. S. (2002). Native-like mean structure in the unfolded ensemble of small proteins. *J. Mol. Biol.* **323**, 153-164.
 29. Lei, H., Wu, C., Liu, H. & Duan, Y. (2007). Folding free-energy landscape of villin headpiece subdomain from molecular dynamics simulations. *Proc. Natl. Acad. Sci. USA* **104**, 4925-4930.
 30. Fierz, B., Reiner, A. & Kiefhaber, T. (2007). Local conformational dynamics in α -helices measured by fast triplet transfer. *To be submitted*.
 31. 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.
 32. Vugmeyster, L., Trott, O., McKnight, C. J., Raleigh, D. P. & Palmer, A. G., 3rd. (2002). Temperature-dependent dynamics of the villin headpiece helical subdomain, an unusually small thermostable protein. *J. Mol. Biol.* **320**, 841-854.

33. 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.
34. Frank, B. S., Vardar, D., Buckley, D. A. & McKnight, C. J. (2002). The role of aromatic residues in the hydrophobic core of the villin headpiece subdomain. *Protein Sci.* **11**, 680-687.
35. Szabo, A., Schulten, K. & Schulten, Z. (1980). First passage time approach to diffusion controlled reactions. *J. Chem. Phys.* **72**, 4350-4357.
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. Lett.* **87**, 2581011-2581014.
37. 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.
38. Perl, D., Jacob, M., Bano, M., Stupak, M., Antalík, M. & Schmid, F. X. (2002). Thermodynamics of a diffusional protein folding reaction. *Biophys. Chem.* **96**, 173-190.
39. Krieger, F. & Kiefhaber, T. Kinetics and barriers of loop formation in natural loop sequences. *To be submitted*.
40. Möglich, A. & Kiefhaber, T. The peptide backbone is the major interaction site for urea and guanidinium chloride. *To be submitted*.
41. Havlin, R. H. & Tycko, R. (2005). Probing site-specific conformational distributions in protein folding with solid-state NMR. *Proc. Natl. Acad. Sci. USA* **102**, 3284-3289.
42. Kubelka, J., Hofrichter, J. & Eaton, W. A. (2004). The protein folding 'speed limit'. *Curr. Opin. Struct. Biol.* **14**, 76-88.
43. Graham, R. & Lewis, J. R. (1978). Synthesis of 9-oxoxanthen-2-carboxylic acids. *J. Chem. Soc. Perkin Trans. 1*, 876-881.
44. Pace, C. N. (1986). Determination and analysis of urea and guanidine hydrochloride denaturation curves. *Meth. Enzymol.* **131**, 266-280.
45. Bolen, D. W. & Santoro, M. M. (1988). Unfolding free energy changes determined by the linear extrapolation method. 2. Incorporation of ΔG°_{N-U} values in a thermodynamic cycle. *Biochemistry* **27**, 8069-8074.
46. Koradi, R., Billeter, M. & Wüthrich, K. (1996). MOLMOL: a program for display and analysis of macromolecular structures. *J. Mol. Graph.* **14**, 51-55, 29-32.

Table 1. Stability of donor-acceptor labelled HP35 variants and their contact formation dynamics in the denatured state.

HP35 variant	D	Thermodynamic Stability				Contact Formation in the Denatured State			
		5.0 °C		22.5 °C		5.0 °C		22.5 °C	
		m_D (kJ mol ⁻¹ M ⁻¹)	$\Delta G_D^{5°C}$ (kJ mol ⁻¹)	[D] _{1/2} (M)	E_a^{app} (kJ mol ⁻¹)	k_c^0 (10 ⁶ s ⁻¹)	m_c (J mol ⁻¹ M ⁻¹)	k_c^0 (10 ⁶ s ⁻¹)	m_c (J mol ⁻¹ M ⁻¹)
Xan0-Nal23	GdmCl	2.9 ± 0.3	6.4 ± 0.6	2.2 ± 0.2	29 ± 1	2.7 ± 0.1	417 ± 29	7.7 ± 0.3	503 ± 20
Xan0-Nal23	urea	n.d.	n.d.	n.d.	21 ± 1	2.8 ± 0.1	171 ± 45	7.5 ± 0.3	186 ± 21
Xan7-Nal23	GdmCl	3.0 ± 0.2	6.4 ± 0.4	2.1 ± 0.3	28 ± 1	4.0 ± 0.2	508 ± 19	10.0 ± 1.8	506 ± 34
Xan0-Nal35	GdmCl	3.2 ± 0.2	8.3 ± 0.5	2.6 ± 0.3	27 ± 1	2.9 ± 0.3	572 ± 80	5.5 ± 0.3	522 ± 30
Nal23-Xan35	GdmCl	3.0 ± 0.2	7.0 ± 0.6	2.3 ± 0.4	28 ± 1	3.2 ± 0.1	416 ± 14	7.7 ± 0.8	429 ± 23

The thermodynamic stability was calculated from unfolding transitions at 5.0 °C, monitored by CD at 222 nm (Fig. 2 and Material and Methods). In all cases the slope of the native baseline was assumed to be zero. With urea no reliable data (n.d.) could be obtained, due to the broadness of the transition. Contact formation in the denatured state was measured at 5.0 °C (Fig. 3-5) and 22.5 °C (Fig. 6). The rate constants were extrapolated to 0 M denaturant (Eq. 1), giving k_c^0 and m_c . * Apparent activation of energies of contact formation obtained from T-dependencies measured in 4.0 M GdmCl or 6.0 M urea (not corrected for changes in solvent viscosity).

Figure Captions

Figure 1. Ribbon diagrams of the donor-acceptor labelled villin headpiece subdomain (HP35) variants used to measure triplet-triplet energy transfer (TTET). The triplet donor xanthone (Xan) is shown in red, the triplet acceptor naphthalene (Nal) in blue. The models are based on the X-ray structure of HP35 [18] (PDB 1YRF) and were prepared using MolMol [46].

Figure 2. Structure and stability of HP35 Xan0-Nal23. (A) One-dimensional NMR spectrum in 5 % (v/v) D₂O at 25.0 °C. (B) GdmCl-induced unfolding transition at 5.0 °C monitored by CD at 222 nm. The line represents the fit of a two-state model with the data given in Table 1. In all experiments 10 mM potassium phosphate, pH 7.0 was used as buffer.

Figure 3. TTET in the HP35 variant Xan0-Nal23 at 5.0 °C. (A) Representative decays of the Xan triplet absorption band (590 nm) at different GdmCl concentrations (violet to red: 1.0 M, 1.7 M, 2.1 M, 2.5 M, 3.0 M, 4.2 M). The intrinsic lifetime of Xan0 is represented by the donor-only decay (grey). In donor-acceptor labelled HP35 the triplet state decays faster at higher GdmCl concentrations, due to TTET upon contact formation in the denatured state. The solid lines show fits of double exponential functions, with the rate constants given in (B) and the normalized amplitudes in (C). The slow phase corresponds to the triplet lifetime of Xan0 in the native state (grey). In the denatured state contact formation between Xan0 and Nal23 (black) can be described according to Eq. 1 (line), the parameters are given in Table 1. For comparison the fraction of unfolded molecules obtained by CD (Fig. 2B) is shown in blue.

Figure 4. TTET in the HP35 variant Xan0-Nal35 at 5.0 °C. (A) Representative decays of the Xan triplet absorption band (590 nm) at different GdmCl concentrations (violet to red: 1.3 M, 2.2 M, 2.6 M, 4.2 M, 5.5 M, 7.4 M). The intrinsic lifetime of Xan0 is represented by the donor-only decay (grey). At low GdmCl concentrations amplitude is missing and fast TTET is observed. Upon unfolding with GdmCl, TTET becomes slower and the whole Xan triplet decay can be resolved. The solid lines show fits of double exponential functions with the rate constants given in (B) and the normalized amplitudes in (C). The phases reflect contact formation in the native state (green, closed symbols) and contact formation in the denatured state (black). An additional faster TTET process in the native state cannot be resolved, as amplitude is missing (green, open symbols). Contact formation in the denatured state is described by Eq. 1 (black line), the parameters are given in Table 1. For the native state contact formation the fit gives $k_c = (2.9 \pm 0.3) \cdot 10^7 \text{ s}^{-1}$ and $m_c = (750 \pm 140) \text{ J mol}^{-1} \text{ M}^{-1}$ (green line). For comparison the fraction of unfolded molecules obtained by CD (Fig. S1) is shown in blue.

Figure 5. TTET in the HP35 variant Nal23-Xan35 at 5.0 °C. (A) Representative decays of the Xan triplet absorption band (590 nm) at different GdmCl concentrations (violet to red: 0 M, 1.3 M, 1.7 M, 2.7 M, 3.9 M, 6.2 M). The intrinsic lifetime of Xan35 is represented by the donor-only decay (grey). Efficient TTET is observed at all GdmCl concentrations, independent of whether HP35 is folded or not. (B) Rate constants obtained from the double or triple exponential fits and (C) the normalized amplitudes of the phases. Contact formation in the native state comprises two phases (red, green), the third phase can be attributed to contact formation in the denatured state (black). The

parameters for Eq. 1 describing the denatured state dynamics (line) are given in Table 1. The open symbols indicate that this rate constant was fixed for fitting the corresponding traces. For comparison the fraction of unfolded molecules obtained by CD (Fig. S1) is shown in blue.

Figure 6. Contact formation in the denatured state of the HP35 variants Xan0-Nal23 (orange), Xan7-Nal23 (green), Xan0-Nal35 (violet) and Nal23-Xan35 (brown) at 22.5 °C. (A) GdmCl-dependencies of the rate constants for contact formation, k_c , which can be described by Eq. 1 (parameters given in Table 1). (B) Comparison of k_c with those determined for end-to-end contact formation in unstructured poly(Gly-Ser) and poly(Ser) peptide chains of different length [8]. Tails were found to slow down k_c up to a factor of 2.5, indicated by the dotted line [12].

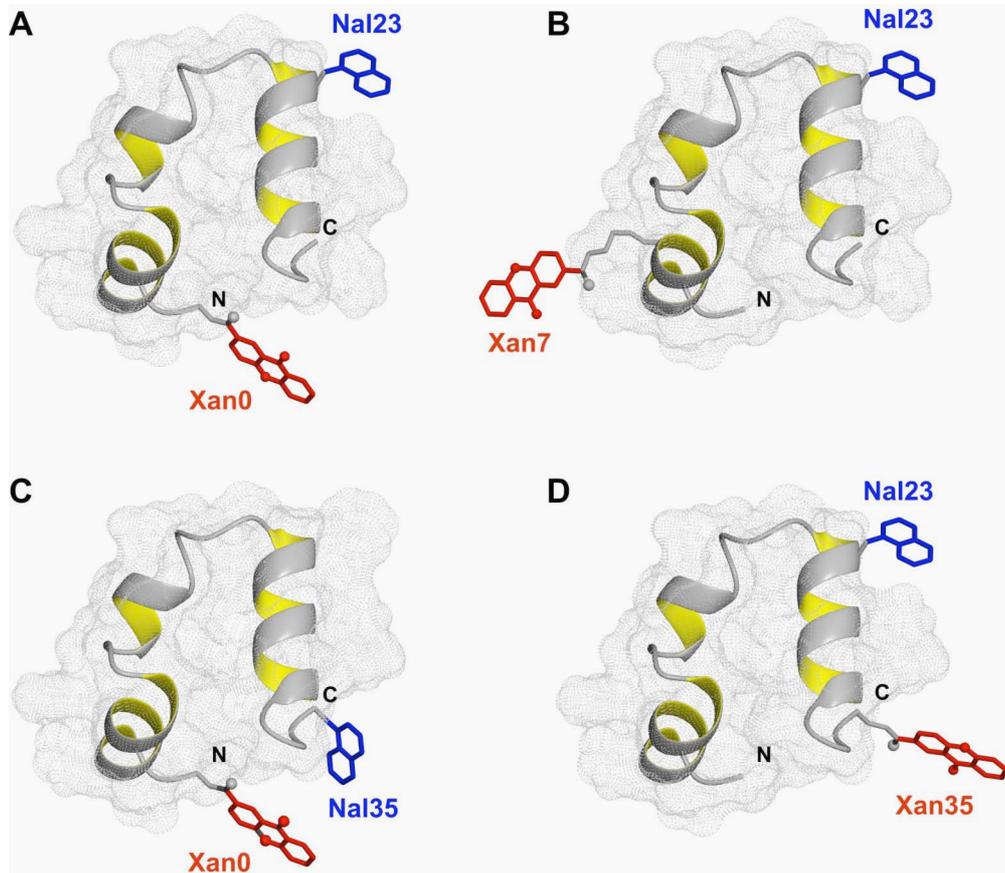


Figure 1

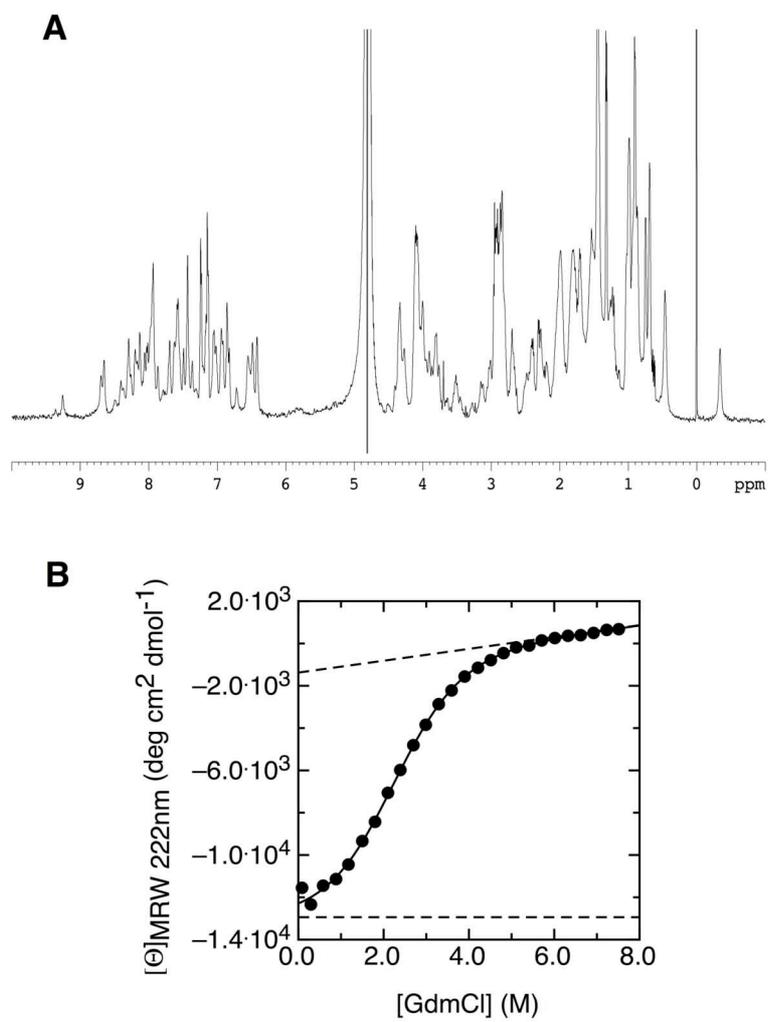


Figure 2

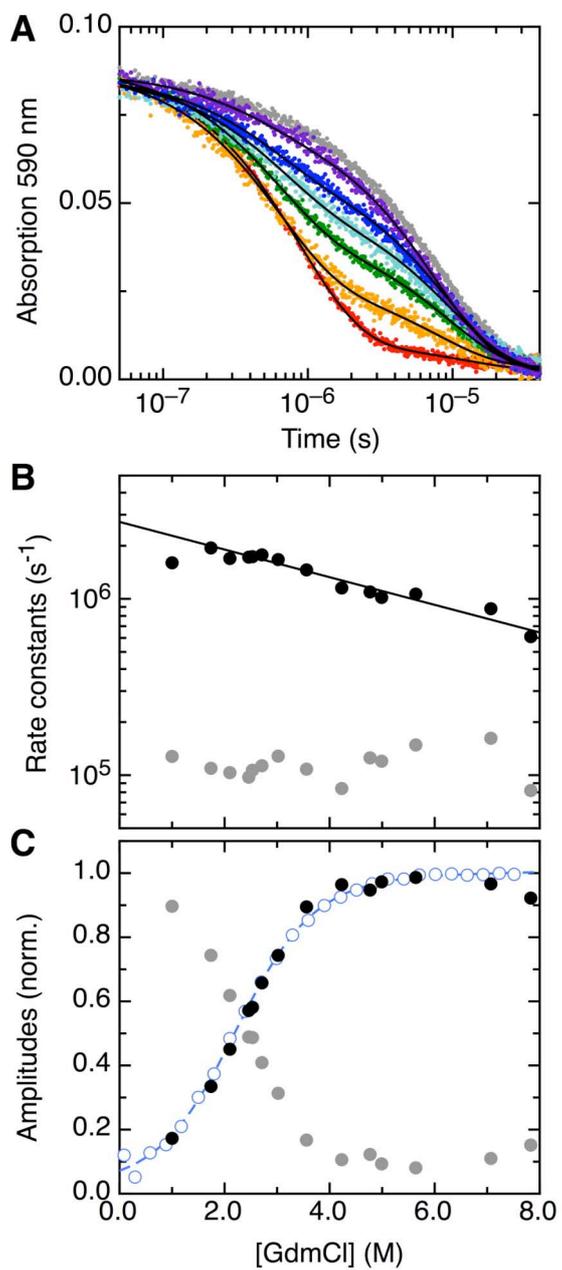


Figure 3

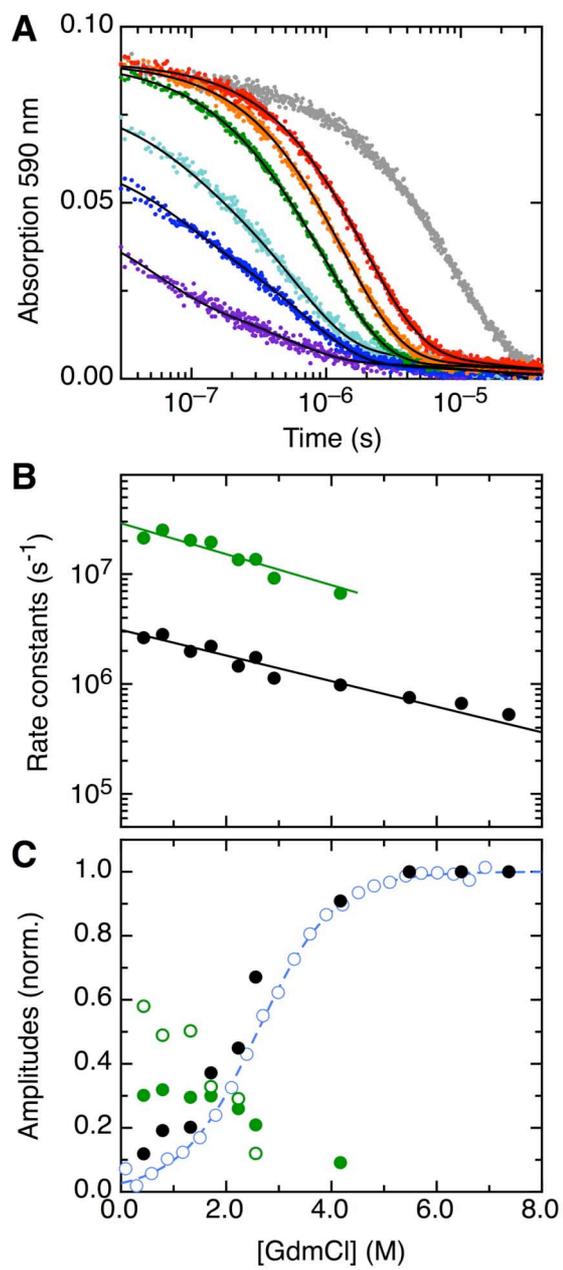


Figure 4

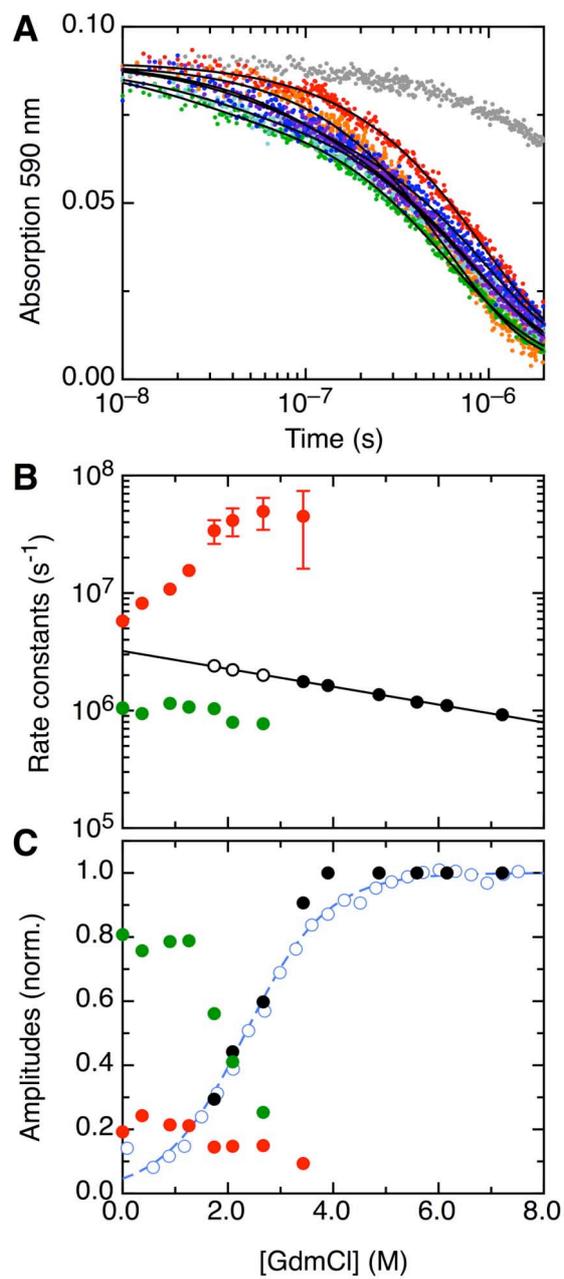


Figure 5

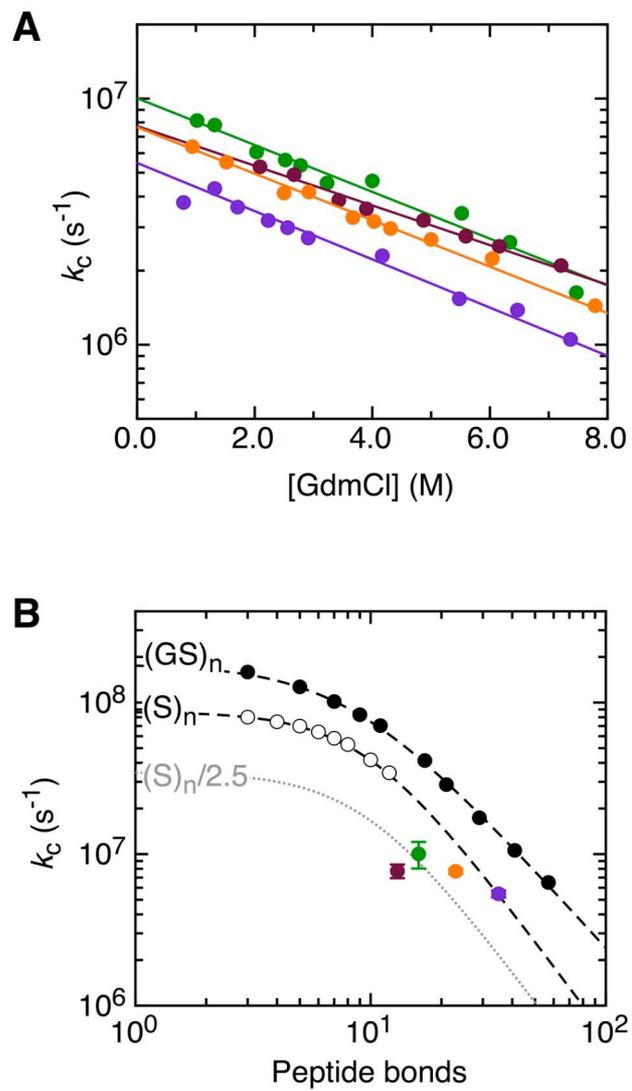


Figure 6

Supporting Information

1. Abbreviations of Chemicals
2. CD and NMR Characterisation of Double-labelled HP35 Variants
3. TTET Data on Xan7-Nal23

1. Abbreviations of Chemicals

GdmCl: guanidinium chloride; HATU: *O*-(7-azabenzotriazol-1-yl)-*N,N,N',N'*-tetramethyluronium hexafluorophosphate; HBTU: *O*-(benzotriazol-1-yl)-*N,N,N',N'*-tetramethyluronium hexafluoro-phosphate; HOBt: *N*-hydroxybenzotriazole; Nle: norleucine; PyBOP: (benzotriazol-1-yloxy)tripyrrolidinophosphonium hexafluoro-phosphate; TFA: trifluoroacetic acid; TIPS: tripropylsilane.

2. CD and NMR Characterisation of Double-labelled HP35 Variants

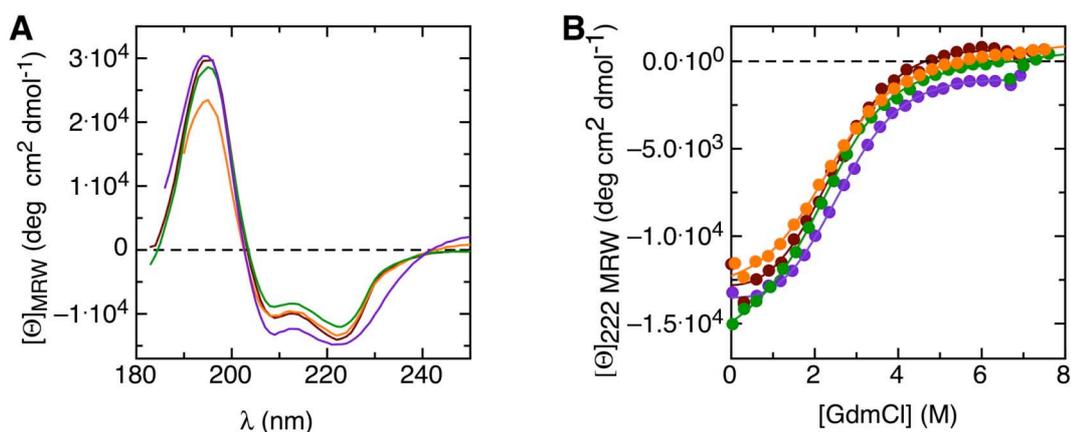


Figure S1. (A) Far-UV CD spectra of the HP35 variants Xan0-Nal23 (orange), Xan7-Nal23 (green), Xan0-Nal35 (violet) and Nal23-Xan35 (brown) at 5.0 °C. (B) GdmCl-induced unfolding equilibrium transitions at 5.0 °C monitored by CD at 222 nm. The lines show fits of a two-state model, with the parameters given in Table 1. The slopes of the native baselines were assumed to be zero. For the measurements 10 mM potassium phosphate, pH 7.0 was used as buffer.

The deviations observed for Xan0-Nal35 are likely due to contributions of Xan or Nal, which absorb strongly in this spectral region ($\epsilon \approx 10^3 - 10^5 \text{ M}^{-1} \text{ cm}^{-1}$). In this variant they seem to sense an asymmetric environment.

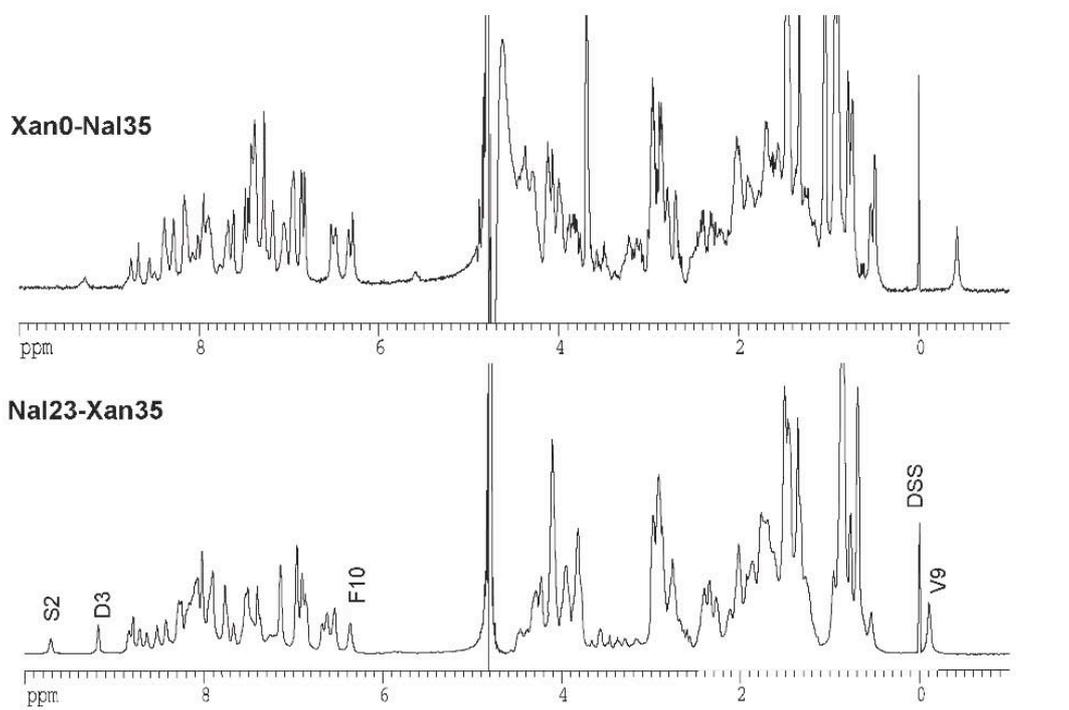


Figure S2. One-dimensional ^1H NMR spectra of Xan0-Nal35 and Nal23-Xan35 in 5 % (v/v) D_2O at 25.0 $^\circ\text{C}$, pH 7.0. The spectrum of Xan0-Nal23 is shown in Fig. 2. The assignment of the amide protons of Ser2 and Asp3, of the aromatic side chain proton of Phe10 and of the Val9 methyl group protons was confirmed with NOESY spectra. See also Reference [15].

3. TTET Data on Xan7-Nal23

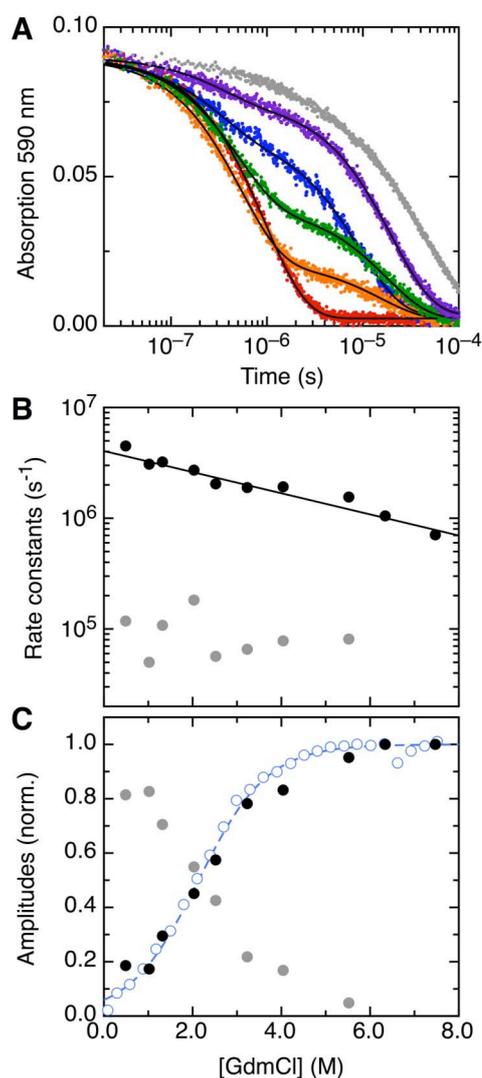


Figure S3. TTET in the HP35 variant Xan7-Nal23 at 5.0 °C. (A) Representative decays of the Xan triplet absorption band (590 nm) at different GdmCl concentrations (violet to red: 1.0 M, 1.3 M, 2.5 M, 3.2 M, 6.3 M). The intrinsic lifetime of Xan7 is represented by the donor-only decay (grey). In donor-acceptor labelled HP35 the triplet state decays faster at higher GdmCl concentrations, due to TTET upon contact formation in the denatured state. The solid lines show fits of double exponential functions with the rate constants given in (B) and the normalized amplitudes in (C). The slow phase reflects the triplet lifetime of Xan7 in the native state (grey). In the denatured state contact formation between Xan7 and Nal23 (black) can be described according to Eq. 1 (line) with the parameters given in Table 1. For comparison the fraction of unfolded molecules obtained by CD (Fig. S1) is shown in blue.