Fakultät für Chemie

Lehrstuhl für Biophysikalische Chemie

Dynamics in Folded and Unfolded Peptides and Proteins Measured by Triplet-Triplet Energy Transfer

Natalie D. Merk

Vollständiger Abdruck der von der Fakultät für Chemie
der Technischen Universität München zur Erlangung des akademischen Grades eines

Doktors der Naturwissenschaften

genehmigten Dissertation.

Vorsitzender: Univ.-Prof. Dr. Aymelt Itzen

Prüfer der Dissertation:

1. Univ.-Prof. Dr. Thomas Kiefhaber (Martin-Luther-Universität Halle-Wittenberg)
2. Univ.-Prof. Dr. Michael Groll

Die Dissertation würde am 09.02.2015 bei der Technischen Universität München eingereicht und durch die Fakultät für Chemie
am 05.03.2015 angenommen.

Content

1. Introduction

1.1 Proteins

Proteins, the major class of macromolecules in living organisms, are found in all cells. They are not only responsible for the form and structure of cells, but also serve as 'molecular machines' and fulfill a magnitude of functions such as catalysis, signal transduction, transport and regulation. Proteins are linear polymer chains, composed of 20 proteinogenic α-L-amino acids connected by peptide bonds. The side chains of the amino acids have different structures, determining their physical and chemical properties. The information about the sequence of amino acids of a protein is provided by protein-encoding genes. After transcription of a gene's DNA, the messenger RNA (mRNA) carries the information out of the nucleus to the ribosome, where the protein chain is assembled. In higher organisms, further post-translational modifications such as the attachment of phosphates, lipids or carbohydrates often take place.

In order to fulfill its biological function, the polypeptide chain has to adopt a well-defined three-dimensional structure, predetermined by the amino acid sequence. This process is referred to as protein folding. Proteins that do not fold properly tend to aggregate and were found to be the cause of several diseases such as Alzheimer, Parkinson's and Creutzfeldt-Jacob disease[1].

1.2 Protein folding

During protein folding, the polypeptide chain, starting from the unfolded or denatured state (U), has to search for energetically favorable interactions to from secondary structured elements like loops, helices or β-sheets. In the native state (N), the secondary elements are arranged in a well-defined three-dimensional structure (tertiary structure). This process occurs spontaneously, but may also be catalyzed *in vivo* by disulfide isomerases and peptidyl-prolyl isomerases[2]. The thermodynamic states U and N are separated by a free energy barrier and are in equilibrium with a postulated state, the activated complex or transition state, which is situated on the saddle point of the barrier[3]. Proteins exist as an equilibrium mixture between their native and denatured state and continuously fold and unfold.

Modern studies about protein folding, starting with Christian Anfinsen in the 1960ies[4], aim to answer basic questions of how the amino acid sequence of a protein determines its three-dimensional structure and about the mechanism of the folding process. Final goals are the

prevention of protein misfolding as well as the prediction of the native structure from the amino acid sequence and, in reverse, the design of proteins with new properties.

1.2.1 The unfolded state of proteins

In theory, the unfolded state is defined as an unstructured random coil[5, 6]. This ideal polymer with *n* bonds of fixed length *l* performs a random walk in a three-dimensional space. The angles at the bond junctions can adopt all values with the same probability, independent of each other. The average end-to-end distance of the polymer $<r^2>$ is described as

$$< r^2 > = n \cdot l^2 \qquad (1.1)$$

The distribution of the end-to-end distances $W_{(r)}$ is described as

$$W_{(r)} = \left(\frac{3}{2\pi<r^2>}\right)^{\frac{3}{2}} \cdot e^{\left(\frac{3r^2}{2<r^2>}\right)} \qquad (1.2)$$

The random coil model is not realistic as it assumes equally probable angles, since in real polypeptide chains, the possible angles of the bond junctions of the chain backbone are restricted. The geometry for each amino acid, and therefore the overall conformation of the chain, is specified by peptide dihedral angles ψ and ϕ the angle of the peptide bond ω (Figure 1.1).

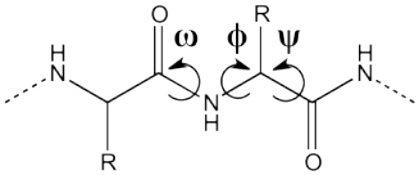


Figure 1.1 Peptide dihedral angles ψ and ϕ and angle of the peptide bond ω. Side chains are represented as R

The Ramachandran plot visualizes backbone dihedral angles ψ against ϕ of amino acid residues in a protein structure[7]. Assuming a hard sphere atomic model, a conformation is not allowed if it results in an atomic clash. This eliminates a large part of the conformational space and therefore reduces the conformational entropy U. If hydrogen bonding constraints are taken into account, the region for allowed ψ and ϕ is even more restricted[8] (Figure 1.2).

This results in stiffer chains than expected from the random coil model and therefore longer end-to-end distances.

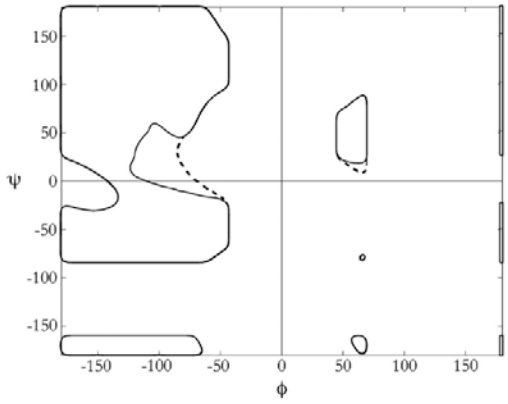


Figure 1.2 Ramachandran plot of an alanine dipeptide. Allowed conformations are within the dashed lines. The solid line shows the restricted region, when hydrogen-bonding constraints are taken into account. Figure from[8]

In order to adapt the random coil model to stiffer chains, Flory introduced the characteristic ratio (C_n)[9].

$$< r^2 > = C_n \cdot n \cdot l^2 \qquad (1.3)$$

C_n is a measure for the dimensions of a stiff chain and increases with chain length for short chains. For long chains ($n \rightarrow \infty$), C_n reaches a constant limiting value C_∞; for this case equation 1.3 can be rewritten.

$$< r^2 > = C_\infty \cdot n \cdot l^2 \qquad (1.4)$$

For an ideal Gaussian chain with equally probable angles, C_∞ is 1, which would result in equation 1.1. For real chains, C_∞ is larger than 1 with little contribution from flexible glycine residues and a large contribution from stiffer amino acids.

Other measures for chain stiffness are the Kuhn length *b*, the hypothetical length of chain segments that can move without experiencing chain stiffness[10, 11] and the persistence length l_p which is a measure for the distance that an infinitely long chain continues in the same direction. They are defined by

$$b = C_\infty \cdot l \qquad (1.5)$$

$$2 \cdot l_p = b \qquad (1.6)$$

Kuhn was also the first to discuss excluded volume effects[10] which means that one part of a chain can not occupy space that is already occupied by another part of the same chain. The result is increased chain dimension and a larger end-to-end distance. Flory included the contribution from excluded volume effects by an approximate exponent[12] (equation 1.7).

$$\sqrt{< r^2 >} \propto \ l \cdot n^{v} \qquad \text{with } v = 0.59 \qquad (1.7)$$

Several theories have been developed to describe the macroscopic dynamics of a polymer chain. In the Rouse model[13] and the Zimm model[14], the polymer is treated as a set of beads connected by harmonic springs. In the Rouse model, the dynamic behavior is defined by Brownian motions and described by the Langevin equation (1.8).

$$\ddot{x} = - M^{-1} \frac{\partial U_{(x)}}{\partial_x} - \gamma \dot{x} + \ M^{-1} F_{fluct}(t) \qquad (1.8)$$

Here M is the particle mass, x is the reaction coordinate, $U_{(x)}$ denotes the energy, γ is a friction coefficient and $F_{fluc}(t)$ is the random force which represents the thermal motion of the environment.

In the Zimm model, additional hydrodynamic interactions between different parts of the chain and the solvent are taken into account.

The Szabo, Schulten, Schulten theory[15] (SSS) describes the kinetics of end-to-end contact formation by a single exponential decay (equation 1.9).

$$\sum(t) \approx \sum_{approx}(t) = e^{\left(-\frac{t}{\tau}\right)} \qquad (1.9)$$

Here τ is the average reaction time and $\sum(t)$ is the probability that loop formation not yet happened at time *t*.

Under physiological conditions, most molecules are folded; therefore, investigation of the unfolded state in this environment is difficult. Anyway, U can be investigated under denaturing conditions such as high temperature, high pressure, extremes of pH or the presence of denaturants. Tanford and others investigated polypeptide chains under strong denaturing conditions and found them to be highly expanded and solvent exposed[16]. However, it is not

clear whether the unfolded state under denaturing conditions is comparable to the unfolded state under physiological conditions.

Investigations of the unfolded state under native conditions were done using protein fragments[17-20], short and flexible model peptides[21-25], proteins destabilized by reduced disulfide bonds or mutations[26, 27] as well as intrinsically disordered proteins[28-30].

Hereby, it was shown that unfolded species under native conditions deviate significantly from random coil behavior and are more compact than the urea-denatured state. Residual structure has also often been found in the unfolded state of proteins.

However, U resists distinct structural characterization but represents a highly dynamic ensemble of multiple states with low barriers between the conformations.

1.2.2 The native state of proteins

The three-dimensional structure of a protein under physiological conditions, which is related to its biological function, is referred to as the native state (N). Since Anfinsen's experiments, it is generally accepted that the native structure and the fundamental mechanism of folding is determined by the protein's amino acid sequence[31]. The linear arrangement of amino acids in a protein and the location of covalent linkages such as disulfide bonds is referred to as primary structure.

Hydrogen bonds (H-bonds) between the polypeptides backbone carbonyl oxygen and N_{amide} proton lead to the formation of secondary structure elements such as α-helices, β-sheets and turns[32, 33]. Forming H-bonds is energetically favorable. Actually, backbone polar groups deprived of a H-bond partner (intramolecular or from the solvent) act as destabilizing influences in the molecule[8].

Water is a poor solvent for apolar molecules. Thus solvent exposed hydrophobic side chains result in an ordering of water and a loss in solvent entropy. Placing multiple non-polar side chains in close proximity in the protein's core minimizes unfavorable, ordered water. This important driving force for the correct folding of the native state is referred to as the hydrophobic effect[34, 35]. Weak, short range interactions resulting from transient induced dipoles in the electron cloud surrounding an atom that are referred to as van der Waals interactions contribute to the strength of the hydrophobic effect.

In the next step, the tertiary structure of a protein and therefore its final specific geometric shape is determined by electrostatic close-range and long-range interactions (salt bridges) between the charged side chains of the amino acids as well as the free ionized groups of the

amino- and carboxy-termini[36, 37]. A self-contained tertiary structured unit, which can fold independently and also in isolation, is called a domain[38].

Several methods such as X-ray crystallography[39] and nuclear magnetic resonance (NMR) spectroscopy[40] are currently used to determine the native structure of a vast number of proteins with atomic resolution. However, research suggests that proteins are not rigid. Their biologically active state is an ensemble of native conformations, which are subject to thermal fluctuations and motions on a wide range of timescales. Single bond vibrations and exchange of loosely associated water molecules occur on the timescale of femto- to picoseconds[41]. Using vibrational and NMR spectroscopy molecular motions on nanosecond or on micro- to millisecond timescales can be monitored[42, 43], revealing movements of surface loops and side chains, but also of the core structure. Thus, internal protein motions correspond to the inter-conversion of protein conformations and the sampling of multiple sub-states which is vital for a number of protein functions including molecular recognition processes[44], binding affinities and enzyme catalysis[45].

1.2.3 Protein stability

The folding transition of most small, single-domain proteins from U to N is a reversible reaction and highly cooperative, which means that molecules within a population are either completely folded or unfolded. Partially unfolded polypeptides exist in a negligible amount. Thus, their folding reaction can be described by two-state behavior:

$$U \underset{k_u}{\overset{k_f}{\rightleftharpoons}} N \qquad (1.10)$$

The folding rate constant is k_f and the unfolding rate constant is k_u. The ratio of the equilibrium concentrations of N and U and also of the folding and unfolding rate constant k_f and k_u is K_{eq} and determines the free energy of the folding reaction ΔG^0 by

$$\Delta G^0 = -RT \cdot ln\, K_{eq} = -RT \cdot ln \frac{[N]}{[U]} = -RT \cdot ln \frac{k_f}{k_u} \qquad (1.11)$$

The free energy of protein folding is usually small and in the range of -15 to -60 kJ/mol[46]. This is due to compensating effects from enthalpy and entropy. Whereas formation of intramolecular interactions such as hydrogen bonds and salt bridges is energetically favorable,

the deprivation of interactions from the protein to the solvent is unfavorable. Additionally, the native state is destabilized by a loss of chain entropy, while release of water molecules from the hydrophobic protein's core achieves a gain of entropy.

The change of the free energy difference between U and N is given by the Gibbs fundamental equation[47].

$$d\Delta G^0 = \Delta V^0 dp - \Delta S^0 dT + \sum_i \Delta\mu_i^0 dn_i \qquad (1.12)$$

ΔV^0 is the difference in volume, ΔS^0 the difference in entropy and $\Delta\mu^0$ the difference in chemical potential. ΔG^0 can be changed by varying the pressure p, the temperature T or by adding denaturants or other co-solutes to the solvent and thereby changing its composition n_i. Including the effect of denaturant on protein folding, equation 1.12 can be rewritten.

$$d\Delta G^0 = \Delta V^0 dp - \Delta S^0 dT + m d[D] \qquad (1.13)$$

The change of ΔG^0 with the denaturant concentration linearly depends on the proportionality constant m, which in turn correlates with the change in solvent accessible surface (SASA) between N and U.

The mechanism by which common denaturants such as urea or GdmCl destabilize protein structure is controversial. Research suggests that urea directly binds to the protein's backbone[48], while GdmCl leads to higher solvation of the protein surface[5, 6].

In 1973, Christian Anfinsen stated that the three-dimensional structure of a native protein in its normal physiological milieu is the one in which the Gibbs free energy of the whole system is the lowest[31]. Today, we know several exceptions from this rule such as intrinsically disordered proteins or misfolded proteins involved in prion diseases.

1.2.4 Barriers in protein folding

In 1969, Levinthal stated his famous paradox. He found that because of the large number of degrees of freedom in an unfolded polypeptide chain, the molecule has an astronomical number of possible conformations, but folds spontaneously on a millisecond to microsecond timescale[49]. He suggested that the paradox can be solved if protein folding is guided by the rapid formation of local interactions, which then determine the further folding of the peptide.

Today, we know that proteins fold on a timescale that ranges from a few microseconds[50] to several hundred seconds[51, 52]. Different kinds of motions occur during protein folding such as

side chain rotations within picoseconds, loop motions within nanoseconds[21, 53, 54] or the rearrangement of whole domains on microsecond timescale[55]. Slow, mostly rate-determining steps in protein folding are proline *cis-trans* isomerization and formation of disulfide bridges[56]. There is a high energy barrier for isomerization (16-20 kcal/mol in model compounds) of X-Pro bonds[57] which results in flipping times of ten to thousand seconds at room temperature[58]. In vivo, enzymes such as peptidyl-prolyl *cis-trans* isomerases lower the energy barrier for isomerization and speed up folding[59, 60].

Formation of disulfide bridges often takes place during synthesis[61, 62]. Thereby, depending on the number of cysteine residues and on the native structure, errors in disulfide pairing can occur[63]. Repairing of these errors by breaking up wrongly formed bridges proceeds in a trial and error manner[64].

Assuming two-state folding, the folded and the unfolded state are separated along the reaction coordinate by a free energy barrier, with the transition state (TS) located on the barrier top (Figure 1.3).

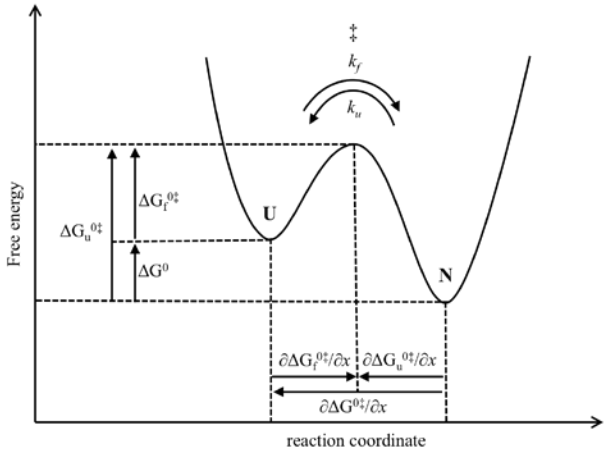


Figure 1.3 Hypothetical free energy profile of a two-state transition. The total free energy difference ΔG^0 is the energy difference of the unfolded (U) and the native state (N). The activation free energies $\Delta G_u^{0‡}$ and $\Delta G_f^{0‡}$ are the height of the barrier.

According to the transition state theory of Eyring[3], the rate constant for a reaction depends on the free energy barrier $\Delta G^{0‡}$

$$k = \frac{k_B \cdot T}{h} \cdot e^{\left(-\frac{\Delta G^{0‡}}{RT}\right)} \qquad (1.14)$$

where h is the Planck constant and k_B is the Boltzmann constant. This reflects the maximum rate constants for elementary reactions in organic chemistry where covalent bonds are formed and broken. For folding reactions equation 1.14 can be rewritten to

$$k = k_0 \cdot e^{\left(-\frac{\Delta G^{0\ddagger}}{RT}\right)} \qquad (1.15)$$

Here k_0 is the pre-exponential factor and represents the maximum rate constant in absence of free energy barriers. According to the original Eyring theory, the pre-exponential factor corresponds to the frequency of a single bond vibration ($\approx 6\times10^{12}$ s^{-1} at 25 ºC). Since the rate constant for protein folding reactions is strongly influenced by intrachain diffusion processes, k_0 here is expected to be slower ($\approx 10^7$ to 10^8 s^{-1})[53].

In order to characterize the free energy barrier and the transition state rate equilibrium, free energy relationships (REFERs) can be applied. Here, it is assumed that effects of perturbations on the kinetics of protein folding reactions are linearly related to the corresponding effects on thermodynamics. The factor α_x is introduced as a measure for the sensitivity of the transition state to perturbations ∂x relative to the ground states (equation 1.16)[65].

$$\alpha_x = \frac{\partial \Delta G^{0\ddagger}/\partial x}{\partial \Delta G^{0}/\partial x} \qquad (1.16)$$

For $\alpha_x = 1$, the transition state has the same properties as the native state, while for $\alpha_x = 0$, the transition state resembles the unfolded state.

Applying the Gibbs fundamental equation 1.13 to the transition state, one can write

$$d\Delta G_f^{0\ddagger} = \Delta V_f^{0\ddagger} dp - \Delta S_f^{0\ddagger} dT + m_f[D] \qquad (1.17)$$

Variation of the pressure p gives information about the Volume ($\Delta V_f^{0\ddagger}$) of the transition state[66]. By changing the temperature T information about the entropy ($\Delta S_f^{0\ddagger}$), enthalpy ($\Delta H_f^{0\ddagger}$) or heat capacity ($\Delta C_{p\,f}^{0\ddagger}$) of the transition state[67-69] can be obtained. Denaturant dependencies give information about changes in SASA[70].

The ϕ–value analysis is used to map the structure of transition states[71-73]. Analogical to the α_x-value of REFERs, ϕ is a measure for the sensitivity of TS to perturbations caused by site-

directed mutagenesis. A ϕ-value of 1 indicates that the interaction to the mutated side chain is already formed in the transition state.

Protein folding reactions may sometimes show non-linear REFERs; in this case, α_x is sensitive to changes in ΔG^0. Non-linear REFERs are difficult to characterize but can give valuable information about the shape of the free energy barrier. Causes for nonlinear behavior may be movement of the position of the TS along the reaction coordinate[74-76], structural changes in the unfolded or native state[77, 78] as well as the presence of parallel or sequential reaction pathways[79, 80].

1.2.5 The effect of friction on protein folding kinetics

In order to understand protein folding kinetics, friction as an additional limiting factor has to be taken into account. Kramers' transition state theory[81] considers the influence of the solvent to chemical reactions. It describes the crossing of a hypothetical particle over an energy barrier with the height *E*.

$$k = \frac{\omega_0 \omega_{\ddagger}}{2\pi\gamma} \cdot e^{\left(-\frac{E}{RT}\right)} \qquad (1.18)$$

ω_0 and $\omega_{\ddagger}$ represent motion of the system in the starting well and the on barrier top. γ is a friction coefficient for motion in the solvent and according to the Stokes' law, it is proportional to its viscosity η. Assuming a pre-exponential factor that is related to ω_0 and $\omega_{\ddagger}$, equation 1.18 can be rewritten (equation 1.19).

$$k = k_0 \cdot \left(\frac{\eta_0}{\eta}\right)^{\beta} \qquad (1.19)$$

Thus, the reaction rate is inversely coupled to the solvent viscosity η. For β = 1, the dependence is linear, which means that the reaction is diffusion controlled. Experimentally, this was found to be true for the loop formation of long unfolded model poly(Gly-Ser)$_x$ peptides with more than 15 peptide bonds[82]. For short chains with numbers of peptide bonds smaller than 10, β deviates from 1 reaching values from 0.81 to 0.96[21]. Similar observations were made for motions of simple polymers in organic solutions[83-85] and for dynamics of native proteins[86-88].

In 1992, the concept of internal friction was introduced[89] to explain deviation from the Stokes' law. Eaton and co-workers suggested two sources of friction to influence protein

conformational changes. One is the friction of the solvent, which retards the motion of atoms on the surface of the protein; the other is the internal friction of the protein, which slows the motion of protein atoms relative to each other. The source of internal friction might be steric effects or dissipative interactions within the polypeptide chain[90].

If one assumes that the friction of the protein and solvent is additive, internal friction can directly be integrated in the Kramers equation 1.18. However, researchers argued that the folding time $\tau = k^{-1}$ in the limit of small solvent viscosity gives a clearer indicator of the internal friction contribution to protein dynamics, since it is interpreted as the speed of the rate-limiting step when the diffusional motions in the chain become extremely rapid[91, 92]. Therefore, it was proposed that the folding time is the sum of two separate time scales; one solvent-controlled relaxation time τ_{solv} and a solvent-independent τ_{int} (equation 1.20)[93].

$$\tau = \tau_{int} + \tau_{solv} \cdot \left(\frac{\eta}{\eta_0}\right)^{-\beta} \qquad (1.20)$$

Contributions from τ_{int} can be directly determined by measuring τ at different viscosities and extrapolation of $\tau_{solv} \rightarrow 0$. Plotting τ against the viscosity η would result in a linear (for $\beta = -1$) or sloping (for $\beta > -1$) curve with the intercept representing τ_{int}.

Several researches used this strategy setting β to -1. Thereby, for tryptophan Cage folding τ_{int} was found to be about 0.7 µs independent of the temperature[94]. 8 µs was reported for the folding of cytochrome c at 20 ºC[91]. For intact proteins folding at millisecond or slower rates the viscosity dependence is often seen to be linearly related to the viscosity without contributions of internal friction[92, 95, 96]. That would result in a β-value of -1 and $\tau_{int} = 0$.

Studying the transition path time and the folding time of a two-state, all-α-helical designed protein, Chung and Eaton found β-values of -0.19 and -0.3, respectively, and an intercept of zero[97]. However, they measured at very high concentrations of viscogen, an experimental setup in which the assumptions of solvent-independent internal friction, two-stateness and unchanged curvatures of the free-energy surface may not longer hold.

Today, it is not established definitely whether the data are better fit by a power law with β deviating from -1 or by a linear relation. Anyhow, for an accurate determination of internal friction in protein folding, fast folding reactions at low ranges of viscosity should be investigated. Therefore, techniques with high time-resolutions are needed.

1.3 Triplet-triplet energy transfer as suitable tool to study protein folding

In order to understand protein folding it is essential to gain knowledge about the dynamics in polypeptide chains. Triplet-triplet energy transport (TTET) is a method that allows for measurements of absolute rate constants for site-specific contact formation within the chain. Therefore, it is a suitable tool to obtain information about dynamics in the unfolded state such as loop formation as well as local unfolding or folding events in the native state (Figure 1.4). In principle, TTET is based on the contact formation between two labels that are attached at specific sites of a polypeptide chain.

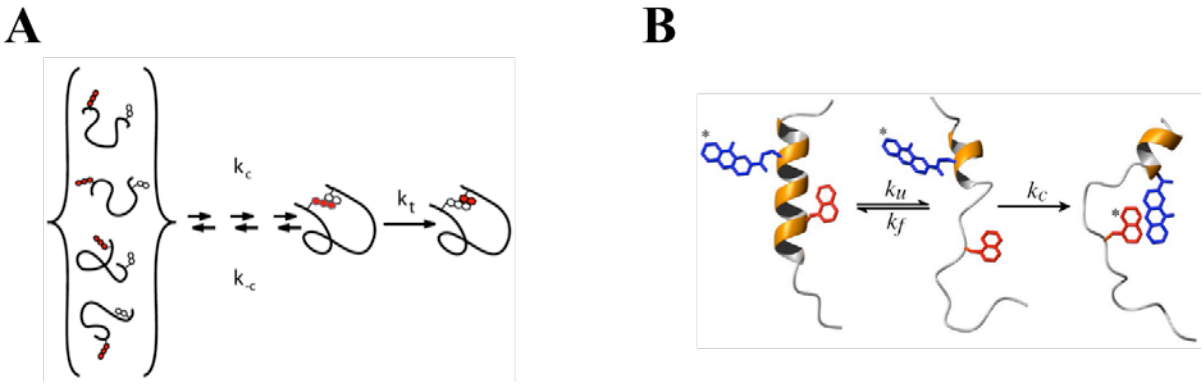


Figure 1.4 TTET as a method to study loop formation (A) as well as local unfolding/folding events (B). Figure A taken from [53], B taken from[98].

For TTET, the donor molecule xanthone is selectively excited to the S_1 state by a short laser pulse at 355 nm. This leads to a fast (~ 2 ps) intersystem crossing (ISC) to the triplet state T_1 with a quantum yield of ~ 99 %[99, 100]. The intrinsic lifetime of the triplet state of xanthone is between 20 and 30 µs in oxygen-free solutions[82]. If the donor (xanthone) and the acceptor (naphthalene) lie within van der Waals range (distance < 5 Å) during this time segment, an energy transfer occurs radiation-free according to the Dexter mechanism[101]. In comparison to Förster resonance energy transfer (FRET) which occurs through dipole-dipole interactions over larger distances and gives information about distance distributions, TTET is a two-electron exchange process and therefore needs actual contact formation.

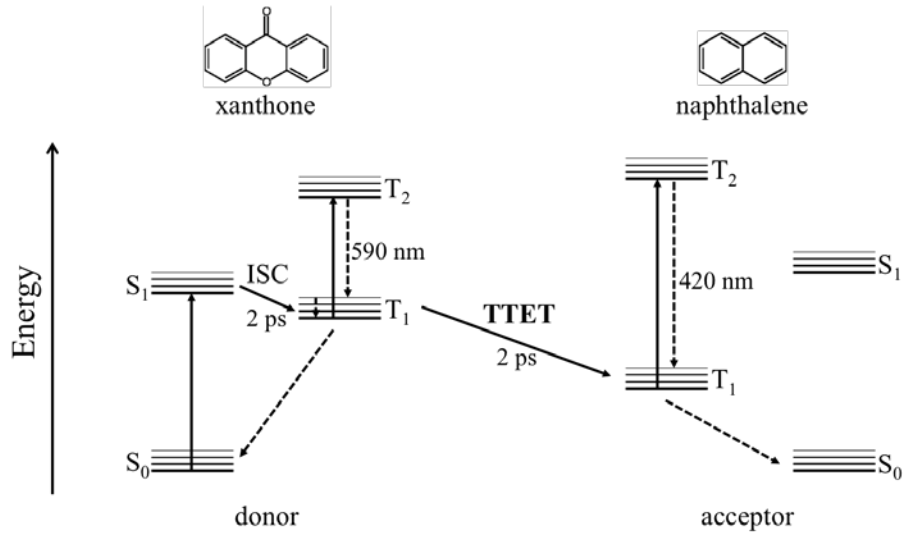


Figure 1.5 Jablonski diagram for TTET between the donor xanthone and the acceptor naphthalene. T_1 is the first electronically excited triplet state; S_0 is the singlet ground state.

The triplet transfer can be monitored spectroscopically as the triplet states of xanthone and naphthalene show intense absorption bands at 590 nm and 420 nm respectively (Figure 1.6).

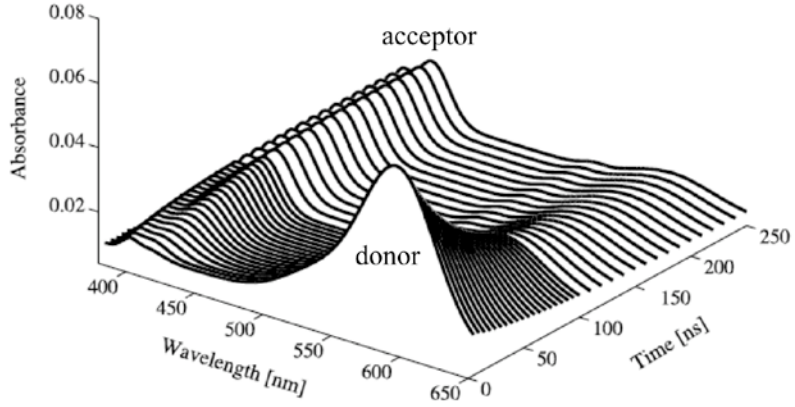


Figure 1.6 Monitoring of TTET by decay of xanthone triplet absorbance at 590 nm and concomitant increase of naphthalene triplet absorbance at 420 nm. Figure adapted from[53]

The energy transfer reaction is diffusion controlled and inversely viscosity-dependent ($k \sim 1/\eta$) with an activation energy of zero ($E_A = 0$) as shown under pseudo first order conditions with the labels free in solutions. The bimolecular transfer rate constant is 4×10^9 M^{-1} s^{-1} [53] as predicted for a diffusion-controlled reaction by Smoluchowski[102].

Since the photochemistry of TTET between xanthone and naphthalene is very fast and diffusion-controlled, it allows for measurements of absolute rate constants in the range of ~10 ps to ~30 µs. Due to the strong triplet absorbance signals, peptide concentrations of 10 – 50 µM are sufficient which also rules out contributions from intermolecular transfer reactions.

Transfer processes through the bonds of the peptide's backbone can occur over distances up to 8 bonds[103, 104]. Therefore, contributions from through-bond TTET are excluded when at least one amino acid is placed between the labels, leading to a distance of 11 bonds at minimum.

In addition to the triplet quenching effect of oxygen, some natural occurring amino acids also influence the TTET reaction. Methionine, cysteine and histidine have been shown to efficiently quench the xanthone triplet with a bimolecular rate constant close to the diffusion limit, while tryptophan and tyrosine perform TTET themselves. In order to perform TTET studies in natural peptides and proteins, it is therefore essential to find systems where these quenching amino acids are absent or can be mutated to non-quenchers without affecting the protein's stability and the folding kinetics and mechanism.

1.3.1 Loop formation in polypeptide chains measured by TTET

Loop formation dynamics in unfolded polypeptide chains have been studied extensively using TTET. Xanthone- and naphthylalanine-labeled poly-glycine-serine (GS) peptides serve as models for flexible unstructured loops. End-to-end diffusion brings the labels in close contact and leads to a transfer of the triplet state (see Figure 1.4 A), which results in single exponential absorbance decays at 590 nm. The rate constants from the decay directly represent the rate constants of loop formation.

Measuring loop formation rate constants in GS-peptides of different lengths (1 – 28 repeats) revealed stronger chain length dependence for longer peptides (number of peptide bonds N > 20) compared to shorter chains[53] (see Figure 1.7, filled circles).

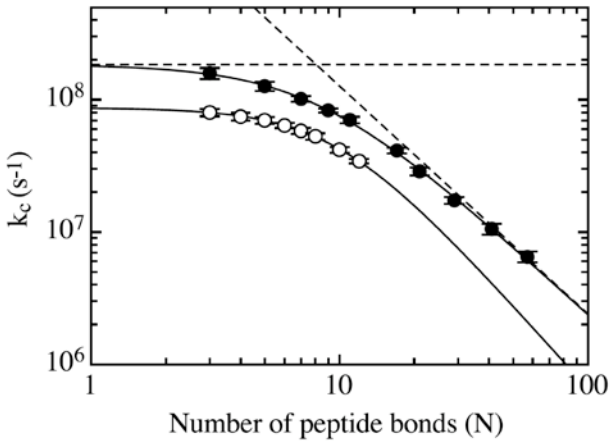


Figure 1.7 Scaling of end-to-end diffusion rate constants k_c with loop size. Filled circles represent k_c of poly-glycine-serine peptides and open circles represent k_c of poly-serine peptides. Continuous lines are fits to equation 1.21. Figure taken from[53].

The length dependence of the rate constants for loop formation can be described by

$$k_c = \frac{1}{1/k_0 + (k_i \cdot N^{\gamma})} \qquad (1.21)$$

and gives a limiting value k_0 for short peptides of 1.8×10^8 s^{-1}. The decrease of k_c with increasing chain length for longer peptides is in good agreement with the Szabo, Schulten, Schulten[15] model for end-to-end diffusion, when excluded volume effects are considered yielding a k_i-value of $6{,}7 \cdot 10^9$ s^{-1} and γ of 1.7. Temperature dependencies of GS-peptides of different lengths also revealed small enthalpic barriers which are higher for short chains (~ 14 kJ/mol for N = 3) and very low (~ 4 kJ/mol) for longer chains[22]. The results agree with polymer theory, indicating that different processes for motions limit the intrinsic dynamics of polypeptide chains over chain segments of different length. Excluded volume effects significantly influence longer chains, while short chains are dominated by chain stiffness.
In order to increase chain stiffness, poly-serine chains were measured (see Figure 1.7, open circles). Surprisingly, loop formation rate constants of short poly-serine peptides are about twofold slower ($k_0 = 8.7 \cdot 10^7$ s^{-1}) compared to GS-chains. This is a rather small effect assuming decreased flexibility and longer donor-acceptor distances. It seems that these effects are compensated by a decreased conformational space[53]. For longer poly-serine chains, a stronger effect of increasing loop size on the loop formation rate constants has been shown ($k_i = 1 \times \cdot 10^{10}$ s^{-1} and $\gamma = 2.1$).

By combination of femtosecond and nanosecond laserflash experiments with short GS- and poly-serine peptides, additional complex kinetics of loop formation on the 50- to 500-ps timescale have been observed[25]. It was suggested that these kinetics arise from conformations that can form loops barrier-free by rotation of just a few bonds.
In order to study the effect of amino acid sequence on dynamics of loop formation, host-guest peptides with the sequence Xan-Ser-Xaa-Ser-Nal-Ser-Gly were synthesized. Xaa stands for one of the canonical amino acids, Xan is xanthone and Nal is naphthylalanine. As glycine is the most flexible amino acid, it leads to loop formation times of 8 ns. Slowest dynamics were found with *trans* proline (50 ns). All other amino acids yielded loop formation times in the range of 12 to 20 ns, indicating that amino acid side chains have only a small effect on contact formation rates[53]. A small but significant difference in rate was observed for amino acids with short side chains such as alanine and serine compared to amino acids with longer side chains such as isoleucine, glutamate, arginine and histidine (see Table 1.1).

Xaa	**k_c (10^7 s^{-1})**
Gly	12 ± 1
Ala	8 ± 0.7
Ser	6.7 ± 0.7
Glu	5.4 ± 0.2
Arg	5.5 ± 0.7
His	4.9 ± 0.4
Ile	4.4 ± 0.3
trans Pro	2 ± 0.3
cis Pro	25 ± 5

Table 1.1 Effect of different amino acids on local peptide dynamics. Table taken from[53]

TTET was also used to study loop formation in natural but unstructured sequences. Observed polypeptides derived from carp parvalbumin, the B1 domain from protein G and a fragment from the DNA-binding Brinker domain. End-to-end loop formation rates were found to be slower and the activation energy higher compared to GS-peptides due to lower glycine content and the presence of larger side chains in the natural sequences[20, 22, 105, 106]. In the case of the Brinker-fragment, repulsive interactions between positively charged residues cause an additional deceleration of loop formation rates[106].

In natural occurring proteins, end-to-end loop formation events are quite uncommon. More probably is contact formation between residues within the polypeptide chain (see Figure 1.8). Comparison of the kinetics of end-to-end (type I loops), end-to-interior (type II loops) and interior-to-interior loops (type III loops) of the same size and amino acid sequence revealed that local chain motions are strongly coupled to motions of other chain segments.

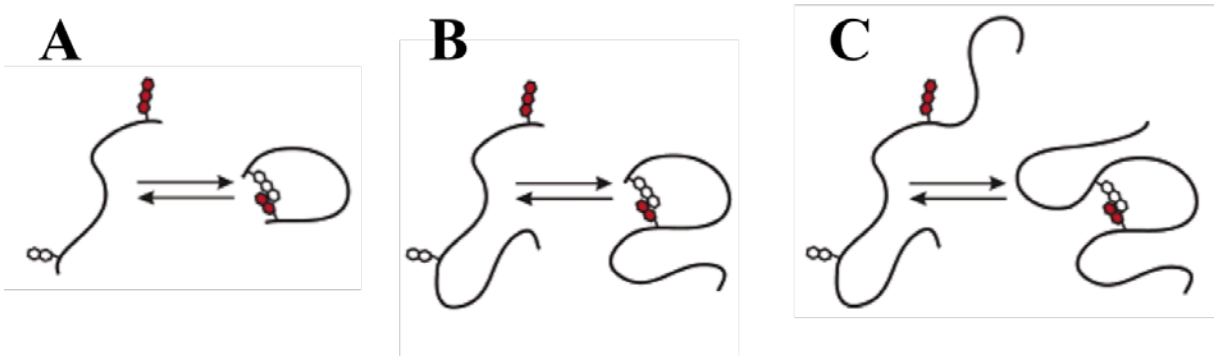


Figure 1.8 Schematic representation of loop types. (A) end-to-end loop type I, (B) end-to-interior loop type II, (C) interior-to-interior loop type III. Figure adapted from[54]

Type II loop formation is slower compared to type I loops, and type III loop formation is slower than type II loops. The strength of the effect is dependent on the size and sequence of the tails. Long and flexible tails show weaker influence on loop formation kinetics than short and stiff tails[54]. It was suggested that end extensions cause differences in internal flexibility at different segments of the chain leading to faster kinetics at the ends compared to segments within the chain.

Loop formation of polypeptide chains is strongly viscosity dependent as expected for a diffusional reaction[23]. The effect of solvent viscosity on unstructured model peptides and natural sequences has extensively been studied by TTET. All measured kinetics could be fitted by equation 1.19. In the presence of small viscogens such as ethylene glycol and glycerol, the rate constants for loop formation of long GS-peptides (N > 15) are inversely proportional to solvent viscosity (β = -1), indicating a fully diffusion-controlled reaction[22]. For shorter and less flexible model peptides as well as natural sequences, β deviates from -1[105]. The same effect was shown for larger viscogens such as polyethylene glycols, indicating that the chain does not experience the full viscous effect[106]. Using equation 1.20, it was shown that kinetic data from TTET measurements at different viscosities are best to fit by a power law with a fractional exponent β, describing the sensitivity of chain dynamics to solvent viscosity. In this case, τ_{int} as determined by extrapolation to zero solvent viscosity, appeared to be on the sub-nanosecond time scale independent of the quality of the solvent[105].

Solvent quality refers to the intramolecular hydrogen bond strength and can be varied by the addition of denaturants, such as urea and GdmCl or stabilizing co-solvents such as TMAO or sarcosine. Denaturants, yielding good solvent and favor peptide-solvent interactions, have been shown to linearly influence loop formation rate constants, describable by equation 1.22.

$$ln\ k_c = \ln k_c^0 - m \cdot \frac{[D]}{RT} \qquad (1.22)$$

Here $k_c^{\,0}$ is the rate constant for loop formation in the absence of denaturant, *[D]* is the denaturant concentration and the *m*-value is a proportionality constant. Urea and GdmCl both show a linear decrease of ln k_c with increasing *[D]*[53]. When denaturant kinetics are corrected for solvent viscosity, the linear relationship gets lost, asymptotically approaching a limiting value at high denaturant concentrations. For GdmCl *m*-values 2-fold higher than for urea were observed[23]. Stabilizing co-solvents favor intramolecular hydrogen-bonding and are therefore poor solvents. After correction for solvent viscosity, loop formation was found to be accelerated in unfolded model peptides as well as in natural sequences and reaches a limiting

value at high co-solvent concentrations[105]. The magnitude of the effect of denaturants and co-solutes is exerted on the peptide backbone, while amino acid side chains are not much affected. The effects of destabilizing and stabilizing co-solutes on loop formation kinetics can be described by the weak binding model from Schellman[107] as well as Tanford's Transfer model[6, 108].

Besides determination of loop formation rates in unstructured peptides, TTET can also be used to investigate dynamics in folded or partially folded structures (see Figure 1.4 B). Therefore, the irreversible transfer reaction must be coupled to a conformational equilibrium. The TTET labels are attached to specific sites in the chain where they cannot come in contact as long as the peptide is folded. TTET can only occur when the local structure between the labels unfolds. Applying the three-state model (equation 3.5), the microscopic rate constants for folding (k_f), unfolding (k_u) and loop formation (k_c) can be determined from the two observable rate constants.

This strategy was used to examine the dynamics and stability of model α-helical peptides. Helix formation is position independent, while unfolding is slower in the peptides center compared to the termini due to higher stability in the center. Monte Carlo simulations using a kinetic Ising model showed that the dynamics of helix formation are governed by a diffusion of boundary mechanism where the helix-coil boundary moves along the polypeptide chain[98]. The boundary diffusion can be slowed down by stabilizing capping motifs. Branched amino acids also interfere with helix formation and unfolding[109]. Information on the ground and transition states of the helix-coil transition was obtained by TTET measurements under high pressure. It could be shown that the transition state has a larger volume than either the helical or the coil state, possibly due to the presence of unsatisfied hydrogen bonds in the transition state[109].

In the villin headpiece subdomain (HP35), conformational fluctuations were investigated by TTET, revealing two native states with slightly different conformations in the C-terminal region. They are connected by an unlocking/relocking equilibrium and unfolding is only possible from the unlocked state[110]. Applying high pressure TTET measurements to the system, it could be shown that the unlocked state of HP35 has a slightly larger volume than the locked state and is therefore no dry molten globule (DMG) state as proposed by Shaknovich and Finkelstein[111]. However, the transition state for unlocking/locking has a largely expanded volume with characteristics in good agreement with the proposed DMG[112].

1.4 Turns in peptides and proteins

Besides α–helices and β–sheets, turns are the third classical secondary structures and the most observed (~ one third of the total molecule) non-repetitive structure elements in globular proteins[113]. They are commonly defined as the regions where the amino acid chain folds back on itself by nearly 180 degrees and therefore allow the protein to form a compact globular state. They are primarily located on the protein surface and likely involved in molecular recognition processes and interaction between proteins, substrates or receptors.

Depending on the number of residues that are necessary to change the overall direction of the polypeptide chain, one discriminates between δ-turns (2 residues), γ-turns (3 residues), β-turns (4 residues), α-turns (5 residues) and π-turns (6 residues)[114]. Unlike loops, which are formed by seven or more amino acids, the backbone groups of turns are in close proximity and may or may not be stabilized by intrachain hydrogen bonding. This compact folding of the backbone leads to an outward projection of the side chains with strict limitations of side-chain orientations.

The most common type of turns is the β-turn and was originally identified by Venkatachalam in 1968[115]. Based on ϕ- and ψ-angles of the residues i+1 and i+2, he proposed three distinct conformations referred to as type I, II and III along with their mirror images I', II', and III' with reversed ϕ- and ψ-values. Each could form a hydrogen bond between the backbone's C=O(i) and N-H(i+3). In 1973, examining the growing number of protein structures, Lewis et al. found that 25% of β-turns did not possess the intraturn hydrogen bond suggested by Venkatachalam[116].

	Dihedral angles (°)			
	i+1		i+2	
Turn	ϕ	ψ	ϕ	ψ
type I	-60	-30	-90	0
type I'	60	30	90	0
type II	-60	120	80	0
type II'	60	-120	-80	0
type IV	-61	10	-53	17
type VIa1	-60	120	-90	0
type VIa2	-120	120	-60	0
type VIb	-135	135	-75	160
type VIII	-60	-30	-120	120

Table 1.2 Dihedral angles of β-turn types stated by Hutchinson and Thornton 1994[117]

They introduced a more general definition of a β-turn, which states that the distance between the Cα(*i*) and the Cα(*i*+3) should not exceed 7 Å and the residues involved were not helical. Finally, in 1994, Hutchinson and Thornton stated the at present most widely used classification of β-turns with 9 distinct types based on ϕ- and ψ-ranges[117] (see Table 1.2).

The two major classes of the four-residue β-turn are types I and II along with their mirror images I' and II'. The main difference between these two types is the relative orientation of the peptide bond between residue *i*+1 and *i*+2 and accordingly the position of their side chains (see Figure 1.9).

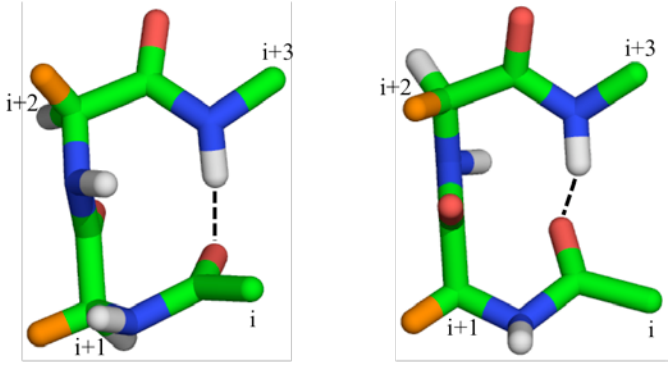


Figure 1.9 β-turns type I (left) and type II (right). Carbon atoms colored in green, side chains colored in orange. Nitrogen atoms colored in blue, oxygen in red and hydrogen in white. Dashed lines represent hydrogen bonds.

1.4.1 The role of β-turns in protein folding

The first studies on β-turns were performed using conformationally constrained cyclic model peptides. Crystal structures[118-120] and NMR studies[121-124] of cyclic peptides confirmed the original classification of β-turn types and showed close similarity in structure to turns observed in proteins. Cyclized model turns were also used to spectroscopically characterize β-turns[124-127]. Applying a self-developed algorithm[128] to the CD spectra of 14 β-turn models, Perczel and Fasman succeeded to calculate the pure component CD spectra of different types of β-turns (see Figure 1.10)[129].

Unlike α-helices and β-sheets, β-turns are non-repetitive structure elements and therefore hard to predict. As they are frequently located on the surface of globular proteins, they are primarily composed of hydrophilic amino acids. In 1978, George Rose took advantage of this knowledge and proposed a prediction method where he defined turns as local hydrophobicity minima[130]. However, this does not account for different turn types.

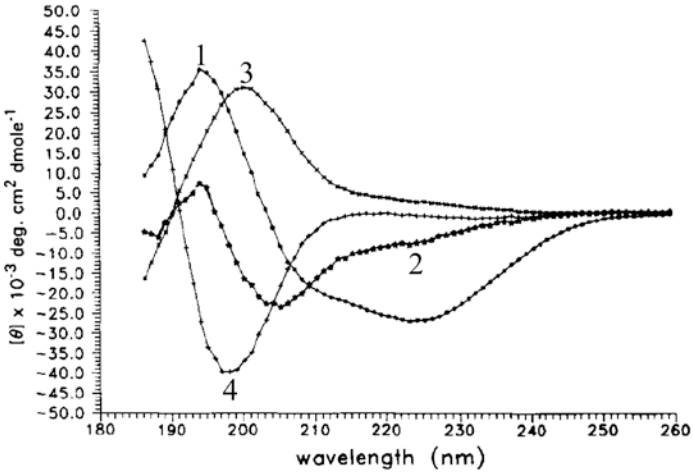


Figure 1.10 Component CD spectra of β-turns type I (1 and 2) and type II (3). 4 is interpreted as the CD spectrum of 'open conformations'. Figure taken from[129]

Over the last two decades, researchers examined residue-specific turn propensities for each turn type using statistical analysis of known protein structures[117, 131], directed evolution and phage-display experiments[132] (for type II β-turns see Table 1.3).

type II turns

Residue	*i*	*i*+1	*i*+2	*i*+3
Ile	1.01	0.64	0.00	1.10
Phe	1.23	0.74	0.37	0.80
Val	1.12	0.43	0.00	1.12
Leu	0.73	0.70	0.21	0.70
Trp	0.49	0.66	0.33	0.99
Met	0.74	0.62	0.12	1.36
Ala	1.05	1.22	0.14	1.07
Gly	0.96	0.18	**9.17**	0.57
Cys	0.13	0.00	0.26	**1.72**
Tyr	**1.59**	0.73	0.20	1.26
Pro	**1.99**	**4.91**	0.00	0.00
Thr	0.91	0.67	0.12	1.12
Ser	0.88	0.96	0.31	**1.38**
His	1.51	0.97	0.65	0.97
Glu	1.00	1.21	0.13	1.21
Asn	1.15	0.65	1.25	0.55
Gln	1.19	1.26	0.56	1.40
Asp	0.34	1.02	0.51	0.89
Lys	1.07	**1.41**	0.26	**1.58**
Arg	0.70	0.81	0.16	1.03

Table 1.3 Positional turn propensities for each amino acid at each of the 4 positions of β-turn Type II. Values taken from[117]

Introducing residues with high turn propensity can also influence whole protein stability as demonstrated by Trevino et al.. They could increase the conformational stability of RNase Sa by 0.7 to 1.3 kcal/mol by optimizing β-turn sequence[133].

Today a bulk of β-turn prediction programs is freely available as webservers on the Internet[134-138]. The most accurate ones use machine-learning neural networks[139] and achieve Matthews correlation coefficients[140] of up to 0.5.

The next hierarchical build-up of secondary structure from a β-turn is the β-hairpin. It is formed by connecting two β-strands in antiparallel direction by loops of various lengths (see Figure 1.11). The most abundant connecting region is a β-turn with residues *i*+1 and *i*+2 forming the loop and *i* and *i*+3 residues sitting at the ends of the β-strands. A recent study showed that from a set of 3977 β-hairpins with a connecting β-turn region, 48% are of type I', 24% are type II', 16% type I and only 6% type II. However the β-strands do not show any specific sequence and structural patterns for β-turn types[141].

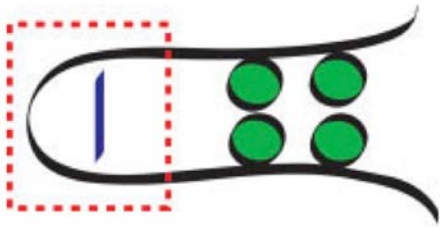


Figure 1.11 Schematic showing of a β-hairpin. The red box represents the turn region. Green circles represent interactions in the flanking regions. Figure taken from[142]

β-hairpins are stabilized by the turn propensities of residues in the loop region as well as cross-strand interactions in between the flanking β-sheets such as aromatic-aromatic, aromatic-polar or hydrophobic interactions, hydrogen bonds or salt bridges[143-146].

Two hypotheses exist for β-hairpin folding. They can either follow the turn zipper mechanism, where folding is initiated in the turn region and zipped from turn to the end[147], or the hydrophobic collapse mechanism. Here, a hydrophobic collapse occurs prior to the formation of hydrogen bonds and the hairpin[148].

Until today, it is controversially discussed whether β-turns play an active or passive role in protein folding. From an active point of view, turns may serve as nucleation sites for protein folding and are formed early in the folding process[149, 150]. However, turn formation might also be a consequence of the association of nonlocal interactions in the flanking strands. Here, turns can only form after other regions have developed[151-153].

In 1988, Dyson et al. published a study that supports the activists view. They measured the turn propensity of the model peptide Tyr-Pro-X-Asp-Val (with X = any amino acid) by NMR

chemical shift. All *trans*-Pro peptides formed ordered β-turn conformations in solution even at high temperatures (360 K), with the highest population for X = Gly[154]. Recent MD simulations of the same peptide showed that turn formation seems to be driven by intra-peptide hydrogen bonding interactions[155].

Several site-directed mutagenesis studies of different proteins and protein domains containing one or more β-hairpins hypothesize that there are turns that form early and turns that form late[156-161]. Interestingly, most of the early forming turns have high-predicted turn propensity and are highly conserved[142].

However this is not true for all cases. The I-band domain 27 from Titin has a β-turn of high propensity and conservation that does not nucleate protein folding[162].

In order to understand the mechanisms of β-turn folding and its contribution to protein folding, more studies on excised model peptides as well as whole proteins, using ever more powerful methods, are necessary.

1.4.2 Carp β-parvalbumin as an appropriate model to study protein folding by TTET

Parvalbumins represent a family of homologous, intracellular Ca^{2+}-binding proteins of low molecular weight. They were first reported to occur in the skeletal muscle fibers of fish and amphibians[163], but are also found in lower amounts in the fast-twitch muscles of higher vertebrates[164, 165]. Here, they interact with the sarcoplasmic reticulum and function in calcium buffering involved in the relaxation process of muscles[166-169].

Parvalbumins from fish were among the first identified allergen molecules and are the major fish allergens[170], constituting remarkable resistance to heat, denaturants and proteolytic enzymes when Ca^{2+} is bound[171]. The loss of Ca^{2+} leads to a change in conformation and decreased allergenicity[172, 173].

Based on the amino acid sequence, parvalbumins can be classified into two evolutionary lineages. The α-group representing proteins with 109 amino acids are less acidic with a pI of 5.0 or higher; the more acidic β-forms consisting of 108 amino acids with a pI of 4.5 or lower[174]. The first structure of parvalbumin determined by Kretsinger and Nockolds in 1973 was that of the β-form of carp (*cyprinus carpio*). It revealed three typical helix-loop-helix domains (EF-hands)[175]. Two of these EF-hands, the CD- and the EF-motif, are capable of binding Ca^{2+} with dissociation constants of 10^{-7} – 10^{-9} M, while the AB-motif lost this ability[176] (see Figure 1.12).

Due to the two high affinity Ca^{2+} binding sites Parvalbumin is frequently used as a simple model for Ca^{2+} binding proteins. Hereby, a major point of interest is the investigation of interactions between calcium binding sites[177] as well as the effects of the intrinsic calcium binding ability on other proteins, peptides or membranes with physiological significance as reviewed in[178].

Removal of the two Ca^{2+}-ions by chelation with ethylene glycol tetraacetic acid (EGTA) leads to reduced helical content and the population of a partially folded intermediate state, which makes parvalbumin an attractive model protein for the investigation of the properties and structure of intermediate molten globule-like states[179].

The primary sequence of carp β-parvalbumin (PV) is rich in alanine, aspartic acid, glutamic acid and phenylalanine. It possesses only one of the xanthone triplet-quenching amino acids histidine and cysteine, respectively. Methionine, proline, tyrosine and tryptophan are lacking completely[180], which makes it a suitable model to study protein folding using TTET.

Figure 1.12 Structure of carp β-parvalbumin (PV). The helix-loop-helix motif AB is shown in green, CD in blue and EF in red. Calcium ions are represented as black spheres. Figure was prepared using MacPyMOL and the protein databank (pdb) file 4CPV[181, 182].

Different fragments of PV were used to study the dynamics in unfolded natural sequences via TTET[20, 22, 105, 106, 183]. PV contains phenylalanines that flank each loop region of the EF- and CD-motifs and point to the hydrophobic core of the protein. The EF-loop (PV residues 85-102) and the DE-loop (PV residues 66-85) were synthesized via solid-phase peptide synthesis (SPPS) and the respective flanking phenylalanines were replaced by the TTET labels. Single-exponential kinetics were observed with time constants for loop formation of 50 ns for the

EF-loop and 70 ns for the DE-loop. The faster kinetics in the EF-loop are due to its high glycine content.

TTET measurements in PV were also done in order to determine the effect of long range interactions on the dynamics of the E-helix[105]. Therefore, loop formation kinetics in PV fragments were compared to those in full length PV, in the unfolded state as well as the Ca^{2+}-free intermediate state. In the intermediate state, loop formation is slowed down due to long-range interactions, whereas in the unfolded state, long range interactions are not present and do not influence loop formation kinetics.

SPPS is limited to a polypeptide length of 70 – 100 amino acids. Therefore, the full-length protein containing the TTET labels for these measurements was produced using native chemical ligation. An expressed fragment was condensed with a synthesized fragment bearing the triplet donor xanthone. The triplet acceptor was introduced by using a thiol reactive mechanism to couple a naphthyl moiety to the side chain of cysteine. However, this method is extensive and time-consuming. Therefore, in order to perform further TTET measurements on full-length PV and other suitable proteins, a new strategy to site-specifically introduce TTET labels in bigger proteins is required.

1.4.3 Site-specific modification of proteins via incorporation of unnatural amino acids and click chemistry

Since Cricks 'central dogma of molecular biology' stated in the 1950s it is known that the information about the amino acid sequence of proteins is provided in the DNA in form of the four bases adenine (A), cytosine (C), guanine (G) and thymine (T). After transcription, where T is exchanged by uracil (U) the ribosome translates the messenger RNAs into proteins[184]. Thereby, a triplet base sequence (codon) corresponds to one amino acid, which ultimately leads to 64 possible coding units with only 61 encoding for the 20 canonical amino acids and three stop codons, that are used to terminate the protein expression[185] (see Figure 1.13).

In order to gain more functionalization, proteins *in vivo* are often posttranslationally modified. By enzyme-catalysation, the canonical amino acids can be phosphorylated, glycosylated, acetylated, oxidatively modified and even brominated[186]. Additionally, two non-canonical amino acids, selenocysteine and pyrrolysine, are found to be incorporated into proteins co-translationally, regarded as natural expansion of the genetic code.

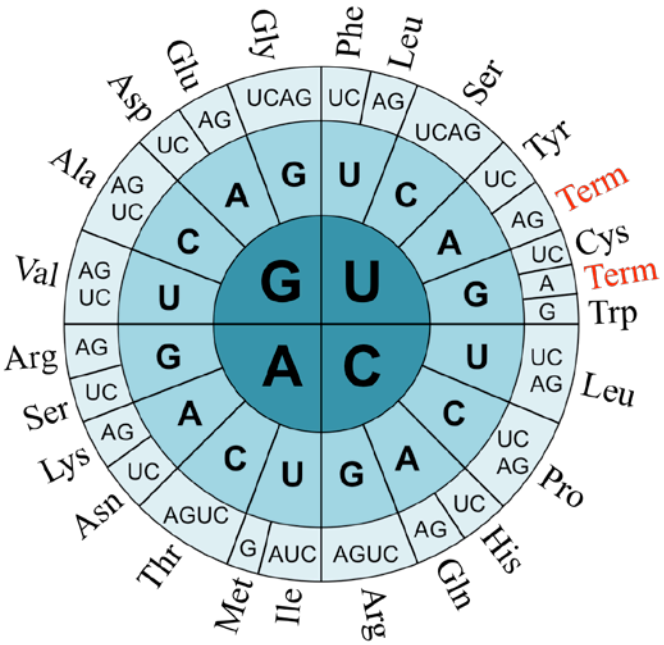


Figure 1.13 Presentation of the degenerated genetic code in RNA format. Inner circle: 5'-nucleobase of the triplet, middle circle: middle nucleobase of the triplet and outer circle: 3'-nucleobase of the triplet. Stop codons denoted as Term.

In research, site-specific modification of proteins is a powerful tool for investigating protein structure and function, as well as generating proteins with new or enhanced properties. A particularly attractive approach therefore is the incorporation of unnatural amino acids (UAAs), as it allows for recombinant expression of site-specifically modified proteins in high yields in bacteria, yeast or mammalian cells with only minimal effects on the structure of the wild-type protein.

An early method for the genetic incorporation of UAAs into proteins is the selective pressure incorporation (SPI). A bacterial strain, auxotrophic for a structurally related canonical amino acid, is used and the cells are grown in an environment containing limited concentrations of the amino acid to be substituted. UAAs that are structurally similar to the limited amino acid are recognized by that wild-type aminoacyl-tRNA synthetase (aaRS) when added in excess. Subsequently, the UAAs will be incorporated at every position of the cognate amino acid not only in the recombinantly expressed protein, but in the whole bacterial proteome[187]. SPI has been applied to analogues of Pro, Tyr, Phe, Leu and Val but is most suited for analogues of Met (see Figure 1.14), as Met has low abundance in the proteasome and hence less toxicity when replaced globally[188].

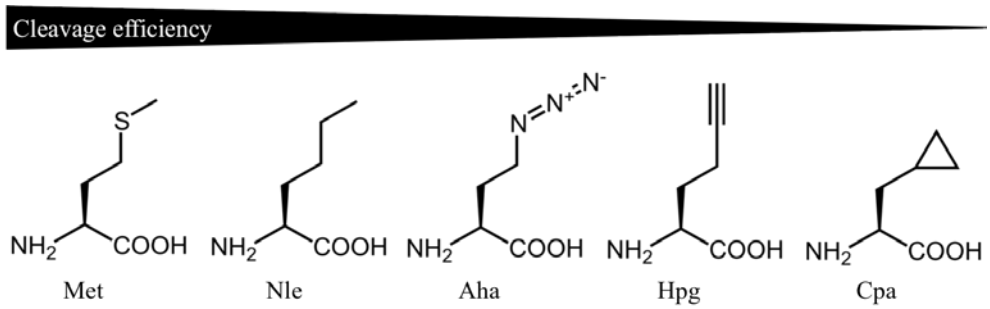


Figure 1.14 Chemical structures of Met analogues. Nle: norleucine, Aha: azidohomoalanine, Hpg: homopropargylglycine, Cpa: cyclopropylalanine

Met is always the first amino acid incorporated in the N-terminal position of a protein, but is in most cases co-translationally cleaved off by the Met aminopeptidase (MetAP). Whether or not the initial Met is removed depends on the second amino acid. Bulky residues on position 2, such as Arg or Lys, block N-terminal residue excision, whereas small amino acids like Ala or Gly facilitate the process. Met analogues incorporated in the N-terminal position show varying cleavage efficiencies. Norleucine (Nle) excision is almost as efficient as Met, azidohomoalanine (Aha) is less efficiently excised, whereas homopropargylglycine (Hpg) and cyclopropylalanine (Cpa) are unfavorable substrates for MetAP[189].

The SPI method is restricted to close analogues of the cognate amino acid. Furthermore, substitution at multiple sites throughout the protein and the proteome can limit the utility of this technique. Also, it is not possible to incorporate several different UAAs into the same protein.

In order to overcome these limitations, the lab of P.G. Schultz invented another technique by expanding the genetic code beyond the canonical amino acids[190]. They reprogrammed the cell's translational machinery by directed evolution of tRNA/aaRS pairs and achieved to incorporate UAAs in nonsense codons[191]. It is essential that the tRNAs and the aaRSs are orthogonal to their counterparts in the host cell, which means that they do not cross-react (see Figure 1.15).

Until today, 70 UAAs were successfully incorporated in bacteria, yeast and mammalian cells by evolving orthogonal tRNA/aaRS pairs derived from leucyl[192], tryptophanyl[193], tyrosyl[194], lysyl[195], pyrrolysyl[196] and prolyl[197] pairs. These orthogonal tRNA/aaRS pairs recognize all three stop codons (amber codon UAG, ochre codon UAA and opal codon UGA) as well as 4-residue frameshift codons[195]. Furthermore, combinatory incorporation of two or more different UAAs in the same protein was achieved by either combination of the SPI method and stop codon suppression[198] or simultaneous suppression of different codons in one protein[199].

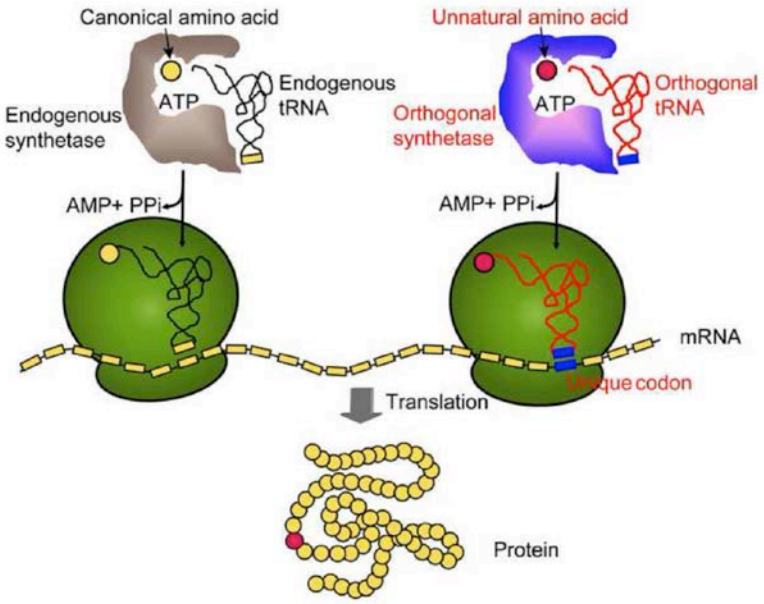


Figure 1.15 Scheme of the incorporation of an unnatural amino acid via an orthogonal tRNA/amino acyl synthetase pair. Figure taken from[200]

A matter of particular interest is the incorporation of non-canonical, chemically reactive amino acids that can be used to selectively conjugate the protein to other small molecules. Suitable chemical reactions take place under physiological conditions, are highly regio-specific and orthogonal to the biochemistry in living cells. The most widely used reactions include the Staudinger ligation[201], inverse electron demand hetero Diels-Alder reactions[202] and Cu^{+}-mediated Huisgen 1,3-dipolar cycloadditions (CuAAC or click chemistry)[203]. In click chemistry, an azide and an alkyne form, in presence of Cu(I), a 1,2,3-triazole ring, which is stable under physiological conditions (see Figure 1.16).

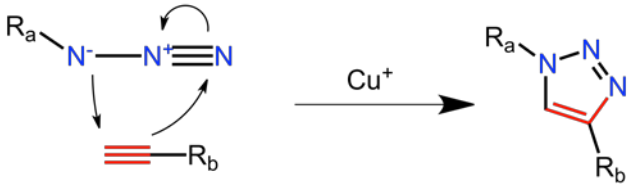


Figure 1.16 Copper(I)-catalyzed azide-alkyne Huisgen cycloaddition (CuAAC).

The Huisgen cycloaddition gained great importance in chemical posttranslational modifications due to the easy incorporation *in vivo* of azide- and alkyne-moiety bearing UAAs such as Aha and Hpg. To date, the CuAAC-reaction has been successfully adopted for a tremendous number of most diverse biological, biochemical and chemical fields, as reviewed in[204] and is an important aid in site-specific modifications of proteins.

2. Aim of Research

The aim of this thesis was to answer fundamental questions concerning dynamics in the unfolded state of proteins as well as during protein folding. Herefore, the method of triplet triplet energy transfer (TTET) was used. TTET yields absolute rate constants for site-specific contact formation within a polypeptide chain on the nanosecond to microsecond timescale. Coupled to a conformational equilibrium, it gives information on conformational dynamics in proteins and peptides.

Intrachain diffusion in unfolded polypeptide chains

The unfolded state is the starting point for protein folding. Therefore, the characterization of dynamics in the unfolded state is of particular importance as they determine the speed of the conformational search for favorable interactions. In earlier experiments, TTET was used to study the kinetics of end-to-end loop formation in poly(glycine-serine) and polyserine model peptides as well as in natural sequences derived from unstructured loops or intrinsically disordered proteins. By analyzing the effect of solvent viscosity and temperature, it was possible to assess contributions from solvent interactions and to characterize the barriers for end-to-end loop formation kinetics. It was shown that end-to-end loop formation in long, very flexible poly(glycine-serine) model chains is inversely proportional to solvent viscosity end exhibits activation energies on the range of RT, which suggests fully diffusion-controlled reaction behavior. End-to-end loop formation in shorter model loops and natural sequences show fractional viscosity dependencies and significant activation energies due to increased chain stiffness. In this work, we aimed to determine the effect of increased chain length and stiffness to the rate constants of end-to-end loop formation by measuring TTET in three other natural peptide fragments, derived from the intrinsically disordered protein AT phd. Furthermore, we wanted to characterize the activation energies and the viscosity dependencies of these peptides.

In natural occurring proteins, end-to-end loop formation events are quite uncommon. A more relevant case is interior contact formation between residues within the polypeptide chains. Previous NMR experiments showed that additional tails lead to a decreased flexibility in the chain center. TTET measurements of model homopolypeptide chains showed decreased loop formation rate constants for interior loop formation compared to end-to-end contact formation, which is in accordance with this finding. The aim of this thesis was to investigate the influence of additional tails on the rate constants, activation energies and viscosity

dependencies of loop formation in different unfolded polypeptides derived from natural proteins.

It was suggested that intramolecular peptide motions are subject to two additive frictional forces. The major source of friction derives from interactions with the solvent. However, internal friction effects like intra-chain interactions or steric effects from the amino acid side chains might be taken into account as well. The determination of loop formation time constants at different solvent viscosities allows to test for additive contributions of internal friction to solvent friction through an extrapolation to zero solvent viscosity. We aimed to examine internal and solvent friction in three peptide fragments derived from the intrinsically disordered protein AT phd. In addition, we wanted to investigate the effect of end extensions on frictional forces.

Formation of turns in model polypeptides

Besides α-helices and β-sheets, turns are a third classical secondary structures. It is controversially discussed whether they play an active role in protein folding, serving as nucleation sites, or if they only arise as a consequence of the association of nonlocal interactions in the flanking strands. We wanted to design small β-turn model peptides and measure TTET and far-UV CD to see whether they can fold independently.

In addition, model β-turns type II are well suited to test for the existence of a proposed disfavored bridge region in the Ramachandran plot. The Ramachandran plot visualizes backbone dihedral angles ψ against ϕ of amino acid residues in protein structure. It is controversially discussed whether hydrogen-bonding requirement should be applied as an additional energetic criterion for the Ramachandran plot. If so, residues adjacent to residues situated in the disfavored bridge could only be satisfied by intramolecular H-bonds, which would eliminate a major fraction of the accessible space in the bridge region. In good solvents like high concentrations of urea or GdmCl, which favor the unfolded state, intramolecular hydrogen bonds are disfavored and thus this region should be underpopulated. The *i*+2-glycine residue in a β-turn type II with the sequence Xaa-Pro-Gly-Yaa lies in the disfavored bridge region. By performing TTET and CD measurements in water and in high concentrations of denaturant we wanted to test this theory.

Dynamics in the unfolded state of a full-length protein

In earlier experiments, insights into the dynamics of the unfolded state were obtained by applying TTET measurements to short model peptides or fragments from natural proteins.

However, the native structure in full-length proteins is mainly stabilized by long-range interactions. In order to understand whether some long-range interactions already exist in unfolded or partially folded proteins, it is of great importance to perform TTET measurements in full-length proteins. In previous work, the TTET chromophores were introduced into the amino acid sequence via solid phase peptide synthesis (SPPS). However, SPPS is restricted to proteins of 70 to 100 amino acid length. Therefore, we wanted to establish a method to site-specifically introduce TTET labels during the expression of proteins of higher molecular mass, using unnatural amino acids and bioorthogonal chemistry. We planed to apply this method to test for the influence of long-range interactions on the dynamics of unfolded or partially folded α-helices in carp β-parvalbumin (PV), which is folded in the presence of calcium and is unfolded with residual helical structure in the absence of calcium. Therefore, we planed to compare the results from TTET measurements in an isolated helical fragment and in the full-length PV.

3. Material and Methods

3.1 Used materials

All solvents and chemicals used in this work were purchased from Carl Roth (Karlsruhe, Germany), Merck (Darmstadt, Germany), Sigma-Aldrich (St. Louis, MO, USA) or VWR International, LLC (West Chester, PA, USA) if not stated otherwise. Urea ultrapure was purchased from Gerbu Biotechnik GmbH (Heidelberg, Germany) and Guanidinium chloride (GdmCl) AA grade from Nigu Chemie GmbH (Waldkrainburg, Germany). Fmoc-protected amino acids were purchased from Merck Novabiochem (Darmstadt, Germany) or Iris biotech (Marktredwitz, Germany). Primers were purchased from Biospring (Frankfurt am Main, Germany) and DNA purification kits from Qiagen (Hilden, Germany).

3.2 Solid-phase peptide synthesis (SPPS)

All peptides used in this study were synthesized using standard fluorenylmethoxycarbonyl (Fmoc) chemistry. Synthesis was performed either on an Applied Biosystems 433A peptide synthesizer (Foster City, CA, USA) or an Intavis MultiPep CF Synthesizer (Cologne, Germany) in 0.1mmol scale using a TentaGel R RAM resin from RAPP Polymere (Tübingen, Germany). 1mmol or 0.5mmol of Fmoc-protected α-amino acid was activated either with O-(benzotriazol-1-yl)-N,N,N’,N’-tetramethyluronium hexafluorophosphate (HBTU) or O-(7-azabenzotriazol-1-yl)-N,N,N’,N’-tetramethyluroniumhexafluorophosphate (HATU) in dimethylformamid (DMF). The coupling agent was 2M N,N-diisopropylethylamine (DIPEA). Deprotection was done in 20% (v/v) piperidine. The Fmoc deprotection was controlled by UV feedback monitoring at 301nm and deprotection was repeated until the absorbance was less than 5% of the initial value. An additional capping step with 5% acetic anhydride (Ac_2O) after coupling was performed on the Intavis MultiPep CF synthesizer.

To prevent aspartimide formation for sequences containing an aspartate-glycine motif, a 2,4-dimethoxybenzyl protected aspartate-glycine building block (Fmoc-Asp(OtBu)-(Dmb)Gly-OH) from Merck Novabiochem (Darmstadt, Germany) was used. Aggregation during synthesis was prevented using pseudoproline dipeptides from Novabiochem.

Peptide	double building blocks
poly(Gly-Ser)	Fmoc-Gly-Ser($\psi^{Me,Mepro}$)-OH
DE-Loop N/C+6	Fmoc-Leu-Thr($\psi^{Me,Mepro}$)-OH Fmoc-Lys(Boc)-Thr($\psi^{Me,Mepro}$)-OH
EF-Loop N/C+6	Fmoc-Lys(Boc)-Thr($\psi^{Me,Mepro}$)-OH Fmoc-Asp(OtBu)-(Dmb)Gly-OH Fmoc-Ile-(Dmb)Gly-OH
AT phd 3	Fmoc-Ser(tBu)-Thr($\psi^{Me,Mepro}$)-OH
AT phd 3 N/C+6	Fmoc-Ala-Thr($\psi^{Me,Mepro}$)-OH Fmoc-Ser(tBu)-Thr($\psi^{Me,Mepro}$)-OH

Table 3.1 double building blocks for SPPS

3.3 Peptide modification

3.3.1 Introduction of chromophores for triplet-triplet energy transfer (TTET) into peptides

The triplet acceptor naphthalene was introduced via the unnatural Fmoc protected amino acid 1-(L)-naphthylalanine (BACHEM, Bubendorf, Switzerland) during SPPS.

9-oxoxanthene-2-carboxylic acid (xanthonic acid), which serves as triplet donor, was synthesized as previously described[205] and either coupled to the N-terminus or to the selectively deprotected side chain of diaminopropionic acid (Dpr). The N-methyltrityl (Mtt) protection group of Dpr was removed with 2% (v/v) trifluoroacetic acid (TFA), 2% (v/v) triethylsilane (TES) in dichloromethane (DCM) for 10 minutes. After washing of the resin with DCM, the procedure was repeated 5 times. Xanthonic acid was coupled in threefold excess by HATU activation in the presence of 5 equivalents of N-methylmorpholine (NMM) in DMF for 30 minutes.

For click-peptides xanthone was attached to the side chain of Azidohomoalanine (Aha) by copper-mediated azide/alkyne cycloaddition[203]. For this, a xanthone-alkyne derivative was synthesized by the group of Marcel Mayor (University of Basel). Peptide and label were dissolved in DMSO in a ratio of 1:1, a modicum of $CuSO_4$ and ascorbic acid was added and incubated for 1 hour.

Success of the modification reactions were tested with high-performance liquid chromatography (HPLC) and either Electrospray Ionization (ESI-) or matrix-assisted laser desorption/ionization-time of flight (MALDI-TOF) mass spectrometry.

3.3.2 Acetylation of the peptide amino-terminus

10% (v/v) Ac_2O, 10% (v/v) DIPEA in DMF were added to the resin-bound peptide and shaken for 10 minutes. After washing with DMF, the procedure was repeated 3 times.

3.3.3 Peptide cleavage

The peptide was cleaved from the resin and simultaneously deprotected by 4.8% (v/v) TES, 4.8% (v/v) H_2O, 1.4% (v/v) triisopropylsilane (TIPS) treatment in TFA for 2.5 hours. The cleavage product was precipitated in ice-cold methyl tert-butyl ether, centrifuged and the pellet was lyophilized. As β-turn peptides did not precipitate in ether, they were directly purified from the cleavage solution via HPLC.

3.4 Peptide purification

20 mg of the lyophilized raw product was dissolved in 1 ml TFA and diluted with an acetonitrile/water mixture in accordance with the HPLC starting conditions to 5 ml. HPLC was performed on a 1200 series from Agilent Technologies (Santa Clara, CA, USA) in an acetonitrile/water gradient with 0.1% TFA. Used columns were: Merck RP-8 (LiChrosopher 100, 250 x 25.0mm, 4µm), Phenomenex Jupiter Proteo (C12, 150 x 30 mm, 10 µm) and Phenomenex Kinetex (XB-C18, 250 x 21 mm, 5 µm). Peak fractions were collected and the purity of the fractions checked by analytical HPLC on a Phenomenex Jupiter Proteo (C12, 150 x 4.6 µm, 4 µm) column. Pure fractions were pooled, lyophilized and stored for further usage at -20°C.

3.5 Sample preparation

All measurements, except measurements on Brinker N/C+6, were done in 10mM potassium phosphate buffer, pH 7. For Brinker N/C+6, a 20mM NaCl, 20mM Ac solution pH 5.7 was used. All buffers and samples were filtered with 0.2 µm pore size filters prior to usage.

Peptide concentrations were determined on an Agilent 8453 UV-visible Spectroscopy System (Santa Clara, CA, USA) using an extinction coefficient of 3900 cm^{-1} M^{-1} at 343nm for peptides labeled with Xan. For determination of the concentration of peptides containing tyrosine residues, an extinction coefficient of 1400 cm^{-1} M^{-1} per tyrosine at 280 nm was used. For peptides containing tryptophan residues, an extinction coefficient of 5500 cm^{-1} M^{-1} per tryptophan at 280 nm was used.

Urea and GdmCl concentrations were calculated from refractive indices in the presence (η) and in the absence (η_0) of denaturant measured on an automatic refractometer AR7 Series of Reichert (Depew, NY, USA) after Pace et al.[206].

$$[urea] = 117.66 \cdot (\eta - \eta_0) + 29.753 \cdot (\eta - \eta_0)^2 + 185.56 \cdot (\eta - \eta_0)^3 \qquad (3.1)$$

$$[GdmCl] = 57.147 \cdot (\eta - \eta_0) + 38.68 \cdot (\eta - \eta_0)^2 + 91.6 \cdot (\eta - \eta_0)^3 \qquad (3.2)$$

Solvent viscosities (η) were measured with a HAAKE falling ball viscometer Type C from Thermo Scientific (Waltham, MA, USA). The temperature of the solution was adjusted with a F20, HC/7 water bath from Julabo Labortechnik GmbH (Seelbach, Germany). Sample viscosities were calculated according to equation 3.3

$$\eta = t \cdot (\rho_1 - \rho_2) \cdot K \qquad (3.3)$$

where t specifies the time (s) needed for the ball to traverse a defined distance, ρ_1 and ρ_2 are the density (g/cm^3) of the ball and the solvent and K (mPa cm^3/g) is a ball specific constant.
Later measurements were performed with a Rolling-ball viscometer Lovis 2000 M/ME from Anton Paar (Graz, Austria) with implemented Peltier thermostat and density meter.

3.6 Spectroscopic measurements

3.6.1 Laserflash photolysis

TTET measurements were performed on a commercial Laser Flash Photolysis Reaction Analyzer (LKS.60) from Applied Photophysics (Surrey, UK). Xanthone was excited to the triplet state by a 4 ns laser pulse at 355 nm using a Nd:Yag Brilliant laser from Quantel (Les Ulis, France).
In GS click-peptides, xanthone was excited by a 120 ps laser flash, which allows for the determination of rate constants in the pico- to nanosecond range. In order to separate TTET reactions from photophysics in xanthone, the absorbance changes in the donor-acceptor peptide was normalized against the changes in a donor-only reference peptide.
Relaxation of the triplet state of xanthone and the concomitant formation of the triplet state of naphthalene was followed by the change in absorption at 590 nm and 420 nm, respectively. Absorbance was recorded on an Agilent infiniium oscilloscope with 600 MHz and 4 GSa/s

over at least 5 half-life periods on a logarithmic timebase. Peptide concentrations were between 20 – 70 µM and five to twenty traces were recorded for each measurement and averaged. As a reference, 50 µM xanthonic acid was measured every day and the peptide traces were normalized in relation to the amplitude and analyzed using ProFit (Quantumsoft, Zürich, Switzerland).

3.6.2 TTET data evaluation with the three-state model

In the helical peptides AT phd 3 N/C+6 and PV A-helix, the three-state model for local helix-coil dynamics coupled to TTET can be described by

$$H \underset{k_f}{\overset{k_u}{\rightleftharpoons}} C \xrightarrow{k_c} C^* \qquad (3.4)$$

resulting in two observable rate constants (λ_1 and λ_2) and their corresponding amplitudes (A_1 and A_2). Viscosity and temperature dependencies were fitted globally using the following equations[98]

$$\lambda_{1/2} = \frac{k_u + k_f + k_c \pm \sqrt{(k_u + k_f + k_c)^2 - 4k_u k_c}}{2} \qquad (3.5a)$$

$$A_1 = \frac{1}{\lambda_1(\lambda_1 - \lambda_2)}([C]_0 \cdot k_c \cdot (k_u - \lambda_1) + [H]_0 \cdot k_u \cdot k_c) \qquad (3.5b)$$

$$A_2 = \frac{1}{\lambda_2(\lambda_1 - \lambda_2)}([C]_0 \cdot k_c \cdot (\lambda_2 - k_u) + [H]_0 \cdot k_u \cdot k_c) \qquad (3.5c)$$

3.6.3 Circular dichroism spectroscopy

All CD measurements were performed on an AVIV 410 spectropolarimeter (Lakewood, NJ, USA). Fractional helix contents were calculated after Luo and Baldwin[207] with

$$f_H = \frac{[\theta]_{222nm} - [\theta]_{0\%}}{[\theta]_{100\%} - [\theta]_{0\%}} \qquad (3.6)$$

100% and 0% helix content are given by:

$$[\theta]_{100\%} = (-44000 + 250 \cdot T) \cdot \left(1 - \frac{3}{N}\right) deg \cdot cm^{-2} \cdot dmol^{-1} \quad (3.7)$$

$$[\theta]_{0\%} = -2220 - 53 \cdot T \; deg \cdot cm^{-2} \cdot dmol^{-1} \quad (3.8)$$

$[\Theta]_{222\,nm}$ is the observed mean residual ellipticity at 222 nm, T is the temperature in ºC and N is the chain length in residues.

3.6.4 NMR spectroscopy

NMR spectra were recorded on NMR spectrometers with ^{1}H resonance frequencies of 500, 600 and 750 MHz (Bruker, Karlsruhe, Germany) with inverse TXI (^{1}H, ^{13}C, ^{15}N) probes (600, 750 MHz) or inverse TCI (^{1}H, ^{13}C, ^{15}N) cryoprobe (500 MHz), all equipped with z gradient coils.

For all compounds, spectra were recorded at temperatures of 288 K, 298 K and 308 K; for some compounds additional spectra were run at 293 K and 303 K. The temperatures were calibrated on a methanol-d_4 sample[208].

1D spectra were recorded with a W5 Watergate sequence (Bruker pulse program *zggpw5*)[209], 2D TOCSY spectra (Bruker pulse program *mlevgpphw5*, ca. 10 kHz spinlock field)[209, 210] with long mixing time (60 or 75 ms; for sidechain assignment / amino acid type identification) or shorter mixing time (12, 20 or 40 ms, for resolution of overlapping H_N signals). Where sample concentration was sufficient, additional 2D ^{1}H,^{13}C-HSQC spectra were acquired for amino acid type verification (Bruker pulse program *hsqcetgp*)[211].

Temperature coefficients for H_N signals were calculated from chemical shifts from 1D or 2D signals; all signals showed a linear dependence within the observed temperature range.

$^{3}J_{HN,Ha}$ coupling constants were directly measured from high-resolution 1D NMR spectra after appropriate Lorentz-to-Gauss apodization to reduce linewidths.

3.7 Computational methods

3.7.1 Calculation of the solvent accessible surface area (SASA)

SASAs of unfolded peptides were calculated using the mean accessible surface area (ASA)-values for the individual amino acids obtained from[212] with a “hydrogen bond dial” of 1.5 kcal/mol per hydrogen bond for the backbone and the side chain. Xanthone had a constant

maximum ASA of a N-terminal residue and naphthylalanine was set equal to tryptophan. Acetylated N-termini and amidated C-termini were included.

3.8 Molecular biology methods

3.8.1 Site-directed mutagenesis

Carp-β-parvalbumin (PV) variants N7-Aha and N7-Aha-E16-Nal were obtained by site-directed mutagenesis of the double PV mutant C18S/H26F (PV pseudo WT) in the vector pET 11d provided by C. Nyffenegger[105]. Mutagenesis was performed using the QuickChange® Site-Directed Mutagenesis Kit from Stratagene (La Jolla, CA, USA). The Primer sequences are given in Table 3.2.

PV variant	Direction	Primer sequence
PV N7-DO	5'➔3'	cgc tgg tgt tct gat gga cgc tga cat cgc tgc
	3'➔5'	gca gcg atg tca gcg tcc atc aga aca cca gcg
PV N7-E16	5'➔3'	gac atc gct gct gct ctg tag gca agc aaa gcc gc
	3'➔5'	gcg gct ttg ctt gcc tac aga gca gca gcg atg tc

Table 3.2 Primer sequences used for site-directed mutagenesis.

Polymerase chain reaction (PCR)[213] was performed on an Advanced Primus 25 PCR thermocycler from PEQLAB (Erlangen, Germany) using a Pfu Turbo Polymerase supplemented in the kit. The temperature program is given in Table 3.3.

Temperature	Time	Cycles
95 ºC	30 s	1x
95 ºC 55 ºC 68 ºC	30 s 1 min 6 min	16x
4 ºC	24 h	1x

Table 3.3 PCR Temperature program

Parental strands were digested with 1 µl *DpnI* for 1 hour at 37 ºC. Supercompetent XL1-Blue *E.coli* cells (genotype: endA1 gyrA96(nal^R) thi-1 recA1 relA1 lac glnV44 F' [::Tn10 proAB$^+$ lacI^q Δ(lacZ)M15] hsdR17(r_K^- m_K^+)) were transformed with the *DpnI* treated PCR product and incubated on 0.1 mg/ml ampicillin (Amp) containing lysogeny broth (LB) agar plates overnight at 37ºC. Overnight cultures, inoculated from grown colonies, were grown in LB and

0.1 mg/ml Amp to amplify the mutated plasmid. Extraction and Purification of the plasmid was done using a QIAprep Spin MiniPrep Kit from QIAGEN (Hilden, Germany). The success of the site-directed mutagenesis was confirmed by sequencing performed by GATC Biotech (Konstanz, Germany).

3.8.2 Cloning of PV constructs

Cloning of PV constructs in the vector pQE16_RS and pQE_80L (from Qiagen, Hilden, Germany) was performed using the In-Fusion® Advantage PCR Cloning Kit from Clontech (Mountain View, CA, USA) according to the manufacturer's instructions. Subsequently, a thioredoxin and a His-tag were fused to the C-terminus of the protein using the same kit. Gene and vector backbone were PCR amplified. Linearized pQE16_RS was a kind donation of M. Hösl[214]. Primer sequences are given in Table 3.4.

Name	Direction	Primer sequence
Amplification carp β-PV		
pQE16_RS_PV_infu_fw	5'➔3'	gag gag aaa tta act atg gct tt gct ggt gtt ctg a
pQE16_RS_PV_infu_rev	3'➔5'	ctc agc taa tta agc tca tgc ttt aac cag agc ggt aa
Linearization carp β-PV in pQE16_RS		
pQE16_RS_PV_lin_fw	5'➔3'	gct ttc gct ggt gtt ctg aac ga
pQE16_RS_lin_rev	3'➔5'	agt taa ttt ctc ctc ttt aat gaa ttc tgt gt
Amplification thioredoxin-His-Tag		
pQE16_RS_His_PV_infu_fw	5'➔3'	gag gag aaa tta act atg ggc cat cac cat cac cat
pQE16_RS_His_infu_rev	3'➔5'	aac acc agc gaa agc gga tcc acg cgg aac cag g
Linearization pQE80_L		
D66 pQE16 infu fwd	5'➔3'	gct taa tta gct gag ctt gga ctc c
111 pQE16 infu rev	3'➔5'	aca cag aat tca tta aag agg aga aat taa ct
Amplification carp β-PV-thioredoxin-His		
pQE homology His fw	5'➔3'	gag gag aaa tta act atg ggc cca tca c
pQE homology PV rev	3'➔5'	ctc agc taa tta agc tca tgc ttt aac ca

Table 3.4 Primer sequences for *in vitro* homologous recombination by In-Fusion PCR cloning

Gene amplification was done with a *Pfu* DNA Polymerase from Promega (Madison, WI, USA) and vector linearization was done with Phusion High-Fidelity DNA Polymerase from New England Biolabs (Ipswich, MA, USA). The according temperature programs are given in Table 3.5 and Table 3.6.

Temperature	Time	Cycles
95 ºC	2 min	1x
95 ºC 58 ºC 74 ºC	45 s 30 s 50 - 90 s	35x
74 ºC	5 min	1x
8 ºC	24 h	1x

Table 3.5 PCR Temperature program for gene amplification

Temperature	Time	Cycles
98 ºC	3 min	1x
98 ºC 58 ºC 72 ºC	30 s 45 s 1. 5 - 2.5 min	35x
72 ºC	5 min	1x
8 ºC	24 h	1x

Table 3.6 PCR Temperature program for vector linearization

3.8.3 Co-transformation

For expression of recombinant PV N7-Aha-His, $CaCl_2$ competent Met-auxotrophic B834(DE3) *E.coli* cells (F^- *dcm ompT hsdS*(r_B^- m_B^-) gal *met* λ(DE3 [*lacI lacUV5*-T7 gene 1 *ind1 sam*7 *nin*5])) were co-transformed with pQE16_RS_PV N7-Aha-His and pRep4 (Qiagen, Hilden, Germany). Competent cells were thawed on ice. 100 ng per plasmid DNA were added and the mixture was incubated on ice for 30 min. Following, a heat shock at 42 ºC was applied for 2 min, 800 µl SOC medium (SOB medium + 20mM glucose) was added immediately to the transformed bacteria. Cells were recovered for 1 h at 37 ºC and 200 rpm. Finally, the cells were pelleted by centrifugation and resuspended in 100 µl SOB medium (2% bacto-tryptone, 2.5% yeast extract, 10mM NaCl, 25mM KCl, 10mM $MgSO_4$, 10mM $MgCl_2$), plated on agar plates with ampicillin (Amp) and kanamycin (Kan) antibiotics and incubated overnight at 37 ºC.

For expression of recombinant PV N7-Aha-E16-Nal-His electro-competent Met-auxotrophic B834(DE3) *E.coli* cells were co-transformed with pQE80L_PV N7-Aha/E16-Nal-His and pEVOL_NapA (kind donation of P.G. Schultz[194]). Competent cells were thawed on ice. 100 ng per plasmid DNA were added and the mixture was applied to the electroporation cuvette (GenePulser/Micropulser cuvettes, 1 mm gap width; Biorad, Hercules CA, USA). Electroporation was performed in an Electroporator 1000 (Stratagene, La Jolla, CA, USA) by

applying 1650 kV. 1 ml SOC medium was added immediately after electroporation and the suspension was transferred to a sterile 1.5 ml microfuge tube. Subsequently, cells were incubated for 1 h at 37 ºC and 200 rpm. Finally, the cells were plated on agar plates containing Amp and chloramphenicol (CAM) and incubated overnight at 37 ºC.

3.9 Protein expression and purification

3.9.1 Expression of PV N7-Aha-His

2x 1 l LB + Amp + Kan medium was inoculated with 2 ml overnight cultures of transformed Met-auxotrophic B834(DE3) *E.coli* cells and incubated for 4 h at 37 ºC until an OD_{600} of 0.6 was reached. Cells were pelleted at 4500 rpm for 15 min, resuspended in sterile H_2O and pelleted again. Finally, cells were resuspended in 1 l NMM medium (7.5mM $(NH_4)_2SO_4$, 8.5mM NaCl, 22mM KH_2PO_4, 47.6mM K_2HPO_4, 50mg/l all AA –Met, 20mM glucose, 1mM $MgSO_4$, 1µg/ml Ca^{2+}, 1µg/ml Fe^{2+}, 10ng/ml trace elements, 10µg/ml thiamine, 10µg/ml biotin) and incubated at 37 ºC for 30 min. Subsequently, protein expression was induced by adding 100mg/l azidohomoalanine (Aha) and 0.1% lactose in H_2O. The cells were harvested after 4 h of expression and pelleted by centrifugation, frozen in liquid nitrogen and stored at -80 ºC.

3.9.2 Expression of PV N7-Aha-E16-Nal-His

2x 1 l LB + Amp + CAM medium was inoculated with 2 ml overnight cultures of transformed Met-auxotrophic B834(DE3) *E.coli* cells and treated as described in 3.9.1. In order to induce protein expression, 100mg/l Aha (dissolved in H_2O), 215mg/l 2-naphthylalanine (Nal) (dissolved in 10 ml 1M HCl), 10 ml NaOH, 1.3mM arabinose and 0.1% lactose was added. After 4 h of expression, cells were harvested, pelleted, frozen in liquid nitrogen and stored at -80 ºC.

3.9.3 Purification of PV

Harvested cells were resuspended in 50 ml binding buffer (20mM NaH_2PO_4, 0.5M NaCl, 8M Urea, 5mM imidazole, pH 7.4) and lysed by sonication. Lysates were cleared by high speed centrifugation (20000 rpm, 4 ºC, 30 min) and filtered through a 0.45 µm filter. The soluble target protein was purified from the supernatant using 2 sequential 5ml HisTrap Ni-NTA columns (GE Healthcare, Freiburg, Germany) equilibrated with binding buffer. The column

was washed with binding buffer until absorption reached zero again and the flow-through was collected. Bound PV-His was eluted with an imidazole gradient (5 – 500mM). Fractions containing the desired protein were pooled and the protein was refolded by dialysis against a 100fold excess of $NH_4^+HCO_3$ using a Spectra/Por® Dialysis Membrane tubing (Spectrumlabs, Rancho Dominguez, CA, USA) with an appropriate MWCO. Precipitated protein was removed by filtration. Protein purity was analyzed by SDS-PAGE and Coomassie staining (see 3.9.5). Furthermore, successful incorporation of Aha and Nal was tested by mass spectrometry. Purified protein was lyophilized and stored at -20 ºC.

3.9.4 Thrombin cleavage

The thioredoxin-His tag was removed from the C-terminus of the protein via cleavage with thrombin. Lyophilized protein was dissolved in thrombin buffer (200mM Tris-HCl, 150mM NaCl, 2.5mM $CaCl_2$, pH 8.4) and a spatula tip of thrombin was added. After incubation at RT for 48 h, the solution was loaded on a HisTrap Ni-NTA column and the flow-through, containing the cleaved protein, was collected. The bound thioredoxin-His tag was eluted with 500mM imidazole. The buffer was changed to $NH_4^+HCO_3$ via dialysis and the samples were lyophilized and stored at -20 ºC.

3.9.5 Sodium dodecylsulfate polyacrylamide gel electrophoresis

Discontinuous sodium dodecylsulfate polyacrylamide gel electrophoresis (SDS-PAGE) after Laemmli[215] was used to analyze the purity of protein samples. 17% acrylamide gels in combination with 4% stacking gels were prepared using a Serva multiple gel caster (Heidelberg, Germany). Protein samples were mixed 1:1 (v/v) with Laemmli Sample Buffer (Biorad, Hercules, CA, USA) and boiled at 95 ºC for 5 min before being loaded on the gel. Electrophoresis was performed at 100 V, 100 mA. The molecular weight size marker Roti-Mark 10 – 150 from Carl Roth (Karlsruhe Germany) was used as standard.

Gels were stained with coomassie staining solution (0.1% (w/v) Coomassie Brilliant Blue G-250, 80% methanol, 20% acetic acid) for 2 h at room temperature (RT) and destained with destaining solution (10% acetic acid, 5% methanol) prior to gel documentation.

4. Results and Discussion

4.1 Intrachain diffusion in unfolded polypeptide chains studied by TTET

In order to understand how proteins fold and accomplish their biological function, it is essential to investigate the nature of the unfolded state. The unfolded ensemble of a protein consists of multiple conformations, each characterized by the dihedral angles ϕ, ψ and ω of the respective residues. During folding, the polypeptide chain starting from this unfolded state ensemble, has to explore the conformational space to find energetically favorable interactions and form secondary structured elements. Thus, specific contact formation between two sites in an unfolded polypeptide chain represents the elementary step in protein folding and sets an upper limit for the rate at which a protein can adopt its native structure. Measuring TTET in unfolded polypeptide chains allows for the determination of the rate constants for loop formation and therefore presents a powerful tool to understand the earliest steps in protein folding.

4.1.1 Comparison of end-to-end and interior-to-interior loop formation dynamics in unfolded polypeptide chains

In previous experiments, TTET was used to study dynamics in unfolded polypeptide chains using model peptides such as alternating glycine-serine repeats (see Figure 4.1) and some natural sequences (see Figure 4.3). Hereby, the TTET labels were attached to the N-terminus and in close proximity to the C-terminus of the peptide.

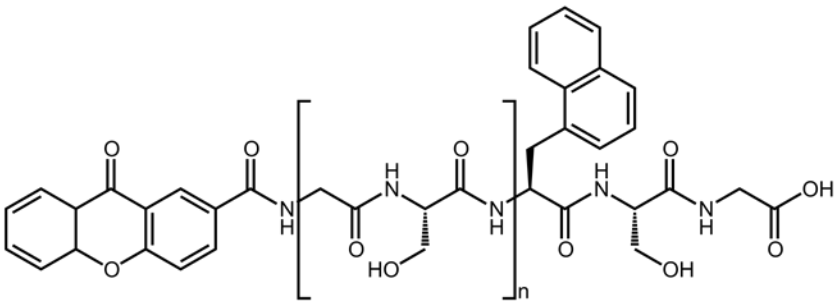


Figure 4.1 Chemical structure of poly (Gly-Ser) peptides. The sequence within the brackets is repeated n times. The triplet donor xanthone (Xan) was coupled to the N-terminus, the triplet acceptor naphthalene (Nal) was introduced as L-1-Naphthylalanine via solid phase peptide synthesis (SPPS).

Due to the high flexibility of Gly, poly(GS) peptides lack secondary structure and can serve as a model for the unfolded state under physiological conditions. The lack of side chains allows for more conformational freedom of the ϕ- and ψ-angles. All measurements in poly(GS) chains revealed single exponential kinetics yielding the rate constants for end-to-end loop formation.

The advantage of using poly(GS) peptides is the ability to establish scaling laws and to compare the results to polymer theory. The repetitive nature of the sequence ensures that alterations in chain dynamics derive from differences in chain length rather than the specific amino acid composition. A significant length dependence of the rate constants for loop formation was found that could be described by equation 4.1 (see Figure 4.6, B, black circles). Furthermore, different chemical processes limit the diffusion over long or short distances. For longer peptides, the decrease of k_c with increasing chain length is in good agreement with the Flory model for the increase in chain dynamics with increasing chain length[216], when excluded volume effects are taken into account. Loop formation over short distances is virtually independent of chain length and controlled by contributions from chain stiffness[53].

$$k_c = \frac{1}{1/k_0 + (k_i \cdot N^{\gamma})} \qquad (4.1)$$

In natural occurring proteins, end-to-end loop formation events as studied above are rather uncommon. A more relevant case is contact formation between internal residues within the polypeptide chain. In previous experiments, TTET was applied to study contact formation dynamics in end-to-interior loops (type II-loops) and interior-to-interior loops (type III-loops)[54]. In these studies, tails were attached to several series of model peptides consisting of either $(Gly\text{-}Ser)_n$ or $(Thr\text{-}Gly\text{-}Gln\text{-}Ala)_n$ sequences. In all peptides, single exponential kinetics for contact formation were observed. Type II loop formation is slower compared to type I loops, and type III loop formation is slower than type II loops. The logarithms of the observed rate constants correlate with the average end-to-end distance of the complete chain ($\sqrt{\langle r^2 \rangle}$) which can be estimated using calculated characteristic ratios for the different sequences. These studies further revealed a limiting value for the effect of additional tails on the kinetics of loop closure. This limiting value can be explained by an increased flexibility of polypeptide chains towards the chain ends, which is in agreement with results from NMR experiments[217, 218] (see Figure 4.2).

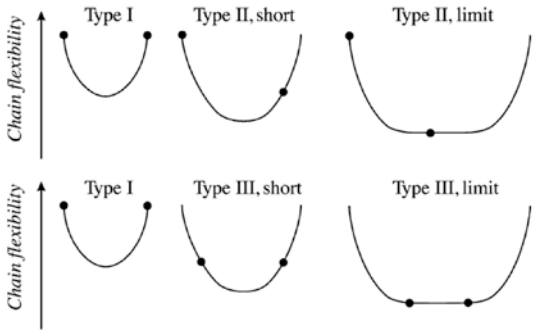


Figure 4.2 Schematic representation of the position-dependence of chain flexibility and the resulting position of the points of contact formation in loops of different types. Figure taken from[54].

This model also explains the different dynamics in type I, II and III loops. In type I loops, both contact points are located in flexible parts of the chain ends which leads to faster dynamics compared to type II and type III-loops, where one or both contact points are located in stiffer segments of the chain. The effect should be saturated for longer loops when chain stiffness has reached a limiting value. However, TTET experiments of model type III-loops did not reach a limiting value. Theoretical considerations based on polymer theory suggest that the effect of additional tails on the dynamics of loop formation is a function of tail length and chain stiffness[219]. This assumption was found to be true for type II loop formation in poly(GS) and poly(Ser) chains[54].

Polypeptide sequences from naturally occurring proteins are usually much more complex than model homo-polypeptide chains and consist of different amino acids with side chains of different size and different chemical properties. In previous studies, end-to-end contact formation was measured in unstructured loop regions derived from natural protein structures (see Figure 4.3): two loop regions connecting helices in the 108 amino acid long calcium binding carp β-parvalbumin (PV DE-Loop and PV EF-Loop)[20, 105] and a fragment from the intrinsically disordered DNA-binding Brinker domain (Brk_{66-89})[106]. In PV, both sequences are flanked by phenylalanines (Phe), which are part of the hydrophobic core and form contact in the native state of PV. The N-terminal and C-terminal Phe was replaced by Xan and Nal, respectively. The PV EF-Loop contains three glycines between the chromophores and is more flexible than the DE-Loop which contains only one Gly residue. Brk_{66-89} includes helices II and III from the DNA-binding motif. In the absence of DNA, α-helix formation in this segment is prevented by repulsive *i*, *i*+4 interactions between lysine and arginine side chains. These interactions are shielded in the presence of negatively charged phosphate groups from the DNA backbone thereby inducing α-helix formation. As Tyr77 and His80 in Brk_{66-89} would interfere with TTET experiments, these residues were replaced by Phe.

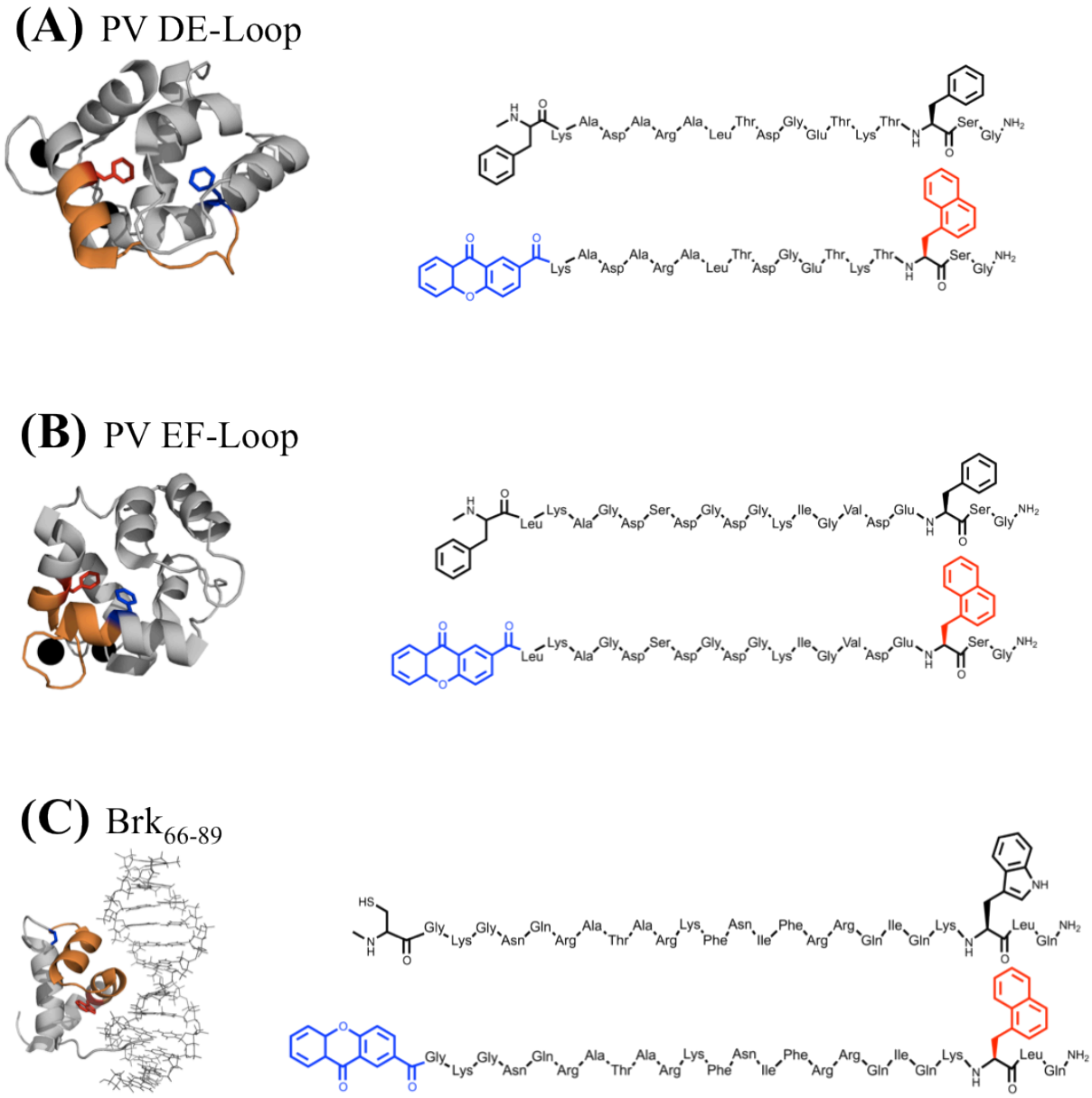


Figure 4.3 Peptide fragments from naturally occurring proteins used to study loop formation in previous experiments. On the left, the structures are shown in the context of the full-length protein, the respective amino acid sequence is given on the right. Xan is shown in blue, Nal is represented in red. (A) DE-Loop of carp β-parvalbumin (PV DE-Loop), (B) calcium-binding EF-Loop of carp β-parvalbumin (PV EF-Loop) and (C) residues 66 – 89 of the DNA-binding domain of the brinker protein (Brk_{66-89}). Figures were prepared using MacPyMOL and the pdb files 4CPV[181] for the parvalbumin sequences and 2GLO[220] for the Brinker sequence.

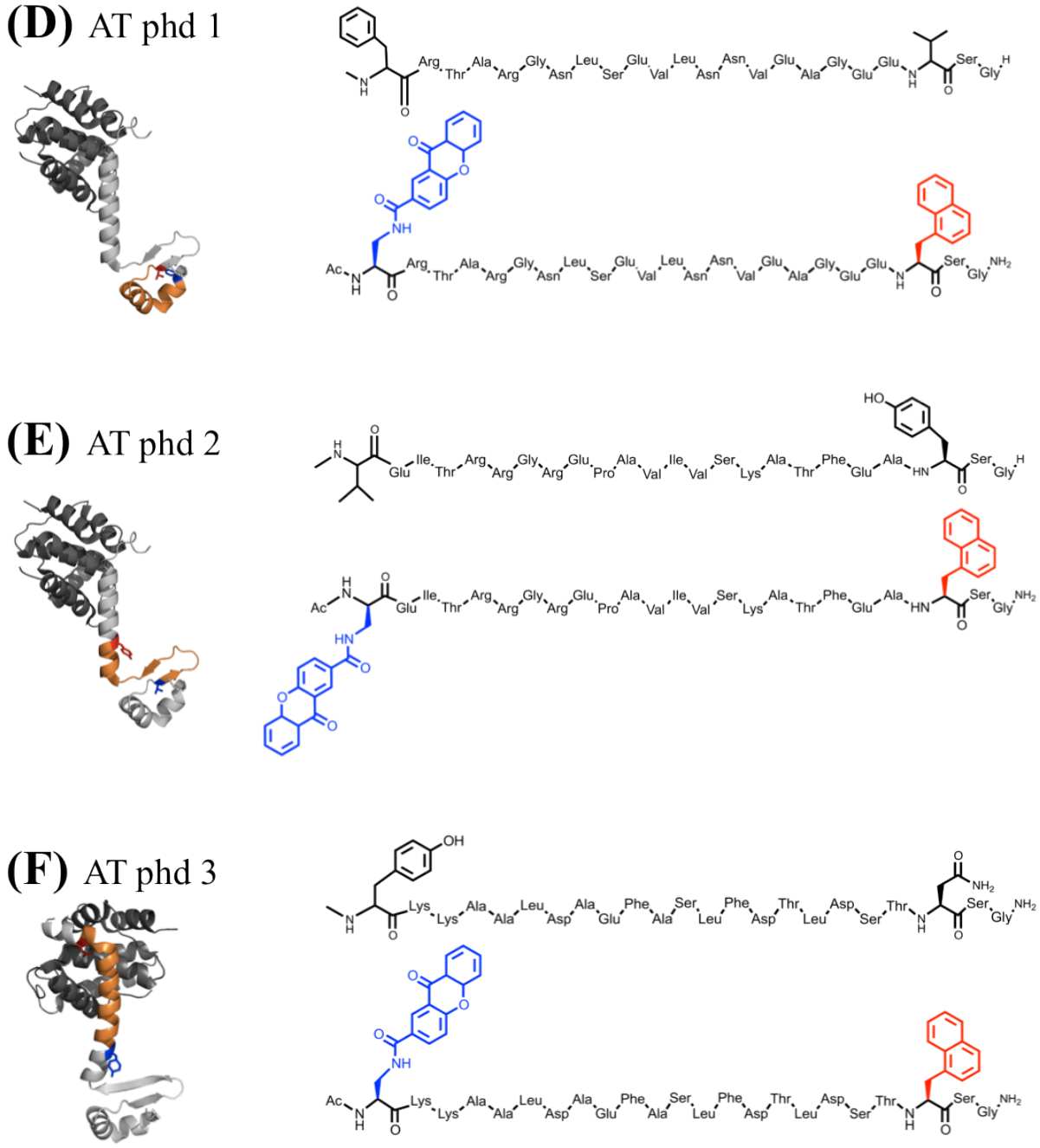


Figure 4.4 Peptide fragments from naturally occurring proteins used to study loop formation in this work. On the left, the structures are shown in the context of the full-length protein, the respective amino acid sequence is given on the right. Xan is shown in blue, Nal is represented in red. (A) Residues 7 – 25 of the intrinsically disordered protein antitoxin phd (AT phd 1), (B) residues 27 - 46 of antitoxin phd (AT phd 2) and (C) residues 48 - 66 (AT phd 3). Figures were prepared using MacPyMOL and the pdb file 3K33[221].

The kinetics of loop formation in the three peptides can be described by a single-exponential fit. A second phase with low amplitude and a time constant around 10 µs corresponds to the intrinsic donor-lifetime and probably arises from small peptide aggregates, in which the labels are unable to form contact[53]. Rate constants for loop formation in all three natural sequences are slower than for poly(GS) chains due to the presence of larger side chains, which were shown to slow down intrachain dynamics in host-guest studies[53].

Chain dynamics in the PV EF-Loop and the PV DE-Loop are comparable to those observed in poly(Ser) chains (see Figure 4.5, B, open circles). Loop formation in the unfolded Brk_{66-89} is dramatically slowed down compared to model polypeptide chains, probably due to repulsive interactions between positively charged side chains.

Peptide	sequence	charges		net charge	mean hydrophobicity
		+	-		
PV DE-Loop[b]	Xan-KADARALTDGETKT-Nal-SG-NH_3^+	4	3	+1	0.38
PV EF-Loop[b]	Xan-LKAGDSDGDGKIGVDE-Nal-SG-NH_3^+	3	5	-2	0.41
Brk_{66-89}[c]	Xan-GKGNQRATARKFNIFRRQIQK-Nal-LQ-NH_3^+	8	0	+8	0.36
AT phd 1	Xan(Dpr)-RTARGNLSEVLNNVEAGEE-Nal-SG-NH_3^+	3	4	-1	0.41
AT phd 2	Xan(Dpr)-EITRRGREPAVIVSKATFEA-Nal-SG-NH_3^+	5	3	+2	0.46
AT phd 3	Xan(Dpr)-KKAALDAEFASLFDTLDST-Nal-SG-NH_3^+	3	4	-1	0.48

Table 4.1 Amino acid sequences of investigated natural fragments. The triplet donor Xan is attached to the N-terminus and the triplet acceptor Nal in the vicinity of the C-terminus. Chain flexibility was increased in all fragments in the proximity of the resin in the course of SPPS by Ser-Gly at the C-terminus. Positively charged residues are represented in red, negatively charge residues in blue. Amino acids with highly hydrophobic side chains are represented in green.

In this work, we wanted to investigate the effect of increased chain length and chain stiffness on end-to-end loop formation by measuring the TTET in three peptide fragments derived from the antitoxin phd (AT phd), a component of the *phd/doc* antitoxin-toxin operon from bacteriophage P1. The C-terminal domain of AT phd is intrinsically disordered in solution and folds into an α-helix upon binding to the toxin Doc[222]. Simultaneously, the binding of Doc to AT phd structures the N-terminal DNA-binding dimerization domain of AT phd which represses the transcription of the operon[221]. Since it contains only one of the TTET interfering amino acids (Tyr47), AT phd is well suited for TTET measurements. We synthesized three fragments of AT phd, which differ in their chain stiffness and attached Xan to the side chain of the unnatural amino acid diaminopropionic acid (Dpr) at the N-terminal position. In addition, Nal was placed in the vicinity of the C-termini (see Figure 4.4). The fragment AT phd 1 contains the residues 7 to 25; Phe6 was replaced by Dpr(Xan) and Val26 was

substituted by Nal. It contains 2 Gly residues, which increase flexibility. AT phd 2 includes the subsequent residues (27 – 46). Here, we replaced Val26 and Tyr47 by the TTET chromophores. The sequence of AT phd 2 contains one glycine and one proline residue. The Pro residue is flanked N-terminally by several charged residues, probably involved in DNA binding, and C-terminally by multiple hydrophobic amino acids that may participate in dimerization. The C-terminal part of AT phd (residues 48 – 66) was named AT phd 3 and Tyr47 and Asn67 were substituted by the TTET labels. AT phd 3 is very hydrophobic and expected to be stiffer than AT phd 1 and 2 as it contains no glycines. We confirmed that all three protein fragments are unfolded under physiological conditions by CD spectroscopy (see Figure 4.5).

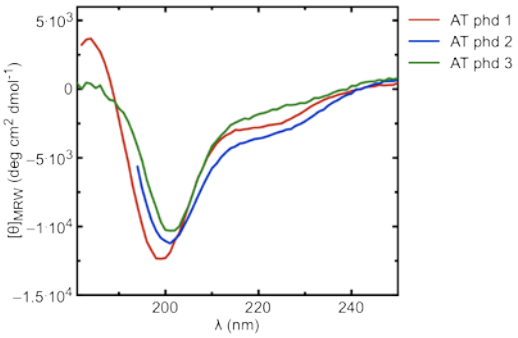


Figure 4.5 Far-UV CD spectra of unfolded fragments derived from AT phd. Spectra were recorded in 10mM PO_4 buffer, pH7 at 22.5ºC.

Loop formation in the AT phd fragments was measured in 10 mM potassium phosphate at pH7 and 22.5 ºC. The time course of the change in absorbance at 590 nm, which corresponds to the intense absorbance band of the xanthone triplet state, is depicted in Figure 4.6, A. Loop formation in AT phd 1 and AT phd 3 could be described by a single exponential decay. An additional slow phase with low amplitude is due to small amounts of aggregated peptide where the labels are unable to form contact. This phenomenon has been observed before in all investigated peptides. A rate constant for loop formation of $(1.6 \pm 0.01) \cdot 10^7$ s^{-1} was found for AT phd 1 and $(1.4 \pm 0.01) \cdot 10^7$ s^{-1} for AT phd 3. Thus, the chain dynamics of both fragments are comparable to the dynamics in poly(Ser) chains of identical length. This is surprising since significantly slower kinetics were expected for the very stiff sequence of AT phd 3, which does not contain one glycine residue.

For loop formation in AT phd 2, an additional fast phase ($(1.5 \pm 0.4) \cdot 10^7$ s^{-1}) with 10 % amplitude was detected, which is comparable with dynamics in poly(Ser) chains. This phase is most probably due to peptides with the Glu-Pro bond in *cis* configuration. The main phase

corresponds to loop formation with the Glu-Pro bond in *trans* configuration and yields a rate constant of $(6.5 \pm 0.04) \cdot 10^6$ s^{-1}, which is significantly slower compared to poly(Ser) chains and comparable to the loop formation rate constant in the Brk_{66-89} fragment. It was shown in S_nPS_n model peptides, that loop formation is significantly faster around *cis* prolyl peptide bonds compared to *trans* prolyl peptide bonds. Furthermore, it was found that loop formation is slower around *trans* prolyl peptide bonds compared to any other amino acid. However, the influence of Pro residues was only observed for short loops containing between 2 and 10 residues[223]. In the case of the larger AT phd 2 loop, the reason for slow loop formation kinetics could also be the formation of residual β-sheet or polyproline II structure in the center of the *trans*Pro peptide and hence less flexibility. However, the CD spectrum of AT phd 2 does not give evidence for residual structure (see Figure 4.5).

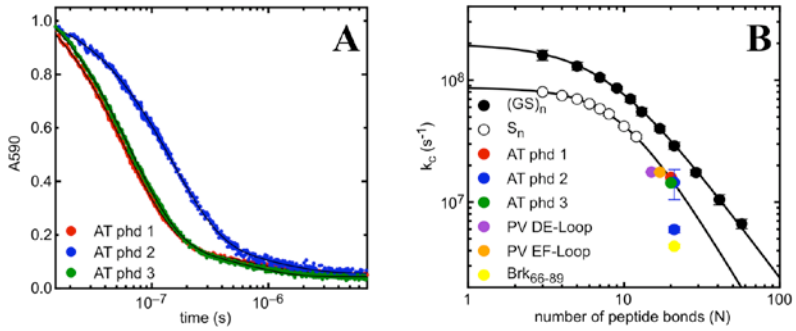


Figure 4.6 Loop formation kinetics in unfolded peptides. (A) Triplet decay curves of xanthone monitored by the change in absorbance at 590 nm. (B) Loop formation rate constants in comparison to poly(Gly-Ser)- and poly(Ser)-peptides. For AT phd 2 a faster and a slower rate constant are plotted corresponding to peptides with the prolyl bond in *cis* and *trans* configuration, respectively. Figure adapted from[53]

Polymer theory predicts that interior loop closure reactions become slower with increasing tail length until a limiting value is reached[224-226]. The effects of additional tails were predicted to depend on the size and the stiffness of both the loop and the tails[219]. We wanted to investigate whether this limiting value can be reached by the attachment of additional tails to the ends of natural sequences of higher stiffness. Therefore, we extended both ends of the natural sequences described above by 6 amino acids corresponding to the natural protein sequence (see Figure 4.7). The TTET labels were incorporated at the same positions as described above; Nal was directly introduced via SPPS and Xan was coupled to the side chain of the unnatural amino acid Dpr. The DE-Loop and EF-Loop from carp β-parvalbumin represent typical loop regions of medium length and flexibility.

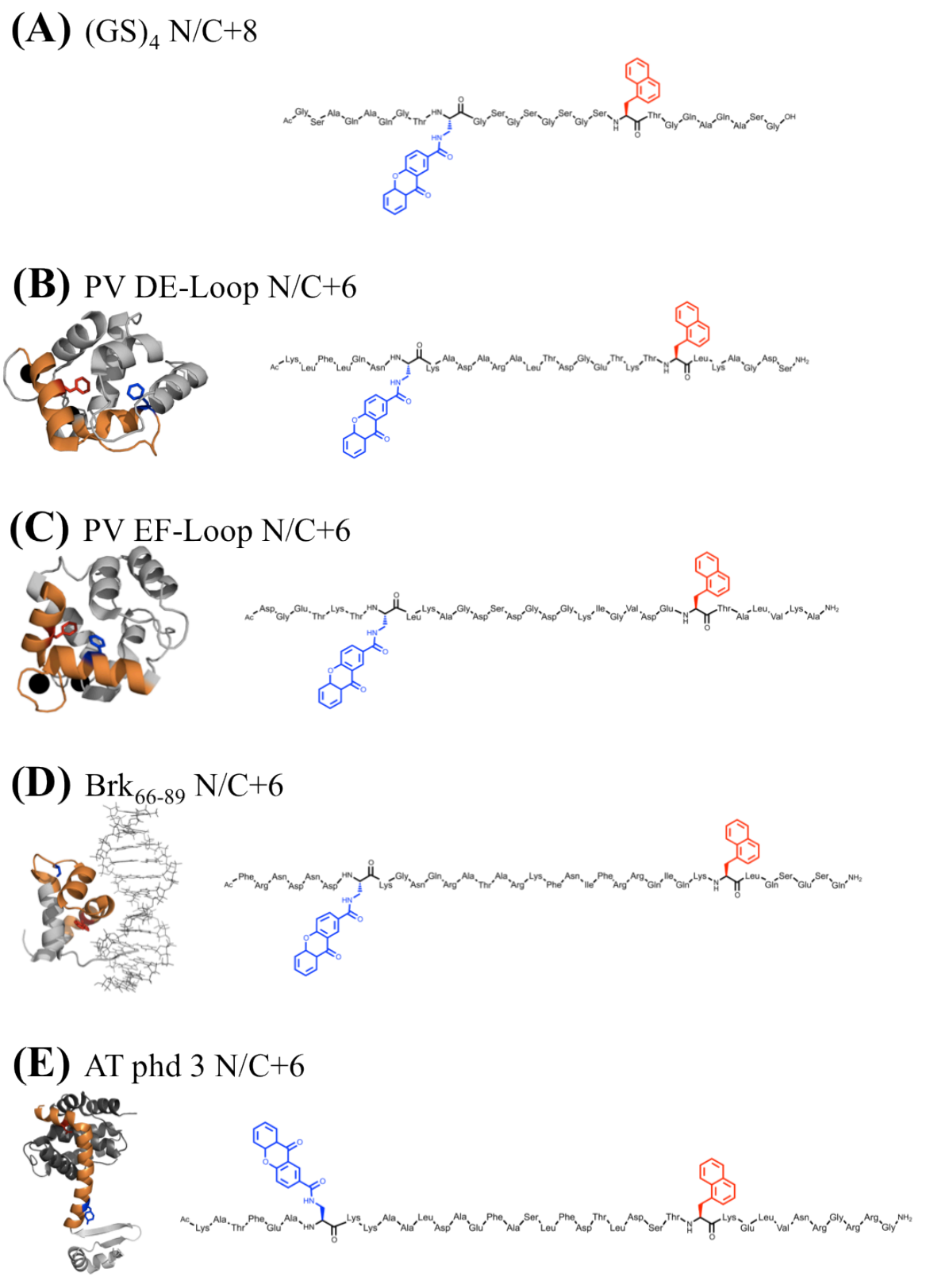


Figure 4.7 Peptide fragments used to study the effect of end extensions on loop formation. On the left, the structures are shown in the context of the full-length protein, the respective amino acid sequence is given on the right. Xan is shown in blue and Nal is represented in red. (A) Loop consisting of four Gly-Ser pairs extended on both ends by 8 uncharged AAs, (B) DE-Loop of carp β-parvalbumin with 6 amino acids of the natural sequence attached to the N- and C-terminus respectively, (C) EF-Loop of carp β-parvalbumin with extended ends, (D) residues 60 – 93 of the DNA-binding domain of the brinker protein and (E) residues 41 - 73 of AT phd 3. Figures were prepared using MacPyMOL and the pdb files 4CPV[181] for parvalbumin fragments, 2GLO[220] for the Brinker peptide and 3K33[221] for AT phd.

Brk_{66-89} and AT phd 3 derive from intrinsically disordered proteins that fold upon binding to their specific interaction partner and are longer and of different amino acid composition than the PV fragments. Brk_{66-89} contains several positively charged amino acid side chains placed in *i*, *i*+4 spacing that prevent α-helix formation in the absence of DNA or phosphate. AT phd 3 is very hydrophobic and represents a sequence of high stiffness due to the absence of glycine residues. Due to its high hydrophobicity, AT phd 3 with end extensions (AT phd 3 N/C+6) showed particularly low solubility in aqueous buffer. In order to increase solubility and enable TTET measurements, the hydrophilic amino acid sequence Gly-Arg-Arg-Gly was attached to the C-terminus. The attachment of two Arg residues increases the net positive charge of the peptide resulting in greater interaction with the solvent and reduction in self-association due to the hydrophobic patches in AT phd 3 N/C+6.

Interior-to-interior loop formation in natural sequences was compared to a $(GS)_4$ peptide with N- and C-terminal extensions of 8 amino acids free of charged side chains[54]. All peptides except AT phd 3 N/C+6 peptides are unstructured under physiological conditions as judged from CD spectroscopy (see Figure 4.8, A). TTET measurements in all peptides except AT phd 3 N/C+6 revealed single exponential kinetics for loop formation. In $(GS)_4$ N/C+8, an additional slow phase appeared due to the low solubility of this peptides and hence the formation of aggregates.

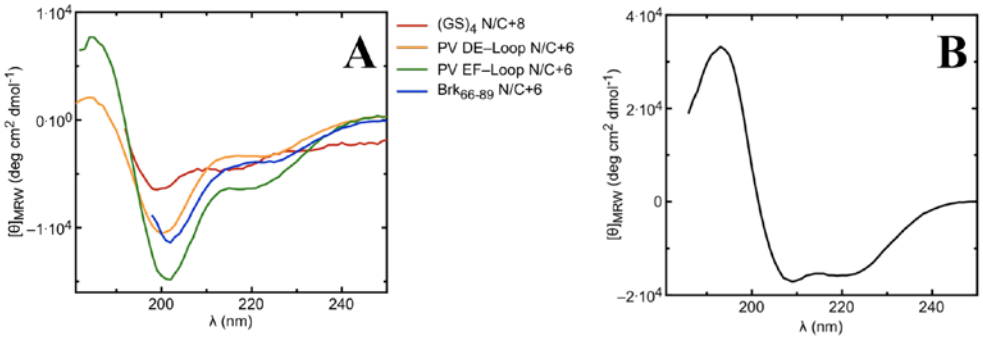


Figure 4.8 (A) Far-UV CD spectra of unfolded polypeptides with end extensions. Spectra of $(GS)_4$ N/C+8, PV DE-Loop N/C+6 and PV EF-Loop N/C+6 were recorded in 10mM PO_4 buffer, pH7 at 22.5ºC; Brk_{66-89} N/C+6 was measured in 20mM Ac, 20mM NaCl, pH5.7, at 22.5ºC. (B) Far-UV CD spectrum of AT phd 3 N/C+6 in 10mM potassium PO_4 buffer at 22.5 ºC.

Surprisingly, the far-UV CD spectrum of AT phd 3 N/C+6 displayed typical helical bands with a maximum of the ellipticity at 190 nm and minima at 208 nm and 222 nm (see Figure 4.8, B). A quantitative analysis of the helix content using the signal at 222 nm gives a value of 46% helix content. AT phd 3 N/C+6 contains the C-terminal part of the antitoxin phd (residues 41 – 73) that was reported to be unstructured under physiological conditions in the

absence of its cognate binding partner Doc[222]. Unlike in other homologous antitoxins[227, 228], the functional toxin-binding region in phd seems to be partly structured even in its free state. Nevertheless, it is possible to determine the loop formation rate constant k_c for AT phd 3 N/C+6 via TTET measurements. Therefore, we applied a three-state system described in equation 4.2. The TTET labels in AT phd 3 N/C+6 are inserted in *i*, *i*+20 spacing and hence point to opposing sites when the helix is folded. The local structure between the labels has to unfold or partly unfold in order to bring the labels together and facilitate the energy transfer reaction.

$$N \underset{k_f}{\overset{k_u}{\rightleftharpoons}} U \xrightarrow{k_c} U^* \qquad (4.2)$$

In equation 4.2, N represents a folded structure with the labels separated and U represents an unfolded or partially unfolded conformation in which the region between the labels is flexible and loop formation can occur. In U^* the irreversible energy transfer from xanthone to naphthalene has taken place. If all three rate constants k_f, k_u and k_c are in the same order of magnitude and both N and U are populated, k_f, k_u and k_c can be directly calculated from the two apparent rate constants and their corresponding amplitudes by equation 4.3[98, 109].

$$\lambda_{1/2} = \frac{k_u + k_f + k_c \pm \sqrt{(k_u + k_f + k_c)^2 - 4k_u k_c}}{2} \qquad (4.3a)$$

$$A_1 = \frac{1}{\lambda_1(\lambda_1 - \lambda_2)}([C]_0 \cdot k_c \cdot (k_u - \lambda_1) + [H]_0 \cdot k_u \cdot k_c) \qquad (4.3b)$$

$$A_2 = \frac{1}{\lambda_2(\lambda_1 - \lambda_2)}([C]_0 \cdot k_c \cdot (\lambda_2 - k_u) + [H]_0 \cdot k_u \cdot k_c) \qquad (4.3c)$$

The TTET data and the rate constants for interior-to-interior loop formation in aqueous buffer at 22.5 ºC and pH 7 are summarized and compared to end-to-end loop formation in Figure 4.9 and Table 4.2.

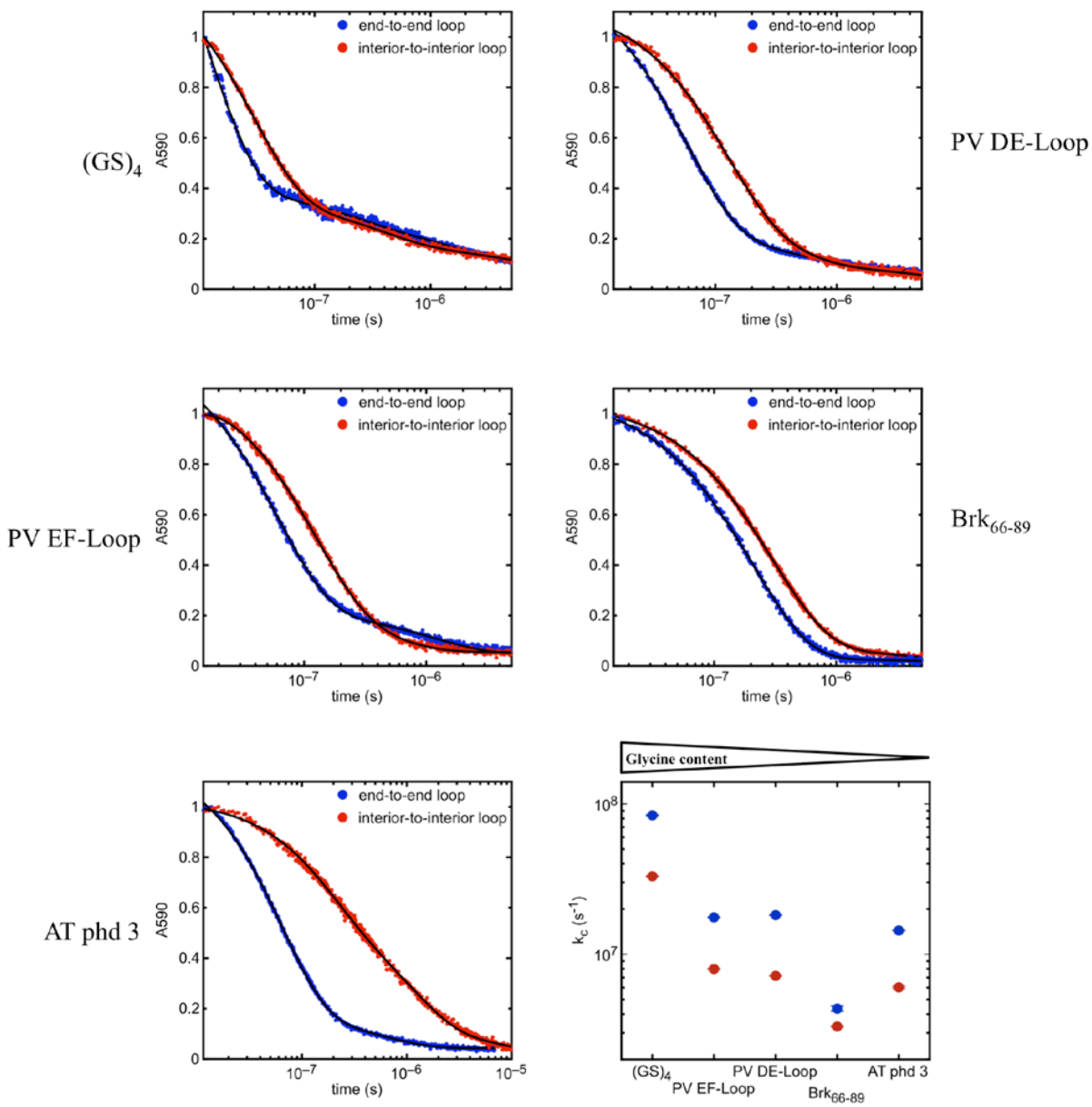


Figure 4.9 Triplet decay curves of Xan monitored by the change in absorbance at 590 nm. The blue traces correspond to end-to-end loop formation reactions and the red curves represent loop formation in the corresponding peptides with end extensions Black lines correspond to exponential fits. AT phd 3 N/C+6 was fitted with equation 4.3. Data of $(GS)_4$ obtained from[22], data of PV DE-Loop and PV EF-Loop from[105] and data of $Brk_{66\text{-}89}$ obtained from[106]. The effect of end extensions on loop formation rate constants is plotted in the panel on the lower right. Rate constants for end-to-end loop formation are shown in blue, rate constants for interior-to-interior loop formation are represented in red.

Additional tails decelerate loop formation in all observed unfolded polypeptide chains. This can be explained by the decreased flexibility in the chain center compared to the chain ends. Peptides with end extensions exhibit loop formation rate constants 2.2 – 2.5-fold slower than the same sequence without end extensions. The effect is similar for all observed peptides except $Brk_{66\text{-}89}$ independent of loop length or flexibility. This is surprising, since polymer theory predicts an increased effect of additional tails with increasing stiffness of the loop

sequences[219]. We therefore expected the effect for the very stiff AT phd 3 to be higher. However, for Brk_{66-89}, the effect is significantly smaller with an interior loop formation rate constant 1.3-fold slower than the end-to-end contact formation rate constant. Brk_{66-89} represents a special case of an unfolded peptide due to repulsing interactions from its charged amino acids.

Peptide	k_c **(end-to-end loop)** $(10^6\ s^{-1})$	k_c **(interior loop)** $(10^6\ s^{-1})$	k_c **(end-to-end)/** k_c **(interior)**
$(GS)_4$[a]	84.0 ± 0.8	32.9 ± 0.2	2.5
PV EF-Loop[b]	17.6 ± 1.1	8.0 ± 0.8	2.2
PV DE-Loop[b]	18.2 ± 1.0	7.2 ± 0.5	2.5
Brk_{66-89}[c]	4.4 ± 1.7	3.3 ± 0.2	1.3
AT phd 3	14.3 ± 0.6	5.9 ± 0.2	2.4

Table 4.2 Loop formation rate constants for peptides with and without end extensions. [a]Data for end-to-end loop formation taken from[22], [b]data for end-to-end loop formation taken from[105], [c]data for end-to-end loop formation taken from[106].

4.1.2 Characterization of the barriers for end-to-end and interior-to-interior loop formation

According to the general rate equation 4.4, the rate constant for a reaction depends on the free energy barrier $\Delta G^{0\ddagger}$. However, the pre-exponential factor k_0, which represents the maximum rate constant in the absence of free energy barriers is not known.

$$k = k_0 \cdot e^{\left(-\frac{\Delta G^{0\ddagger}}{RT}\right)} = k_0 \cdot e^{\left(-\frac{\Delta H^{0\ddagger} - T\cdot\Delta S^{0\ddagger}}{RT}\right)} \qquad (4.4)$$

Therefore, in order to experimentally determine the barriers for reactions, the Arrhenius equation (4.5) is commonly used[229].

$$k = A \cdot e^{\left(-\frac{E_A}{RT}\right)} \qquad (4.5)$$

A is a proportionality constant that varies from one reaction to another and E_A is the activation energy for the reaction. The enthalpic contribution to the free energy barrier (activation enthalpy $\Delta H^{0\ddagger}$) can be determined from the Arrhenius activation energy according to equation 4.6. For fully diffusion-controlled reactions, $\Delta H^{0\ddagger}$ is close to 0.

$$\Delta H^{0\ddagger} = E_A - RT \qquad (4.6)$$

Contributions from the activation entropy $\Delta S^{0\ddagger}$ are included in the Arrhenius pre-exponential factor *A* according to equation 4.7.

$$A = k_0 \cdot e^{\frac{\Delta S^{0\ddagger} + R}{R}} \qquad (4.7)$$

However, it is not possible to directly calculate $\Delta S^{0\ddagger}$ from *A* without knowing k_0. We assume k_0 to be in the order of the rate constant for a single bond rotation in a polypeptide chain, which is around 10^{10} s^{-1}.

We wanted to test whether loop formation reactions are diffusion controlled (E_A ~ 0) or whether additional barriers exist (E_A > 0). Furthermore, we wanted to investigate the effect of loop sequence and additional tails on the barriers for loop formation.

The viscosity of water strongly depends on temperature[230]. Thus, the determined rate constants for end-to-end loop formation have to be corrected against water viscosity changes using equation 4.8 and β-values, which can be determined by TTET measurements at different viscosities (see chapter 4.1.3).

$$k_c' = k_c \cdot \left(\frac{\eta_0}{\eta}\right)^{-\beta} \qquad (4.8)$$

Temperature dependent rate constants for end-to-end loop formation in poly(GS) chains of different length and natural sequences obtained from previous measurements exhibit apparent Arrhenius behavior ($\ln k_c \propto E_A/RT$) in all investigated polypeptide chains[22, 105, 106]. This suggests that changes in heat capacity do not contribute to the activation barrier[231].

In previous measurements, for long, highly flexible poly(GS) chains E_A-values were found to reach a limiting value around 5 kJ/mol which corresponds to a $\Delta H^{0\ddagger}$-value of 2.5 kJ/mol[22]. This is close to zero as expected for purely diffusion-limited reactions. For shorter poly(GS) chains, $\Delta H^{0\ddagger}$-values decrease with increasing chain length.

Entropically, loop formation in highly flexible poly(GS) chains is unfavorable due to their large conformational space, which is indicated by low values for the pre-exponential factor *A*[22]. *A* increases with decreasing chain length in poly(GS) chains, as a consequence, $\Delta S^{0\ddagger}$ increases from values of ~-50 J/mol/K for very long chains to a slightly positive value for

$(GS)_1$. However, even for entropically highly favorable loop formation, $\Delta S^{0\ddagger}$ is expected to be negative. This indicates that the value for k_0 is underestimated and in fact larger than 10^{10} s^{-1}.

Natural sequences investigated in former experiments exhibit both a higher activation energy and a larger pre-exponential factor compared to poly(GS) chains of the same length due to the influence of amino acid side chains. These side chains might form enthalpic intramolecular or peptide-solvent interactions that must be broken upon loop formation[105, 106]. These findings suggest that short and long peptides experience different barriers for loop formation. Barriers for end-to-end loop formation in long, highly flexible peptides mainly contain entropic contributions, while loop formation barriers in short poly(GS) chains and natural sequences are dominated by enthalpic contributions.

We measured temperature dependent end-to-end contact formation via TTET in the three unfolded fragments from AT phd. Measurements were performed in 10 mM potassium phosphate at pH 7 and temperatures between 5 ºC and 30 ºC (see Figure 4.10). Panels A, C and E show the decay of the Xan triplet state monitored by the change in absorbance at 590 nm at the indicated temperatures. Loop formation is significantly accelerated with increasing temperature. Loop formation in AT phd 1 and AT phd 3 could be described by a single exponential fit. For AT phd 2, an additional fast phase of 10 % amplitude was detected which corresponds to peptides with the prolyl bond in *cis* configuration. The main phase is due to loop formation with Pro in *trans* configuration.

Uncorrected (•) and viscosity-corrected (º) lop formation rate constants are plotted in Figure 4.10 panels B, D and F. They all exhibit apparently Arrhenius behavior and were fitted using equation 4.5. Parameters obtained from the fit are summarized and compared to previous results from other natural sequences and a $(GS)_4$ peptide in Figure 4.11 and Table 4.3.

The natural sequences exhibit both a higher activation energy and a larger pre-exponential factor compared to poly(GS) peptides. The activation energy increases with decreasing chain flexibility. This was expected since more flexible chains contain less enthalpic intramolecular interactions that must be broken upon loop formation. The activation enthalpies $\Delta H^{0\ddagger}$ for the natural protein sequences are in the range of 10 – 13.5 kJ/mol, thus comparable to $\Delta H^{0\ddagger}$ reported for the fast collapse of denatured cytochrome c[232]. This indicates that the dynamics of loop formation determine the kinetics for early steps in protein folding. An exception is the E_A-value of AT phd 2 with the prolyl bond in *trans* configuration. A higher activation enthalpy for loop formation in this peptide might be due to the decelerating effect of the *trans* prolyl peptide bond on k_c[223]. Another reason might be the existence of residual structure forming interactions that have to be broken upon loop formation.

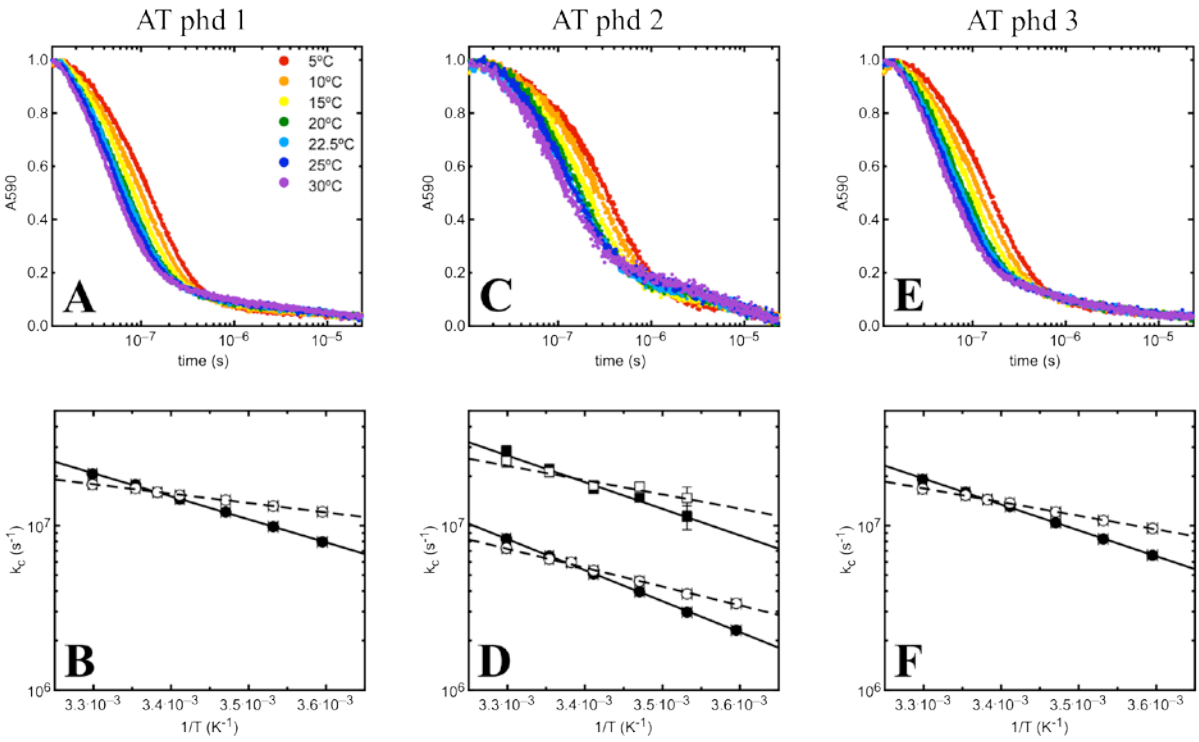


Figure 4.10 Loop formation dynamics in (A, B) AT phd 1, (C, D) AT phd 2 and (E, F) AT phd 3 at different temperatures. Top: Xan triplet absorbance decay at 590 nm Bottom: (•) uncorrected and (^{o}) viscosity-corrected (see equation 4.8) rate constants for loop formation. Solid and dashed lines represent fits to equation 4.5 to the data. For AT phd 2, squares represent rate constants of peptides with the prolyl bond in *cis* configuration and circles show rate constants for peptides with the prolyl bond in *trans* configuration.

Peptide	E_A **(kJ/mol)**	A **(10^9 s^{-1})**	$\Delta H^{0\ddagger}$ **(kJ/mol)**	$\Delta S^{0\ddagger}$ **(J/mol/K)**
$(GS)_4$[a]	6.4 ± 0.2	1.2 ± 0.1	4.0 ± 0.2	-26.3 ± 0.8
PV DE-Loop[b]	14.3 ± 0.7	4.6 ± 1.4	11.8 ± 0.7	-14.8 ± 2.5
PV EF-Loop[b]	12.5 ± 0.7	1.9 ± 0.3	10.0 ± 0.7	-22.1 ± 1.3
$Brk_{66\text{-}89}$[c]	13.3 ± 0.9	1.0 ± 0.4	10.8 ± 0.9	-27.6 ± 3.2
AT phd 1	10.9 ± 0.3	1.4 ± 0.2	8.5 ± 0.3	-24.9 ± 1.0
AT phd 2 *cis*Pro	16.6 ± 3.6	17.0 ± 24.9	14.2 ± 3.6	-3.9 ± 1.5
AT phd 2 *trans*Pro	21.9 ± 0.6	43.6 ± 11.1	19.5 ± 0.6	3.9 ± 2.1
AT phd 3	15.9 ± 0.4	9.3 ± 1.6	13.5 ± 0.4	-8.9 ± 1.4

Table 4.3 Parameters for the temperature dependence of chain dynamics. [a]Data taken from[22], [b]data taken from[105], [c]data taken from[106]. The activation entropy $\Delta S^{0\ddagger}$ was determined assuming k_0 = 100 ps.

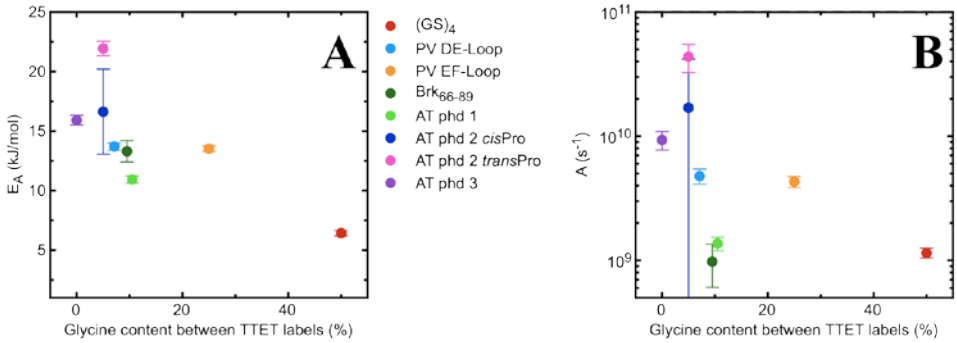


Figure 4.11 Viscosity-corrected (A) activation energies and (B) pre-exponential factors of unfolded peptides of different Glycine content between the TTET labels. Data for $(GS)_4$ taken from[22], for PV DE-Loop and PV EF-Loop taken from[105] and data for Brk_{66-89} taken from[106].

Increased pre-exponential factors for loop formation in natural sequences compared to poly(GS) chains indicate a more restricted conformational space. Hence, the entropic contributions to the free energy barrier for loop formation are lower in less flexible chains. This is expected since stiffer chains can adopt less different conformations and therefore, the conformational space is restricted. However, measurements revealed a smaller pre-exponential factor and therefore a higher loss of entropy upon loop formation for Brk_{66-89} than expected from its glycine content. The reason why loop formation in Brk_{66-89} is entropically very unfavorable might be repulsing interactions in the chain due to positively charged amino acid side chains, which lead to a more extended chain. AT phd 2 *trans*Pro and Brk_{66-89} exhibit significantly slower rate constants for loop formation compared to poly(Ser) chains and other natural sequences of the same length (see Figure 4.5). The higher free energy barrier for loop formation in AT phd 2 *trans*Pro is due to a higher activation energy. For Brk_{66-89}, however, it is due to a higher loss in entropy.

Summarized, the observed barriers for end-to-end loop decrease with increasing chain flexibility; contact formation in the most flexible, long poly(GS) chains is a fully diffusion-controlled reaction[106]. In stiffer chains, the free energy barrier contains mainly enthalpic contributions arising from intramolecular or peptide-solvent interactions of the amino acid side chains that must be broken upon loop formation. Entropic barriers are only relevant for peptides of large accessible conformational space or for chains where loop formation is entropically highly unfavorable.

In order to investigate the effect of end extensions on the free energy barriers for loop formation in unfolded polypeptide chains, we measured temperature dependencies of the five type III loop peptides described above (see Figure 4.7). Loop formation in all unfolded polypeptide chains showed single exponential kinetics.

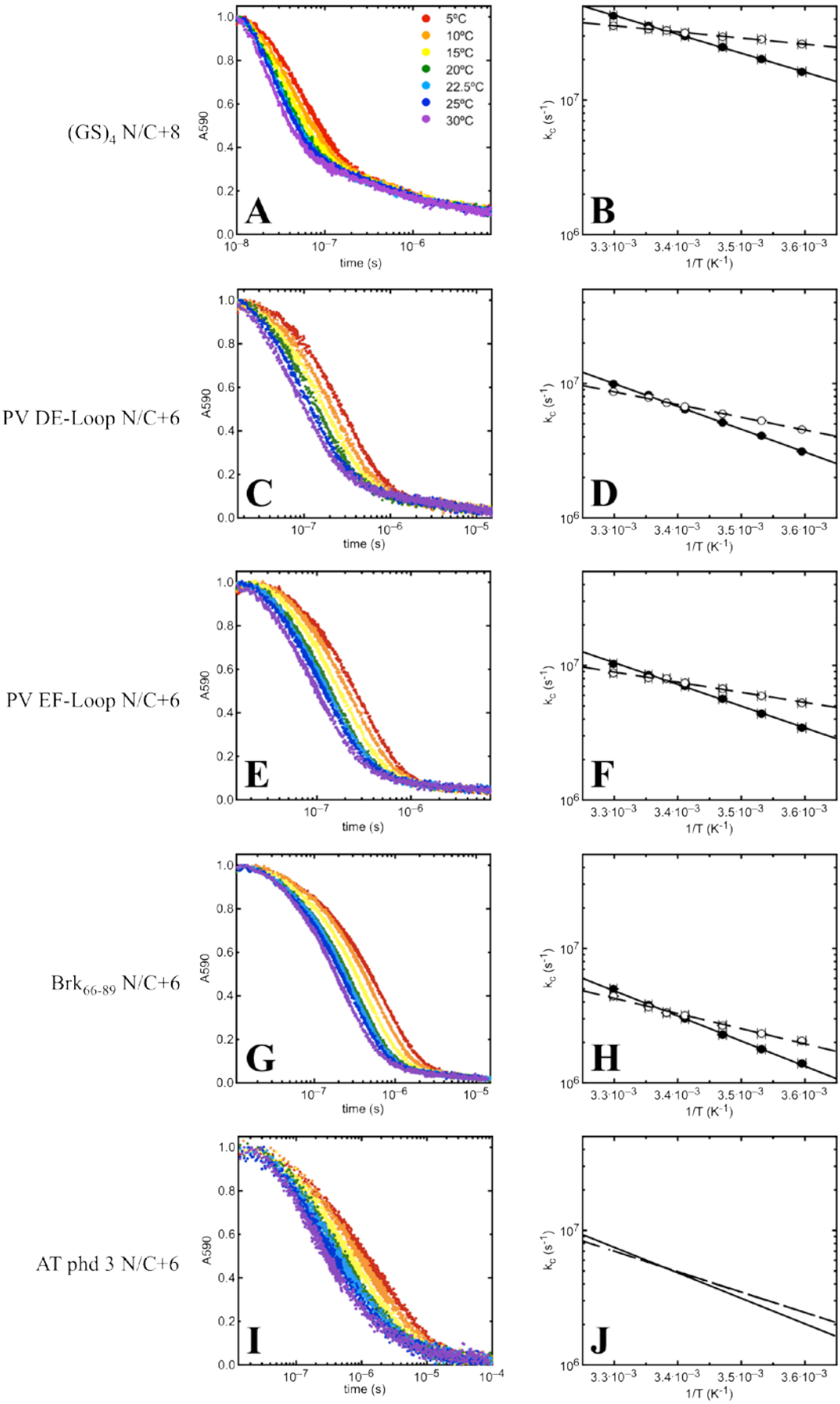


Figure 4.12 Interior-to-interior loop formation dynamics at different temperatures. Left: Absorbance decay of the Xan triplet state at 590 nm. Right: (•) uncorrected and (^{o}) viscosity-corrected (see equation 4.8) rate constants for loop formation. Solid and dashed lines represent fits to equation 4.5 to the data.

For the partially folded AT phd 3 N/C+6, a global fit of all traces using the analytical solution of the three state model (equation 4.3) was applied to obtain k_f, k_u and k_c (see Figure 4.13)[98]. Therefore, we assume a linear effect of 1/T on the logarithm of the rate constants for folding, unfolding and loop formation (Arrhenius equation 4.5).

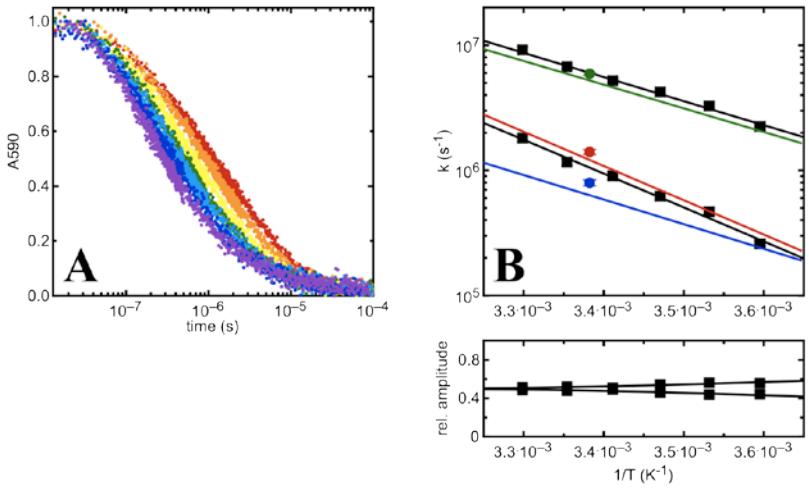


Figure 4.13 Temperature dependence of the dynamics in the partially folded AT phd 3 N/C+6. (A) Decay of the Xan triplet state monitored by the change in absorbance at 590 nm at (•) 5 ºC, (•) 10 ºC, (•) 15 ºC, (•) 20 ºC, (•) 22.5 ºC, (•) 25 ºC, (•) 30 ºC. (B) Upper panel: The two observable rate constants $\lambda_{1,2}$ (squares) of single trace double exponential fits are shown. Global fitting of all traces yields the microscopic rate constants k_u (red), k_f (blue) and k_c (green). Microscopic rate constants obtained from a non-global fit at 22.5ºC are plotted as circles. Lower panel: Relative amplitudes corresponding to $\lambda_{1,2}$.

Arrhenius parameters were obtained by fitting of the viscosity-corrected rate constants for loop formation k_c' using the Arrhenius equation 4.5. They are summarized and compared to parameters from end-to-end loop formation in Table 4.4 and Figure 4.14.

Peptide	E_A **(kJ/mol)**	A **(10^9 s^{-1})**	$\Delta H^{0\ddagger}$ **(kJ/mol)**	$\Delta S^{0\ddagger}$ **(J/mol/K)**
$(GS)_4$[a]	6.4 ± 0.2	1.2 ± 0.1	4.0 ± 0.2	-26.3 ± 0.1
$(GS)_4$ N/C+8	8.7 ± 0.3	1.2 ± 0.1	6.3 ± 0.3	-26.3 ± 0.1
PV EF-Loop[b]	12.5 ± 0.7	1.9 ± 0.3	11.1 ± 0.2	-15.3 ± 0.1
PV EF-Loop N/C+6	14.4 ± 0.5	2.7 ± 0.5	11.9 ± 0.5	-19.2 ± 0.2
PV DE-Loop[b]	14.3 ± 0.7	4.6 ± 1.4	11.2 ± 0.3	-14.5 ± 0.1
PV DE-Loop N/C+6	18.2 ± 0.2	11.7 ± 0.1	15.7 ± 0.2	-7.0 ± 0.1
$Brk_{66\text{-}89}$[c]	13.3 ± 0.9	1.0 ± 0.4	10.8 ± 0.9	-35.7 ± 0.2
$Brk_{66\text{-}89}$ N/C+6	21.6 ± 0.6	22.8 ± 5.5	19.2 ± 0.6	-1.4 ± 0.2
AT phd 3	15.9 ± 0.4	9.3 ± 1.6	13.5 ± 0.4	-8.9 ± 0.2
AT phd 3 N/C+6	29.1 ± 8.0	714 ± 23	26.6 ± 8.0	27.2 ± 3.3

Table 4.4 Parameters of temperature dependence for peptides with and without end extensions. From the Arrhenius pre-exponential factor A we calculated the activation entropy $\Delta S^{0\ddagger}$ using equation 4.7 and assuming $k_0 = 10^{10}$ s^{-1}. [a]Data taken from[22], [b]data taken from[105], [c]data taken from[106].

The activation energy for interior-to-interior loop formation is higher compared to end-to-end loop formation in the observed polypeptide chains. Hence, loop formation in unfolded peptides with end extensions is enthalpically less favorable than end-to-end loop formation. This effect increases with increasing chain stiffness (see Figure 4.14, A). As depicted in Figure 4.2, additional tails decrease chain flexibility in the center of the peptide, which is represented by higher activation energies for loop formation. We could show that this effect becomes more pronounced with increasing chain stiffness in the loop region, as expected from polymer theory[219].

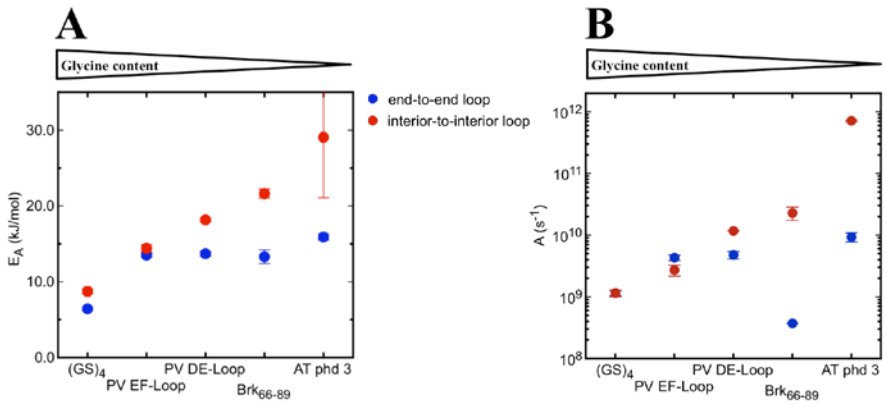


Figure 4.14 Effect of end extensions on (A) the viscosity-corrected activation energy and (B) the viscosity-corrected Arrhenius pre-exponential factor for peptides of different Glycine content.

The pre-exponential factor and hence the activation entropy for loop formation in the highly flexible $(GS)_4$ does not change upon the addition of tails to the termini. However, for the other less flexible chains in our analyses, end extensions cause an increase of *A*, indicating loop formation to be entropically more favorable (see Figure 4.14, B). The reason might be that the decrease in flexibility in the chain center upon the addition of tails to the termini leads to a restriction of the conformational space. This effect also increases with decreasing chain flexibility, but is most pronounced for $Brk_{66\text{-}89}$. The significant loss in entropy upon end-to-end loop formation could be nearly eliminated by the extension of the ends. For the stiffest peptide AT phd 3 N/C+6, calculation yielded a positive value for $\Delta S^{0\ddagger}$, indicating a gain in entropy upon loop formation. However, it is not clear if this value is reliable since we do not know if our assumption for k_0 is correct.

End extension of loop forming unfolded polypeptides lead to an increase in activation enthalpy but also to an increase in activation entropy, making the loop formation reaction enthalpically less and entropically more favorable. Since type I loops are very uncommon in naturally occurring proteins, studies on the dynamics of end-to-end contact formation might

overestimate the speed at which loops can form. The slower rate constants for interior-to-interior loop formation is due to higher stiffness in the chain center, which results in a higher enthalpic, but lower entropic barrier. These two opposing effects increase with decreasing chain flexibility, but sum up to the observed 2.2-2.5-fold decrease in the rate constants between end-to-end and interior loop formation in most unfolded polypeptides independent of chain length or flexibility. However, for peptide chains where loop formation is entropically highly unfavorable e.g. due to highly extended chains, the difference can be smaller. Thus, the investigation of contact formation dynamics in model type I-loops may overestimate the absolute rate constants of the diffusional motions relevant to structure formation during early stages of folding. Anyhow, for the comparison of the dynamics in unfolded polypeptide chains of different length or amino acid composition, end-to-end contact formation experiments are well suited.

4.1.3 Effect of solvent viscosity on end-to-end and interior-to-interior loop formation in unfolded polypeptides

Protein folding reactions are commonly described as a diffusion-driven passage over a free energy barrier along a single reaction coordinate and the reaction rate is inversely coupled to the solvent viscosity η according to Kramers' theory (see equation 4.9)[81].

$$k_c = k_c^{H_2O} \cdot \left(\frac{\eta}{\eta_0}\right)^{\beta} \qquad (4.9)$$

η_0 is the reference solvent viscosity of water and k_c^{H2O} is the rate constant of end-to-end contact formation at η_0. The empirical fractional exponent β reflects the sensitivity of the reaction to solvent viscosity. For β = -1, chain dynamics exhibit a perfect $1/\eta$ viscosity-dependence, indicating that the reaction is diffusion controlled. A β-value of 0 suggests that the reaction is independent of solvent viscosity. By measuring isothermal viscosity dependencies and fitting the resulting effects on the rate constants for loop formation with equation 4.9, it is possible to distinguish between diffusion- and reaction-controlled processes. It was shown in measurements of poly(GS) peptides that it is essential to use a viscogen with a small hydrodynamic radius (r_H), since the β-value decreases for large co-solutes[106]. Thus, the macroscopic solvent viscosity measured in a viscometer and the microscopic viscosity experienced by a polypeptide chain is only identical for small viscogens like glycerol, which allows for determination of the actual viscosity-dependence of loop formation.

In long highly flexible poly(GS) peptides with more than 15 peptide bonds, a β-value of -1 was observed indicating a diffusion-controlled behavior for end-to-end contact formation[82]. For short chains with numbers of peptide bonds smaller than 10, β deviates from -1 reaching values from 0.81 for $(GS)_1$ to 0.94 for $(GS)_4$[21, 22]. This indicates the existence of additional barriers in short poly(GS) chains that limit intrachain diffusion.

β-values of natural sequences were also shown to deviate from -1[105, 106]. In this work, we measured viscosity dependencies for end-to-end loop formation of the three unfolded fragments from AT phd. TTET was performed in different glycerol/water mixtures at pH 7 and 22.5 ºC (see Figure 4.15). Panels A – C show the decay of the Xan triplet state monitored by the change in absorbance at 590 nm at the indicated solvent viscosities. A significant deceleration of the kinetics can be observed with increasing solvent viscosity. Loop formation kinetics of AT phd 1 and AT phd 3 could be described by a single exponential fit. For AT phd 2, a second loop formation phase of 10 % amplitude was detected which corresponds to peptides with the prolyl bond in *cis* configuration. Due to its low population, reliable rate constants for this phase could only be determined at low viscosities. The obtained rate constants are plotted in Figure 4.15 D as a function of solvent viscosity and were fitted using equation 4.9. Parameters obtained from the fit are summarized and compared to previous results from other natural sequences and a $(GS)_4$ peptide in Table 4.5 and Figure 4.16.

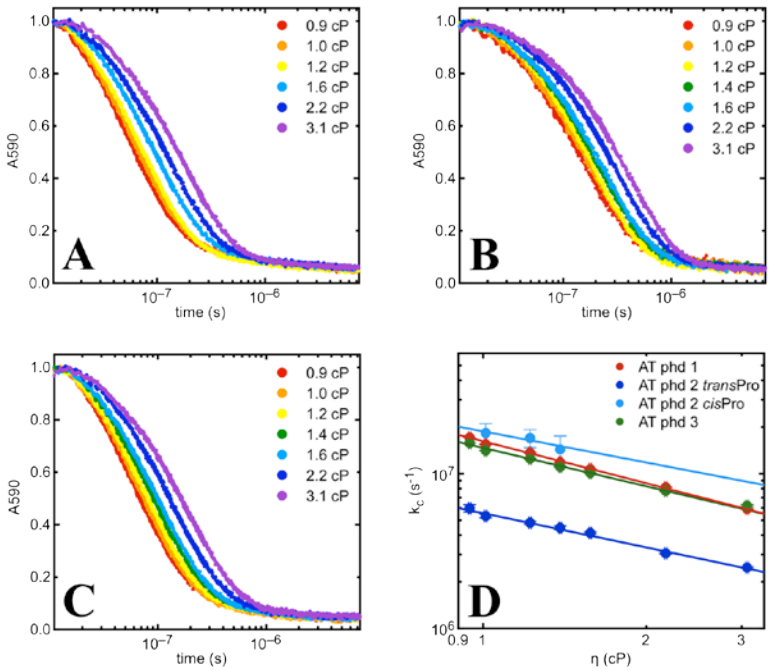


Figure 4.15 End-to-end loop formation in fragments from the intrinsically disordered protein antitoxin phd. Absorbance decay at 590 nm at different solvent viscosities for (A) AT phd 1, (B) AT phd 2 and (C) AT phd 3. (D) Comparison of rate constants for loop formation. Solid lines represent fits of equation 4.9 to the data.

Peptide	k_c^{H2O} (10^6 s^{-1})	-β
$(GS)_4$[a]	81.4 ± 2.7	0.94 ± 0.02
PV DE-Loop[b]	18.2 ± 1	0.89 ± 0.02
PV EF-Loop[b]	17.6 ± 1	0.91 ± 0.01
Brk_{66-89}[c]	4.5 ± 0.1	0.87 ± 0.02
AT phd 1	16.8 ± 0.3	0.88 ± 0.03
AT phd 2 *cis*Pro	19.5 ± 3.5	0.67 ± 0.75
AT phd 2 *trans*Pro	5.8 ± 0.1	0.73 ± 0.02
AT phd 3	15.2 ± 0.3	0.79 ± 0.03

Table 4.5 Parameters for the viscosity dependence of chain dynamics. [a]Data taken from[22], [b]data taken from[105], [c]data taken from[106].

The results reveal that the β-value approaches -1 with increasing glycine content between the labels and thus with chain flexibility. An exception is the β-value of AT phd 2 with the prolyl bond in *trans* configuration. The flexibility of the chain might be lower as expected from the glycine content due to the effect of the *trans* prolyl peptide bond or residual structure. Furthermore, the fit yields calculated loop formation rate constants at the viscosity of water at 22.5°C (k_c^{H2O}). These values may deviate from the experimentally determined rate constants.

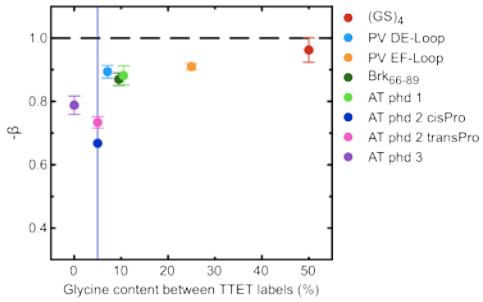


Figure 4.16 β-value of loop formation in unfolded peptides of different Glycine content between the TTET labels. Data for $(GS)_4$ taken from[22], for PV DE-Loop and PV EF-Loop taken from[105] and data for Brk_{66-89} taken from[106].

In order to characterize the contributions from the solvent to interior-to-interior loop formation in unfolded polypeptide chains, we measured viscosity dependencies. Loop formation in all unfolded polypeptide chains showed single exponential kinetics.

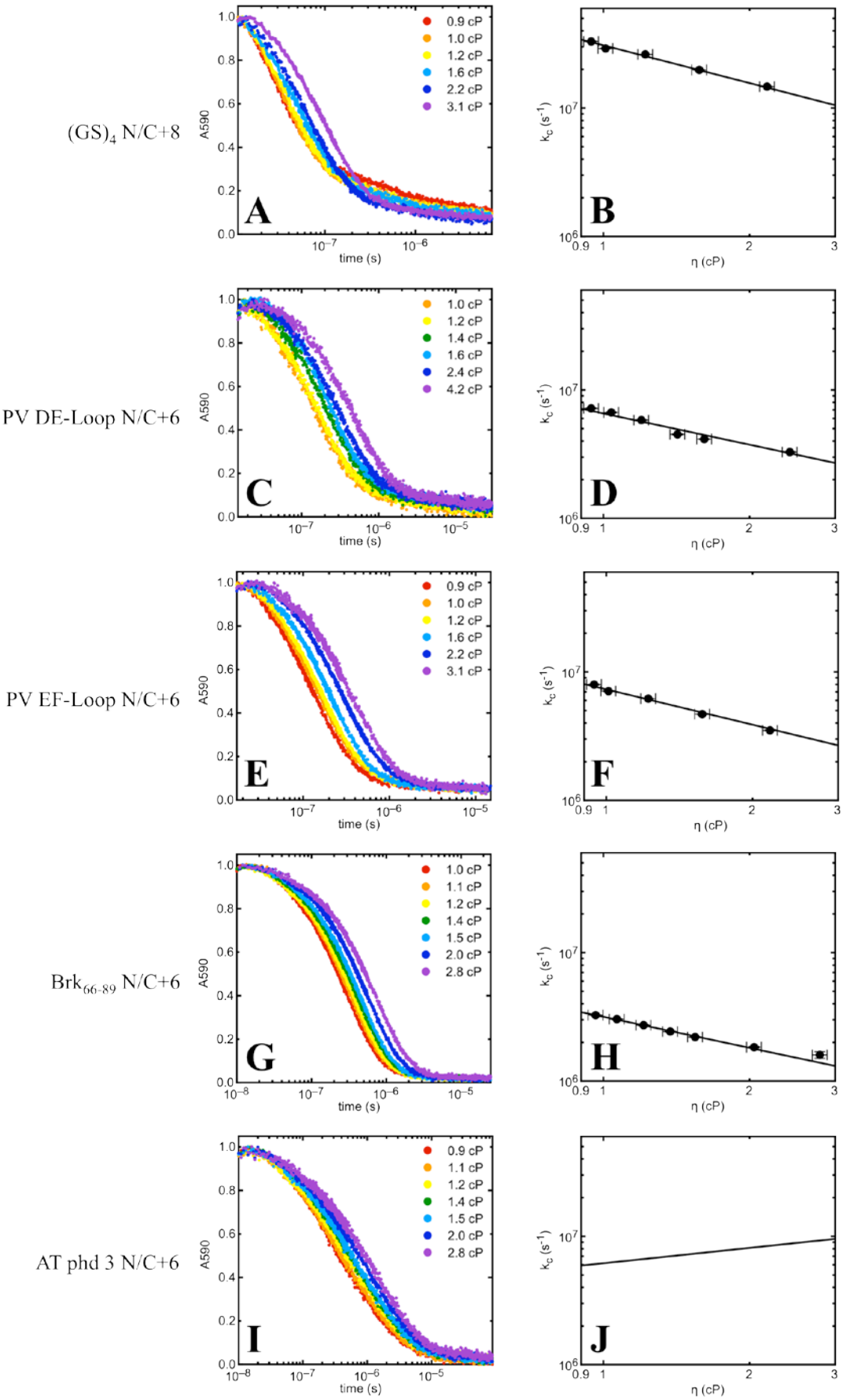


Figure 4.17 Interior-to-interior loop formation dynamics at different viscosities. Left: Decay curve of the xanthone triplet state monitored by the change in absorbance at 590 nm. Right: Rate constants for loop formation. Solid lines represent fits to equation 4.9 to the data.

For the partially folded AT phd 3 N/C+6, a global fit of all traces using the analytical solution of the three state model (equation 4.3) was applied to obtain k_f, k_u and k_c (see Figure 4.18)[98]. Therefore, we assume a linear effect of β on the logarithm of the rate constants for folding, unfolding and loop formation (equation 4.9).

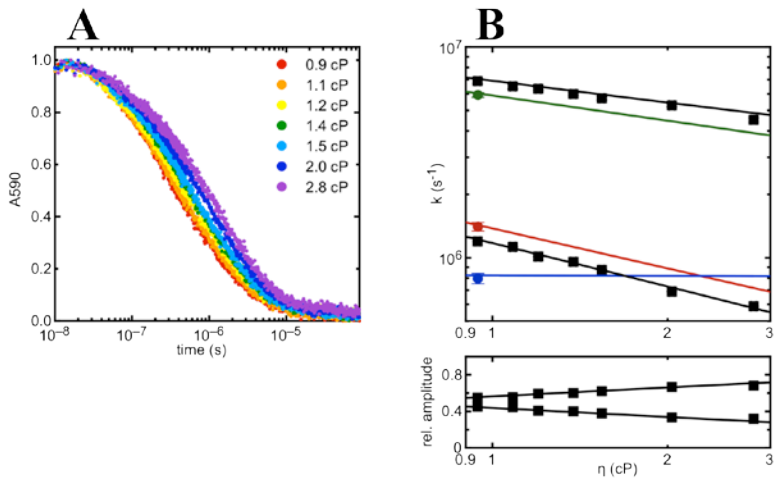


Figure 4.18 Viscosity dependence of the dynamics in the partially folded AT phd 3 N/C+6. (A) Absorbance decay at 590 nm at the indicated viscosities. (B) Upper panel: The two observable rate constants $\lambda_{1,2}$ (squares) of single trace double exponential fits are shown. Global fitting of all traces yields the microscopic rate constants k_u (red), k_f (blue) and k_c (green). Microscopic rate constants obtained from a non-global fit in water are plotted as circles. Lower panel: Relative amplitudes corresponding to $\lambda_{1,2}$.

The β-values were obtained by fitting of the rate constants for loop formation k_c using equation 4.9. Furthermore, the fit yields calculated loop formation rate constants at the viscosity of water at 22.5ºC (k_c^{H2O}). These values may deviate from the experimentally determined rate constants listed in Table 4.2. The fit parameters are summarized and compared to parameters from end-to-end loop formation in Table 4.6 and Figure 4.19.

Peptide	**k_c^{H2O} (end-to-end) (10^6 s^{-1})**	**k_c^{H2O} (interior) (10^6 s^{-1})**	**-β (end-to-end)**	**-β (interior loop)**
$(GS)_4$[a]	81.4 ± 2.7	32.6 ± 0.7	0.94 ± 0.02	0.98 ± 0.03
PV EF-Loop[b]	17.6 ± 1.0	7.7 ± 0.2	0.91 ± 0.01	0.91 ± 0.03
PV DE-Loop[b]	18.2 ± 1.0	6.9 ± 0.1	0.89 ± 0.02	0.81 ± 0.02
Brk_{66-89}[c]	4.5 ± 0.1	3.3 ± 0.03	0.87 ± 0.02	0.80 ± 0.02
AT phd 3	15.2 ± 0.3	6.0 ± 0.1	0.79 ± 0.03	0.40 ± 0.03

Table 4.6 Parameters of viscosity dependence for peptides with and without end extensions. [a]Data taken from[22], [b]data taken from[105], [c]data taken from[106].

For the highly flexible $(GS)_4$ peptide, additional tails at the termini yield a β-value that is even closer to -1 than β of the peptide without end extensions. This indicates that for highly

flexible chains, additional tails favor diffusion-controlled reaction behavior. For the EF-Loop from carp β-parvalbumin, we see no effect of the end extensions on the β-value. However, for stiffer chains, end extensions lead to a deviation away from β = -1 and this effect increases significantly with decreasing chain flexibility (see Figure 4.19). It is notably that for AT phd 3 N/C+6, we do not know whether the helical structure between the labels has to be fully unfolded to allow TTET or if a partially unfolded chain is flexible enough to bring the TTET chromophores in van der Waals contact. Residual structure would decrease the flexibility of the polypeptide chain. The quality of the fits for AT phd 3 N/C+6 could be improved by performing a urea-dependence of the TTET kinetics and globally fitting of all kinetic traces to the three-state-model (equation 4.3) because we assume the polypeptide chain to be completely unfolded at high concentrations of urea. Furthermore, the C-terminal tail contains not 6 but 10 residues due to the C-terminal amino acids required to increase solubility. It was shown that tail length influences loop formation dynamics[54]. This might explain the strong deviation of the β-value from -1.

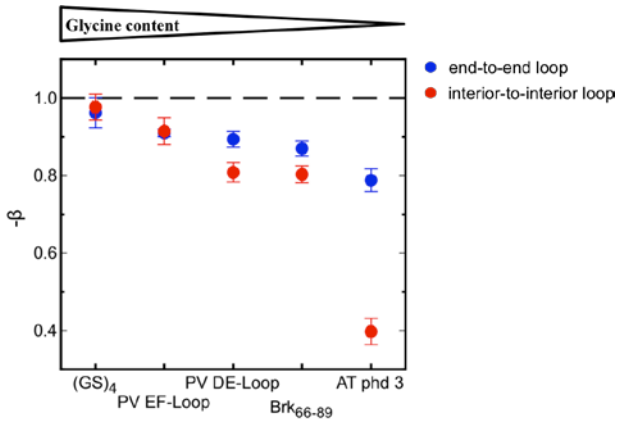


Figure 4.19 Effect of end extensions on the β-value. Peptides without end extensions are shown in blue, data obtained from peptides with end extensions is represented in red.

4.1.4 Contribution of solvent friction and internal friction to the dynamics of loop formation

Kramers' theory[81] assumes that the rate k_f for a reaction depends on the activation free energy $\Delta G^{0\ddagger}$ and a reaction friction parameter γ through $k_f \propto \gamma^{-1} \cdot \exp(-\frac{\Delta G^{0\ddagger}}{k_B T})$. Thus, the dynamics for loop formation should scale inversely with the friction γ, a fact that is supported by computer simulations[233]. However, the source of the friction γ that controls the speed of folding is still a matter of controversial debate. According to the Stokes' law, γ is proportional to the dynamic viscosity of the solvent η, which would result in complete diffusion-controlled

behavior with $k_f \propto \eta^{-1}$. Experimentally, this was found to be true for the loop formation of long unfolded model poly(GS) chains with more than 15 peptide bonds[82]. However, for shorter poly(GS) chains and unfolded polypeptide chains occurring from natural sequences, fractional viscosity dependencies were observed[105, 106, 234] (see also chapter 4.1.3). Similar observations were made for motions of simple polymers in organic solutions[83-85] and for dynamics in native proteins[86-88], suggesting additional frictional forces that hinder intrachain diffusion. These "internal" friction effects may arise from intrinsic properties of the chain, such as steric effects or intermolecular interactions[235]. The term "internal friction" has first been introduced to model the properties of threadlike polymers in solution[236]. In 1978, Haas et al. were the first to apply the concept to unfolded polypeptides, proposing that the folding time is the sum of two separate time scales; one solvent-controlled relaxation time τ_{solv} and a solvent-independent τ_{int} (equation 4.10 with $\tau = k^{-1}$)[93]. η is the viscosity of the solvent and η_0 is a reference viscosity.

$$\tau = \tau_{int} + \tau_{solv} \cdot \left(\frac{\eta}{\eta_0}\right)^{-\beta} \qquad (4.10)$$

According to this equation, contributions from τ_{int} can be directly determined by measuring τ at different viscosities and extrapolation of τ_{solv} to 0. Plotting τ against the viscosity η would result in a linear (for β = -1) or sloping (for β > -1) curve with the intercept representing τ_{int}.
Several groups observed that their data is better fit by a power law with the β-value experimentally determined from viscosity dependences. For the loop formation in poly(GS) peptides of different length studied by TTET, β has been found to be between -0.8 and -0.96 without contributions from internal friction[21, 22]. Studying the transition path time and the folding time of a two-state, all-α-helical designed protein, Chung and Eaton found β-values of -0.19 and -0.3, respectively, and an intercept of zero[97]. Similar results were observed for secondary structure formation rate constants in α-helices and β-hairpins (β = -0.64)[237] or by simulations of the viscosity dependence of first passage times in different model peptides resulting in β-values between 0.59 and 0.73[238].

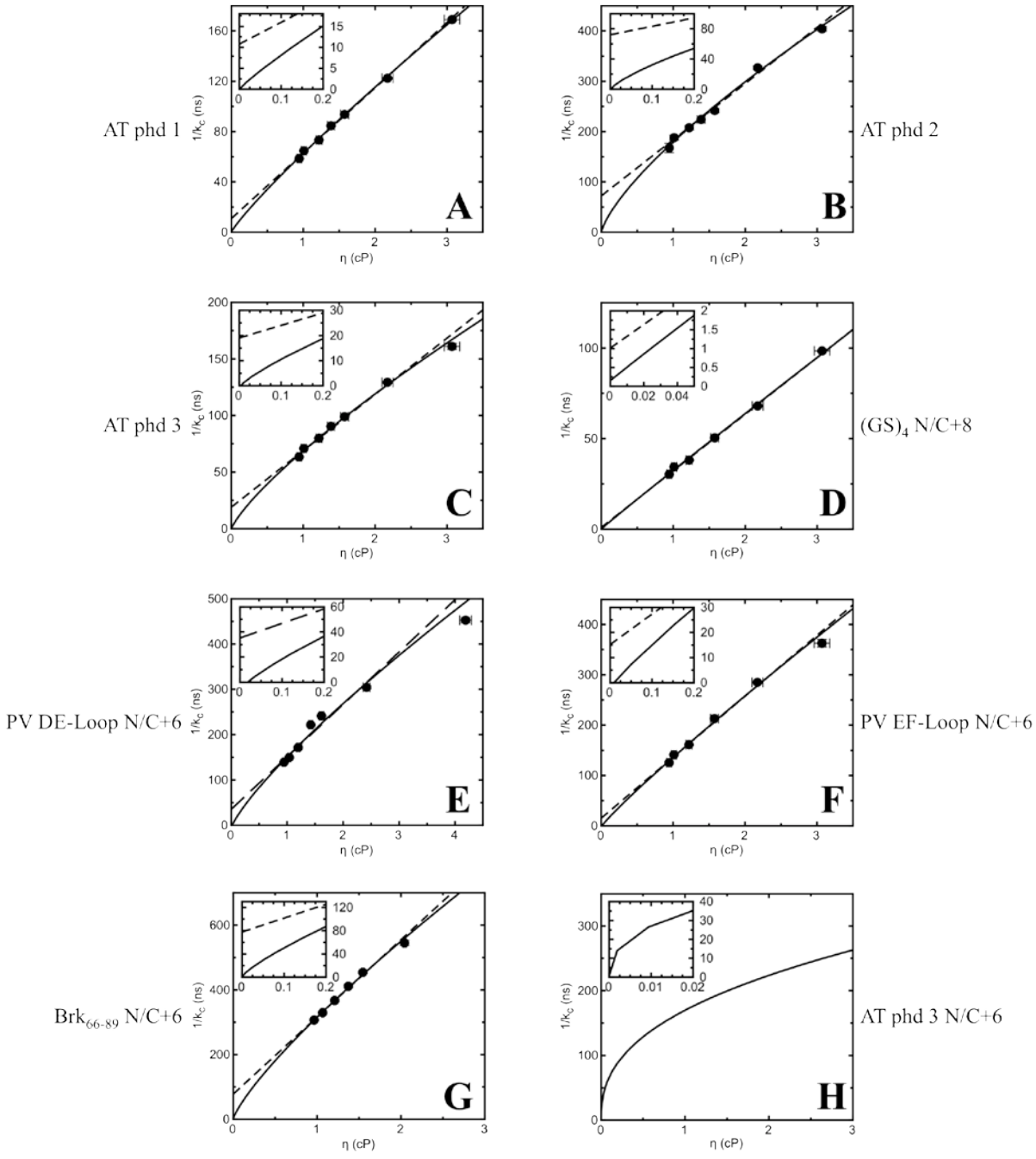


Figure 4.20 Determination of τ_{int} and τ_{solv} for end-to-end and interior-to-interior loop formation. The data is fitted with equation 4.10 using a β-value of 1 (dashed line) and the independently fitted, peptide specific β-value (solid line).

Previous TTET measurements of our group revealed that loop formation time constants at different viscosities are best to fit by a power law with β-values obtained from fits of equation 4.9 to the data[105]. Using the determined β-value at the respective condition, τ_{int} was found to be zero for poly(GS) chains of different length and for the Brinker fragment 66-89[106], the DE-Loop and the EF-Loop of PV[105], independent of solvent quality or temperature. These findings were confirmed by time-resolved FRET measurements of a $(GS)_8$ peptide and the PV EF-Loop at different glycerol concentrations, where the intra-chain coefficient and the frictional coefficients for internal and solvent friction could be determined[227].

In contrast to τ_{int}, τ_{solv} depends on the amino acid sequence, the temperature and the quality of the solvent. Contributions from solvent friction are lower in poly(GS) chains compared to natural sequences and generally high in good solvents, low in poor solvents and intermediate in water[105].

In this work, we wanted to examine internal and solvent friction in three peptide fragments derived from the intrinsically disordered protein AT phd. In addition, we wanted to test the effect of end extensions on frictional forces. TTET was performed in different glycerol/buffer mixtures at pH 7 and 22.5 ºC. The inverse rate constants for loop formation ($\tau = 1/k$) were plotted against the viscosity and the data was fit using equation 4.10 (see Figure 4.20). Dashed lines represent linear fits to the data using a β-value of -1 and solid lines describe the data by a power law with the β-values that were determined experimentally (see Table 4.5 and 4.6). The intercept at zero viscosity gives information about the internal friction (τ_{int}). It was not possible to determine reliable internal or solvent friction values for AT phd 2 *cis*Pro due to the low population of this phase. Setting $\beta = -1$ results in intercepts of $(10.8 \pm 2.8) \cdot 10^{-9}$ s for AT phd 1, $(72.2 \pm 5.1) \cdot 10^{-9}$ s for AT phd 2 *trans*Pro and $(19 \pm 2.8) \cdot 10^{-9}$ s for AT phd 3. However, these fits fail to describe the curvature in the data. A power law using experimentally determined β-values is much better suited to describe the data yielding slightly negative values for τ_{int}, which is physically meaningless. Thus, internal friction as an additive contribution seems to be absent in the dynamics of the three unfolded AT phd fragments. Chain dynamics solely scale with solvent viscosity and are governed by solvent friction (τ_{solv}) and the exponent β. Diverse τ_{solv}-values of $(59.5 \pm 2.4) \cdot 10^{-9}$ s for AT phd 1, $(173 \pm 4.2) \cdot 10^{-9}$ s for AT phd 2 *trans*Pro and $(66.3 \pm 2.8) \cdot 10^{-9}$ s for AT phd 3 were determined, indicating that solvent friction depends on the amino acid composition rather than the chain length.

It was shown earlier that the introduction of additional tails to the chain ends decrease the flexibility in the center of the peptide. We want to test whether increased chain stiffness has an influence on internal friction. Furthermore, we expect the solvent friction to be increased

due to additional dynamical effects from the end extensions. Loop formation time constants ($\tau = 1/k$) of peptides with end extensions are plotted in and fitted using equation 4.10. For the α-helix forming AT phd 3 N/C+6 fragment, the reciprocal microscopic rate constants for loop formation were fitted. Again, the linear fits (dashed lines) result in intercepts that suggest the existence of internal friction but fail to describe the curvature of the data. Therefore, it is essential to use the power law with β-values determined before. Using this strategy, interior-to-interior loop formation yields τ_{int}-values that are slightly negative or zero within the error (see Figure 4.21), indicating that for loop formation in unfolded polypeptides, even in chains of high stiffness, internal friction as an additive contribution does not exist.

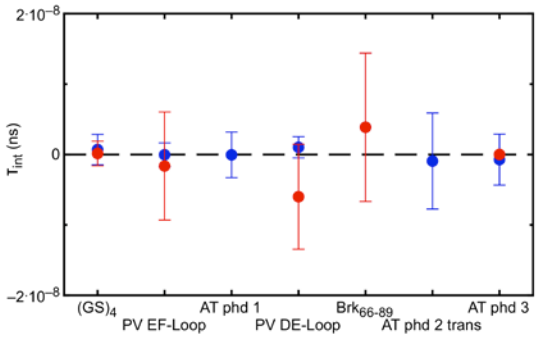


Figure 4.21 Determined values for τ_{int} using equation 4.10 and fitted β-values. Peptides without end extensions are shown in blue, data obtained from peptides with end extensions are represented in red.

This result is inconsistent with findings from single molecule FRET-measurements in unfolded and intrinsically disordered proteins[239], where internal friction was found as an additive contribution to the reconfiguration time, correlating with the compactness of the unfolded protein and approaching zero at high denaturant concentrations. However, these findings are based on linear fits to the data, discarding curvatures in the viscosity dependences. Furthermore, they apply two different polymer models for the distance distribution (random coil model) and the reconfiguration times (Rouse model), which is physically inconsistent. Other groups that have found additive contribution from internal friction to protein folding distinguishing between internal friction and activation energy terms also assumed linearly related viscosity dependences[91, 94].

According to these findings, internal friction does not contribute additively to solvent friction, independent of chain length, chain flexibility, amino acid composition or solvent quality. However, loop formation in unfolded polypeptides of natural sequences or short poly(GS) chains is not a completely diffusion-controlled process, as indicated by β-values deviating from -1. Reasons might be steric hindrance during rotations about the bonds in the

polypeptide chain, the transient formation of inter- and intramolecular hydrogen bonds or other short-lived interactions that must be broken upon folding. These effects depend on the amino acid composition and are more pronounced in natural sequences compared to poly(GS) chains and increase with decreasing chain flexibility. As shown in 4.1.2, unfolded polypeptides exhibit high activation energies upon loop formation. Long, highly flexible poly(GS) chains, where internal friction is expected to be low, show E_A-values of 5 kJ/mol. The E_A-values increase with decreasing chain length; in natural sequences activation energies are even higher and increase with decreasing chain flexibility. Thus, internal friction seems to be no physically additive contribution but might be included in the activation energy. As shown in Figure 4.22, E_A correlates with the sensitivity of the reaction to solvent viscosity (β). In the stiffer $(S)_n$-peptides and in most natural sequences, E_A increases faster with increasing β than in $(GS)_n$ due to increased steric and enthalpic effects from the amino acid side chains.

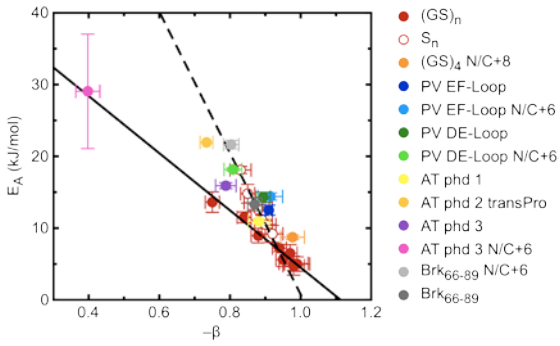


Figure 4.22 Correlation between activation energy (E_A) and sensitivity to solvent viscosity (β). Data for (•) and (◦) obtained from[106].

Thus, internal friction affects the β-values. It is therefore essential to fit protein folding time constants by a power law with peptide-specific β-values.

In contrast to internal friction, solvent friction can be directly determined by measuring viscosity dependencies and fitting of loop formation time constants using equation 4.10 with τ_{solv} representing the slope of the fitting curve. Figure 4.23 represents solvent friction time constants (τ_{solv}) for loop formation in unfolded polypeptides with and without end extensions. Peptides with end extensions experience solvent friction 2.3 – 2.7-fold higher than without end extensions. This effect is similar for all observed peptides independent of loop length or flexibility except $Brk_{66\text{-}89}$. For $Brk_{66\text{-}89}$, the effect is significantly smaller with solvent friction in interior loop formation 1.3-fold higher than in end-to-end loop formation.

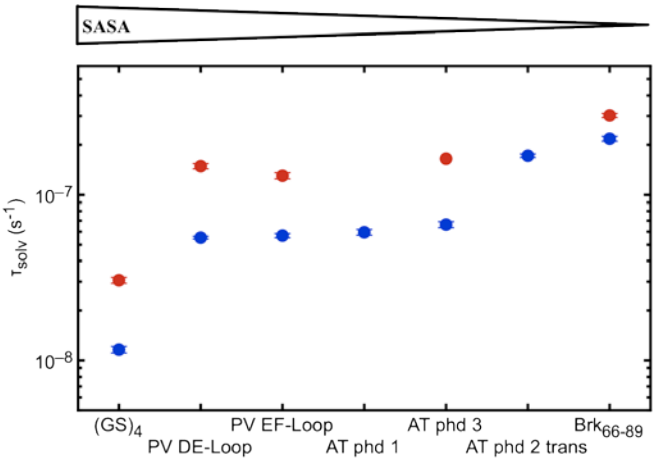


Figure 4.23 Effect of end extensions on solvent friction time constants for peptides with different solvent accessible surface areas (SASAs). Peptides without end extensions are shown in blue, data obtained from peptides with end extensions are represented in red.

Peptide	**SASA (Å^2)**	**τ_{solv} (10^{-9} s)**
$(GS)_4$[a]	1469.8	11.6 ± 0.5
PV DE-Loop[b]	2681.0	55.2 ± 1.0
PV EF-Loop[b]	2769.0	56.8 ± 1.6
$(GS)_4$ N/C+8	3073.2	30.6 ± 1.3
AT phd 1	3339.0	59.5 ± 2.4
AT phd 3	3409.9	66.3 ± 2.8
AT phd 2 transPro	3689.9	173 ± 4.2
Brk_{66-89}[c]	4123.5	217.8 ± 7.4
PV DE-Loop N/C+6	4332.7	149.5 ± 5.5
PV EF-Loop N/C+6	4423.2	130.7 ± 5.7
Brk_{66-89} N/C+6	5556.3	302.2 ± 8.6
AT phd 3 N/C+6	5728.3	166 ± 6.5

Table 4.7 Solvent friction times fitted according to $\boldsymbol{\tau = \tau_{int} + \tau_{solv} \cdot \left(\frac{\eta}{\eta_0}\right)^{-\beta}}$ and SASA of model peptides and naturally occurring protein fragments with and without end extensions. [a]Data taken from[22], [b]data taken from[105], [c]data taken from[106].

The same difference was found for loop formation rate constants in water, suggesting that the effect of end extensions is mainly due to increased solvent friction. As shown above, additional tails increase the stiffness in the chain center. As a consequence, one would expect higher barriers for loop formation due to increased enthalpic effects and internal friction. However, these effects might be compensated by a smaller loss in entropy due to a more restricted conformational space.

Solvent friction arises from interactions of the polypeptide chain with the solvent. As a consequence, solvent friction rate constants should correlate with the solvent accessible surface area (SASA). SASAs of unfolded peptides were calculated using the mean accessible surface area (ASA)-values for the individual amino acids obtained from[212] and plotted against τ_{solv} in Figure 4.24. For poly(GS) peptides, τ_{solv} linearly increases with increasing SASA. Stiffer poly(Ser) chains exhibit slightly higher solvent friction time constants compared to $(GS)_n$ chains of the same SASA. τ_{solv}-values for unfolded polypeptides deriving from natural sequences are similar to those of $(GS)_n$ of similar SASA. This is also true for peptides with end extensions, which indicates that chain stiffness does not influence solvent friction.

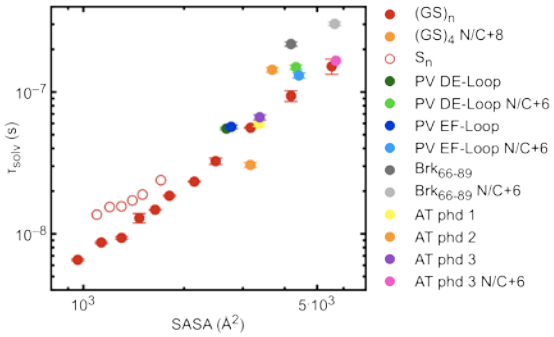


Figure 4.24 Solvent friction time constants of model peptides and naturally occurring protein fragments with and without end extensions plotted against the solvent accessible surface area. Data for (•) and (◦) obtained from[106].

The two peptides deriving from the Brinker protein exhibit higher τ_{solv}-values compared to poly(GS) chains of similar SASA. These fragments have a high ratio of positively charged residues that lead to more extended chains, which might increase the experienced solvent friction.

4.1.5 Conclusion

We measured contact formation in unfolded polypeptide chains via TTET to gain insight into the earliest steps of protein folding. As indicated by an activation energy of $E_A \approx$ RT and an inverse dependence of the bimolecular rate constant to solvent viscosity (β = -1), TTET from xanthone to naphthalene is a diffusion-controlled reaction[21]. Thus, TTET yields absolute rate constants, viscosity dependencies and free energy barriers of loop formation reactions. Contact formation in flexible, long (more than 15 peptide bonds) poly(GS) chains is fully diffusion-controlled, containing mainly entropic barriers that can be overcome by Brownian motion. Loop formation in shorter poly(GS) chains and natural sequences is limited by intrinsic enthalpic barriers as indicated by activation energies $E_A >$ RT and non-linear viscosity dependencies. These barriers arise from steric hindrance during bond rotations or the formation of inter- and intramolecular hydrogen bonds or other short-lived interactions that must be broken upon loop formation[106].

In unfolded natural sequences, β and E_A depend on the chain flexibility; E_A increases with increasing chain stiffness while the sensitivity of the polypeptide chain to solvent viscosity decreases ($-\beta < 1$). On the other hand, entropic contributions to the free energy barrier decrease with increasing chain stiffness due to a more restricted conformational space. These two opposing effects compensate each other, which lead to loop formation rate constants that only depend on the chain length and are comparable to poly(Ser) chains of similar length. Exceptions are AT phd 2 and $Brk_{66\text{-}89}$; loop formation in AT phd 2 is enthalpically disfavored compared to other natural sequences due to the influence of a *trans* prolyl peptide bond or residual structure and in $Brk_{66\text{-}89}$, repulsing interactions of positively charge residues lead to a more extended chain and hence a higher entropic barrier for loop formation. Therefore, loop formation in these peptides is significantly decelerated compared to poly(Ser) chains and natural sequences of similar length.

During protein folding, end-to-end contact formation events are rather uncommon. A more relevant case is interior contact formation between residues within the polypeptide chain. Therefore, we tested the influence of additional tails to loop formation barriers. End extensions decelerate loop formation in all observed unfolded polypeptide chains. Peptides with end extensions exhibit loop formation rate constants 2.2 – 2.5-fold slower than the same sequence without end extensions. It was shown that the flexibility of polypeptide chains is increased towards the chain ends and decreased in the chain center[217]. Thus, in the presence of end extensions, contact points for loop formation are located in stiffer chain segments. As a consequence, enthalpic contributions to interior loop formation barriers are increased

compared to end-to-end loop formation. This effect becomes more pronounced with increasing chain stiffness in the loop region as expected from polymer theory. Entropically, loop formation is favored by the attachment of tails, since increased chain stiffness leads to a restriction of the conformational space and hence to a decreased loss in entropy upon loop formation. These opposing effects sum up to similar differences k_c between end-to-end and interior loop formation independent of chain length or flexibility. Again, this is not true for Brk_{66-89}, where loop formation is entropically immensely favored upon the introduction of end extensions, resulting in rate constants that are only 1.3-fold slower than without end extensions.

It was proposed that in addition to solvent friction, contributions from internal friction affect the relaxation of a polymer towards its equilibrium, which could be determined by kinetic measurements at different solvent viscosities and extrapolation to zero solvent viscosity[93]. Researchers disagree whether protein folding time constants at different viscosities should be described setting $\beta = -1$ or using β-values that were experimentally determined from viscosity dependences. Internal friction has been reported to equal zero in loop formation kinetics following a linear viscosity[96]. On the other hand Soranno et al report a non-zero internal friction for the reconfiguration times of unfolded proteins upon setting $\beta = -1$[239]. Our results show that only a power law with independently fitted β-values for every peptide can correctly describe the curvature of the data. Using this strategy, loop formation time constants extrapolate to a time constant equal zero in the absence of solvent viscosity. This is consistent with findings from time-resolved FRET measurements at different glycerol concentrations[227]. According to our findings, rather than contributing additively to solvent viscosity, internal friction seems to increase the activation energy and to modulate the effect of solvent viscosity on the dynamics of loop formation.

4.2 The formation of turns in model polypeptides studied by TTET

Besides determination of loop formation rates in unstructured peptides, TTET can also be used to investigate folding dynamics of secondary structure elements. In previous studies of our group, TTET was applied to determine the microscopic rate constants for folding, unfolding and loop formation of α-helical model peptides[98, 109]. In this work we want to use TTET to gain knowledge about the formation of turns. Besides α–helices and β–sheets, turns are the third classical secondary structures and the most observed non-repetitive structure elements in globular proteins[113]. They are defined as the regions where the amino acid chain folds back on itself by nearly 180 degrees and therefore allow the protein to form a compact globular state. It is controversially discussed whether turns play an active or passive role in protein folding. From an active point of view, turns may serve as nucleation sites for protein folding and are formed early in the folding process[149, 150]. However, turn formation might also be a consequence of the association of nonlocal interactions between α-helices or β-sheets. Here, turns can only form after other regions have developed[151-153].

4.2.1 Testing for the existence of a disfavored bridge in the Ramachandran plot by experimental examination of β-turn formation

The Ramachandran plot visualizes backbone dihedral angles of amino acid residues in a protein structure[7]. Each peptide unit has only two degrees of freedom, specified by its backbone torsion angles ψ and ϕ, so all sterically allowed conformations of a dipeptide can be described by a two-dimensional plot (for the alanine dipeptide see Figure 4.25). Assuming a hard sphere atomic model, a conformation is not allowed if it results in an atomic clash. This eliminates a large part of the conformational space. The “bridge region” is defined as the bottleneck on the left side of the plot ($\phi < 0^{o}$ and $-20^{o} < \psi < 40^{o}$). The conventional Ramachandran plot can be modified, by applying a hydrogen-bonding requirement as an additional energetic criterion. This would mean that any backbone polar group, which is shielded from solvent and therefore has no hydrogen-bonding partner would be disfavored by about 5 kcal/mol relative to other sterically allowed conformations. As a result, a major fraction of the bridge region shown in Figure 4.25 would be eliminated[8].

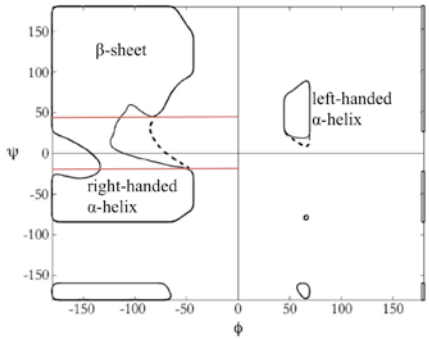


Figure 4.25 Ramachandran plot of an alanine dipeptide. Allowed conformations are within the dashed lines. The solid line shows the restricted region, when hydrogen-bonding constraints are taken into account. The bridge region is depicted inside the red lines. Figure adopted from[8].

Theoretically, when a residue at position *i* is situated in the disfavored bridge, the N-H at (*i*+1) is rendered inaccessible to solvent and can only be satisfied by hydrogen bonds to polar groups other than water. Under native conditions, when intramolecular hydrogen bonds are favored, the accessible conformational space would be expanded because the backbone- or side chain-carboxyl group on another residue could help out being the hydrogen bond partner. Conversely, unfolding conditions would deplete the population in the disfavored bridge.
In order to experimentally verify this proposition, we searched for a structured model peptide with dihedral angles ψ and ϕ positioned in the disfavored bridge. Corresponding to the alanine dipeptide, the glycine dipeptide has an additional disfavored region for $\phi > 0$[238] (see Figure 4.26).

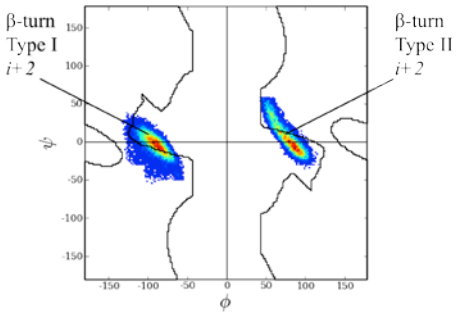


Figure 4.26 Ramachandran plot of the glycine dipeptide. Colored regions represent values for ϕ and ψ angles adopted by residue *i+2* in β-turns type I and type II with high propensity colored in red and low propensity in blue.

Glycine is fundamentally different from all other amino acids in that it lacks a side chain, which eliminates many steric clashes. Therefore, glycine can adopt ϕ and ψ angles in all four quadrants of the Ramachandran plot. Hence it frequently occurs in turn regions of proteins where any other residue would be sterically hindered. The most common type of turns is the

β-turn. It contains four residues and was originally identified by Venkatachalam in 1968[115]. Based on ϕ- and ψ-angles of the residues *i+1* and *i+2*, he proposed three distinct conformations referred to as type I, II and III along with their mirror images I', II', and III' with reversed ϕ- and ψ-values. Each could form a hydrogen bond between the backbone's C=O(*i*) and N-H(*i+3*). In 1973, examining the growing number of protein structures, Lewis et al. found that 25% of β-turns did not possess the intraturn hydrogen bond suggested by Venkatachalam[116]. Glycine residues situated in position *i*+2 of both β-turns type I and type II adopt dihedral angles positioned in the disfavored bridge of the Ramachandran plot. However, β-turns type I may be likely confused with type III β-turns and turns of α-helices. That is why we chose β-turn type II structure as experimental system to answer the following question: Is the disfavored bridge unoccupied in the absence of a hydrogen-bonded β-turn? Or to reword: Does the test peptide form an C=O(*i*) - N-H(*i*+3) hydrogen-bonded type II turn in buffer and, if so, can the turn be abolished in denaturing solvent?

Over the last two decades, researchers examined residue-specific turn propensities for each turn type using statistical analysis of known protein structures[117, 131], directed evolution and phage-display experiments[132]. Proline, as the most restricted residue, was found to occur most frequently on position *i+1*, often followed by a glycine residue on position *i+2*[117]. We synthesized peptides with the sequence Xan-Xaa-Pro-Gly-Yaa-Nal (PG-peptides) and performed TTET measurements under folding and denaturing conditions. Additionally, we recorded far-UV CD spectra of the labeled peptides and of those, where we substituted the TTET chromophores by other amino acids. Furthermore, NMR measurements were performed to gain information about hydrogen bonding in the turn.

4.2.2 β-turn formation in PG model peptides

We used TTET to measure the dynamics of β-turn formation in three PG model β-turn peptides. The proline residue at position *i*+1, because of its cyclic structure is ideally suited for the β-turn, and glycine at position *i*+2 avoids steric interactions between its side chain and the carbonyl oxygen residue of *i*+1. We chose isoleucine (Ile), serine (Ser) and alanine (Ala) as flanking residues *i* and *i*+3 (see Table 4.8). Ile is highly hydrophobic, while the Ser amino acid is hydrophilic but uncharged. Ala bears a neutral side chain. All 3 amino acids have similar values for middle-rate turn propensity at the respective positions[117].

Peptide	sequence
IPGI	Xan-IPGI-Nal-SG-$NH3^+$
SPGS	Xan-SPGS-Nal-SG-$NH3^+$
APGA	Xan-APGA-Nal-SG-$NH3^+$
SGPS	Xan-SGPS-Nal-SG-$NH3^+$

Table 4.8 Amino acid sequence of model β-turn peptides investigated by TTET. The triplet donor Xan is attached at the N-terminus and the triplet acceptor Nal in the vicinity of the C-terminus. Chain flexibility was increased in the proximity of the resin in the course of SPPS by Ser-Gly at the C-terminus.

Figure 4.27 on the left shows the expected β-turn type II structure of IPGI with the predicted backbone hydrogen bond between the CO of residue *i* and the NH of residue *i*+3. However, a β-turn is a crowded motif. The representation on the right might be more realistic.

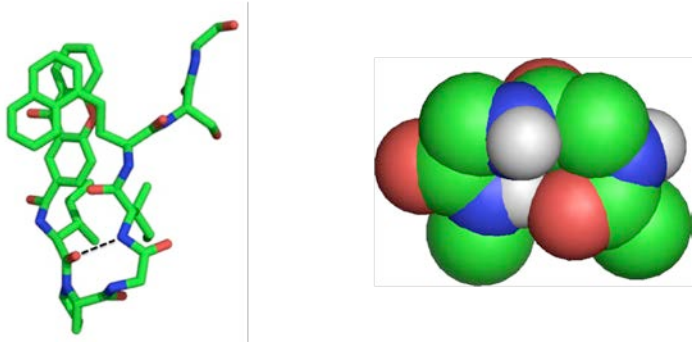


Figure 4.27 Schematic structure of IPGI. Carbon atoms colored in green, nitrogen atoms colored in blue and oxygen in red. The dashed line represents a hydrogen bond predicted by[115].

SGPS serves as a reference peptide that should not adopt β-turn type II structure since Pro on position *i*+2 has a β-turn type II propensity of 0. However, Gly at position *i*+1 shows high frequency of occurence in the enantiomeric β-turn type II'[117].

Figure 4.28 compares the far-UV CD spectra of all observed DA peptides. The far-UV CD spectrum of all three PG model peptides IPGI DA, SPGS DA and APGA DA in phosphate (PO_4) buffer displays two maxima of the ellipticity at ~200 nm and ~230 nm, respectively. Additionally, a minimum of ellipticity in the near-UV region at ~260 nm is observed. The CD band at ~200 nm is characteristic of the peptide backbone folded in a type II β-turn[126, 129]. A strong maximum of ellipticity at ~230 nm and a minimum at ~260 nm was also found in previous experiments of our group for a $(Pro)_3$-peptide and is ascribed to the formation of an excimer between the TTET labels. The pairwise interaction between aromatic molecules can lead to the formation of an excimer (from excited dimer). If one of the molecules is in its electronically excited state (as a result of light absorption), and it is allowed to approach a

ground state species with the appropriate orientation, a stable dimer will form. The dimer is formed reversibly; that is, it can dissociate back into an electronically excited monomer and a ground state species[240]. The far-UV CD spectra of β-hairpin peptides including direct cross-strand Trp–Trp interactions such as the Trp-zipper motif have been reported to exhibit a similar band with a positive maximum at 227–229 nm that was addressed to the edge-to-face interaction between the aromatic chromophores[241-246]. A far-UV CD spectrum that is equally shaped as the spectra of our PG model peptides was found for the β-hairpin forming miniprotein CLN025. It includes a β-turn type II of the sequence Asp-Pro-Glu-Thr with a hydrogen bond formed between Asp (turn residue *i*) and Thr (*i*+3). Furthermore, the β-hairpin is stabilized by an excimer between the side chains of a tyrosine and a tryptophan residue in the flanking strands[247].

As shown in Figure 4.27 on the right, a β-turn is a highly crowded motif. β-turn formation in DA peptides might therefore lead to a close packing of the TTET chromophores. TTET chromophores have been introduced in α-helices in *i*, *i*+4 spacing, which allows close contact in the helical state. However, CD spectra of these peptides did not give evidence for excimer formation[248]. The presence of an excimer band in the far-UV CD spectrum of model β-turn peptides therefore indicates high density and rigidity.

IPGI exhibit higher maxima for the two bands compared to SPGS and APGA, indicating higher β-turn content. This might be due to the longer side chain of Ile compared to Ser and Ala. Bulky amino acids lead to a restriction of the conformational space and therefore stabilize sterically favorable conformations.

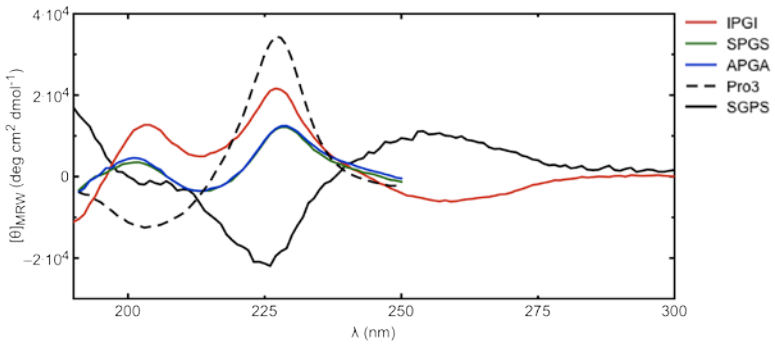


Figure 4.28 Far- and near-UV CD spectra IPGI, SPGS, APGA and SGPS. The maximum of the ellipticity at ~ 230 nm corresponds to excimer formation when the TTET chromophores come in a "sandwich-like" conformation. Previously this was also found by our group for a Xan-$(Pro)_3$-Nal peptide.

The shape of the CD spectrum of SGPS DA resembles a mirror image of the β-turn PG model peptides. However, the two bands addressed to the excimer are slightly blue-shifted and the turn band at ~200 nm is not existent. Nevertheless, the existence of strong CD bands at ~230

nm and ~260 nm indicates the formation of a dense structure that leads to an excimer formed between Xan and Nal and situated in an enantiomeric environment compared to PG peptides.

TTET in the PG model peptides was measured in 10mM potassium phosphate at pH 7 and 22.5 ºC. The time-based change in absorbance at 590 nm for IPGI and APGA is depicted in Figure 4.29, A. In addition to the donor-acceptor peptide (DA), which yields both TTET chromophores we also examined a donor-only variant (DO), where Nal was replaced by Ala. Data from the DO peptide could be described by a double exponential fit where the main phase corresponds to the intrinsic lifetime of xanthone and the faster phase with low amplitude corresponds to xanthone quenching due to the formation of small aggregates. A dramatic loss in amplitude was found for the dynamics in the DA peptide compared to DO due to a fast reaction occurring in the dead time of our experimental setup (~ 10 ns). This phase amounts ~ 90% of the DO amplitude and is due to instant TTET in the main subset of conformations, probably due to the formation of turn structure. The detectable part of the DA dynamics exhibit triple exponential kinetics with the main phase corresponding to a loop formation reaction that brings the TTET chromophores in close contact without adopting β-turn type II structure. Additionally, two phases with lower amplitudes were found that are also present in the DO kinetics and are due to amounts of aggregated peptide where the labels are unable to form contact. TTET measurements of SPGS and SGPS in PO_4 buffer were not possible due to low solubility. However, for these peptides we could determine the loop formation rate constants (k_c) in water by measuring k_c at different viscosities and linear extrapolation to the viscosity of water (see chapter 4.2.3).

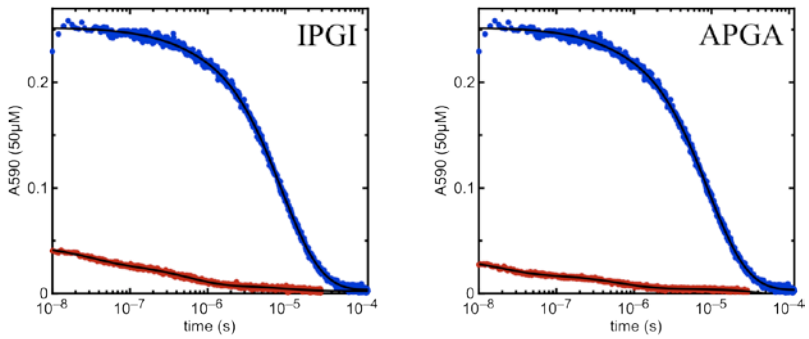


Figure 4.29 Triplet decay curves of xanthone monitored by the change in absorbance at 590 nm of (A) IPGI and (B) APGA in 10mM potassium PO_4 buffer at 22.5ºC. The DA peptide is shown in red and the DO peptide in blue. Black lines represent multi-exponential fits to the kinetics.

In previous work, the effects of prolyl and glycyl residues on chain dynamics have been investigated by measuring the kinetics of loop formation in host-guest peptides of the structure Xan(Ser)$_x$-Xaa-(Ser)$_x$-Nal-Ser-Gly, with Xaa = Pro or Gly[223]. Loop formation was shown to be slower around *trans* prolyl peptide bonds and faster around glycine residues due to their effects on the activation energy. At high viscosities, double exponential kinetics were observed for peptides with Xaa = Pro. The second kinetic phase was faster than loop formation in glycyl peptides with an amplitude of ~15% and was due to *cis* isomers of the Ser-Pro bond. The fast dynamics around *cis* prolyl bonds were ascribed to a higher Arrhenius pre-exponential factor. Loop formation rate constants in water of our β-turn model peptides are compared to k_c in SSPS[223] in Figure 4.32 (chapter 4.2.3).

4.2.3 The effect of viscosity on β-turn formation

A large part of the reaction kinetics in PG turn model peptides occur in the dead time of the TTET set up and therefore escape closer analysis. We measured TTET in the PG-peptides at different viscosities to see if the kinetics can be decelerated to time constants on the nanosecond time scale and therefore detectable by our TTET setup (see Figure 4.30). We chose glycerol as viscosifier since it has a small hydrodynamic radius, which allows for determination of the actual microscopic viscosity experienced by the polypeptide chain[106, 249]. We used glycerol concentrations between 20% and 60%, which resulted in solvent viscosities between 2 cP and 12 cP. Viscosities > 12 cP might affect the bond rotation of Xan and Nal thereby having an unpredictable influence on TTET kinetics.

Figure 4.30 shows the time-based change in xanthone absorbance at the indicated solvent viscosities. A deceleration of the kinetics in the DA peptides can be observed with increasing solvent viscosity. Simultaneously, the amplitude increases slightly. Loop formation kinetics could be described by a single exponential fit. Corresponding to measurements in PO_4 buffer, a slower phases addressed to the formation of aggregates was detected. For viscosities ≥ 5 cP an additional fast phase appeared. Because of its low amplitude of 5% – 10% we assume that this phase is due to loop formation in peptides with the prolyl bond in *cis* configuration. For the DO variants, solvent viscosity has no influence on the intrinsic Xan lifetime but the amplitude dramatically decreases with increasing glycerol concentration. This phenomenon has not been observed in DO variants of less rigid peptides, indicating that the triplet absorbance of Xan, buried in a highly dense structure, is influenced by solvent changes.

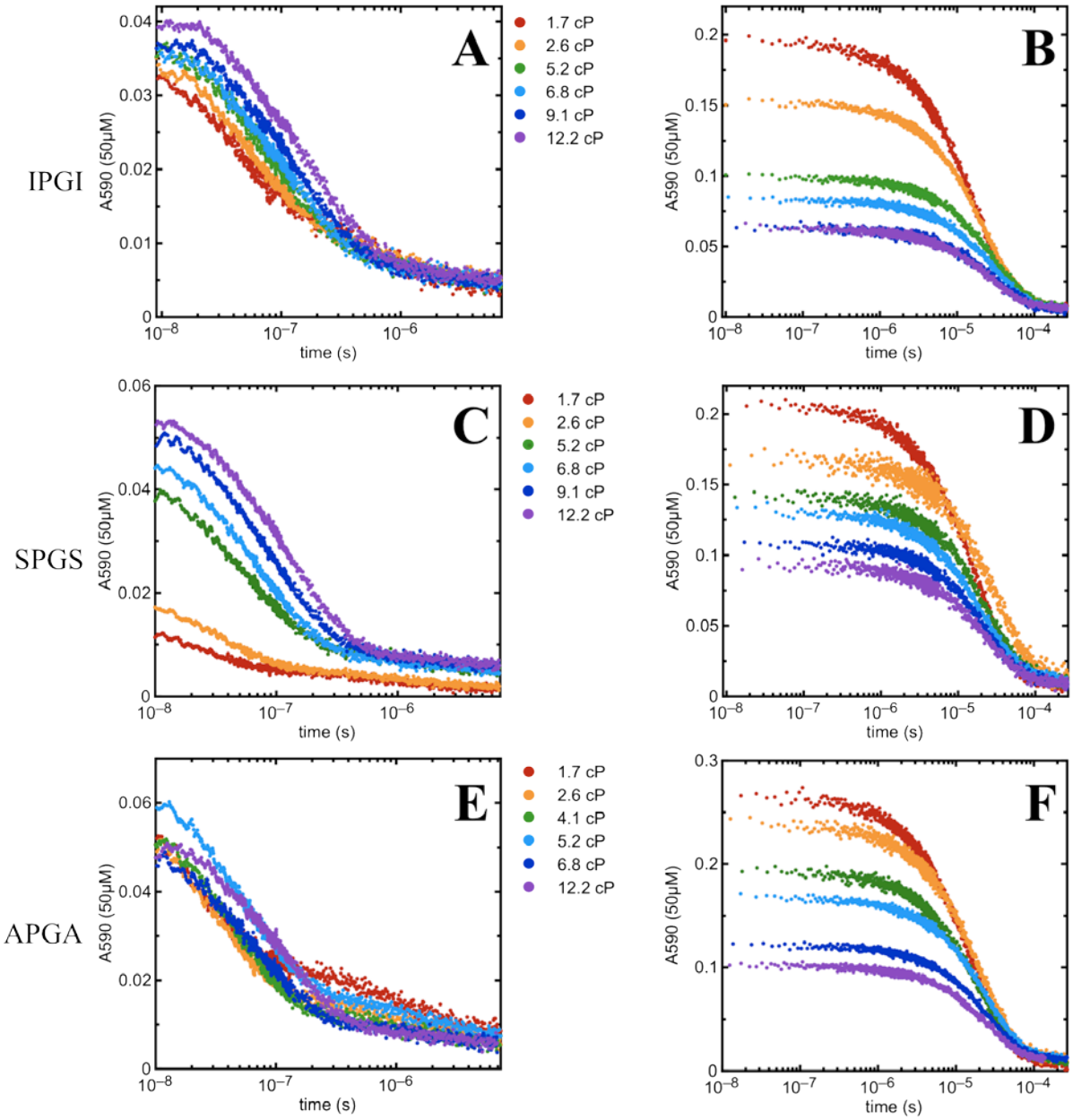


Figure 4.30 TTET measurements of PG turn model peptides at different solvent viscosities. Time dependent absorbance decay at 590 nm for DA peptides (A, C, E) and DO variants (B, D, F).

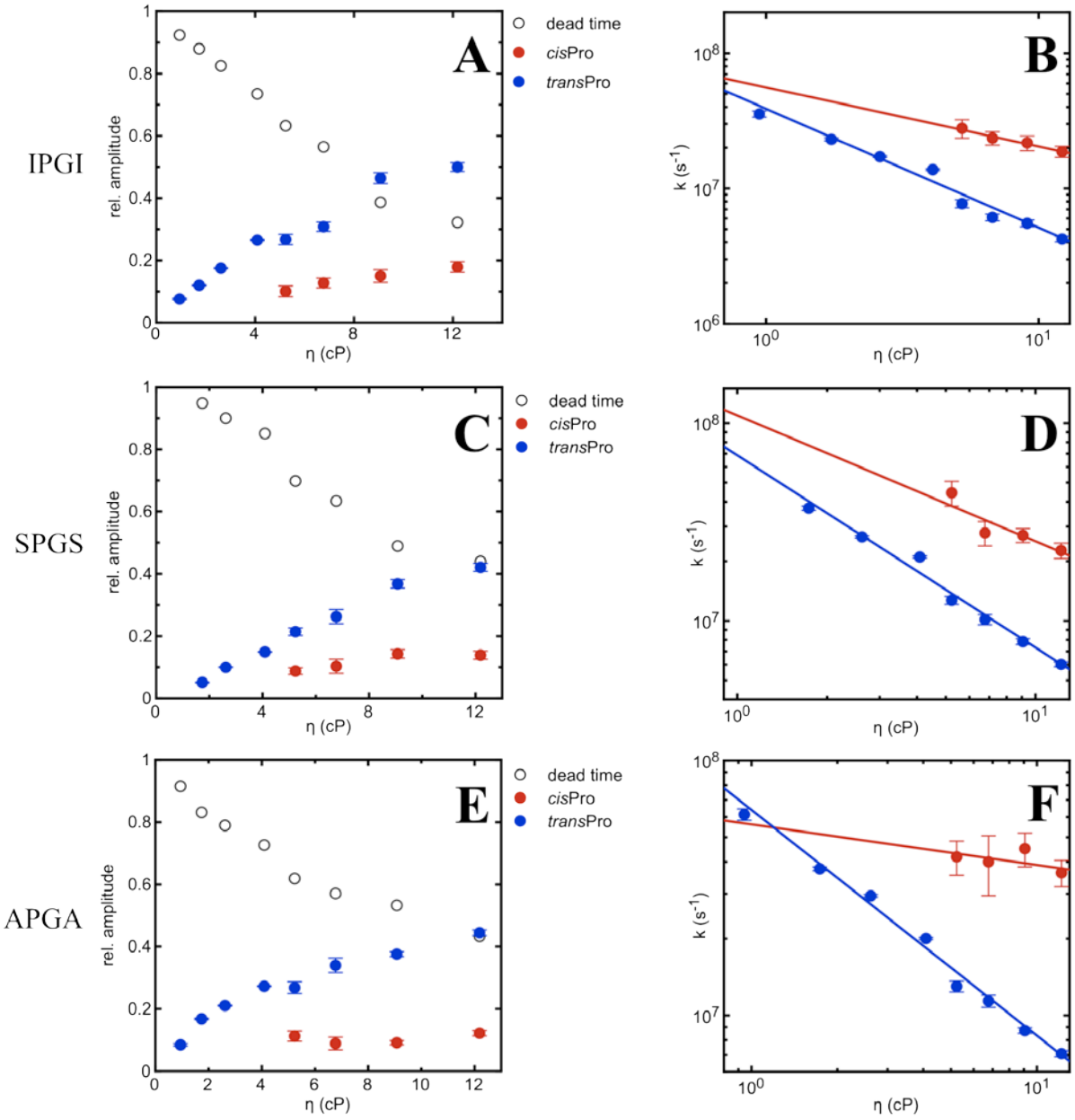


Figure 4.31 TTET measurements of PG turn model peptides at different solvent viscosities. Relative amplitudes (A, C, E) and rate constants (B, D, F) obtained from exponential fits to the data. The colors of the rate constants and their relative amplitudes are equivalent.

Hence, the dead time amplitude is significantly decreased with increasing glycerol concentration.

The rate constants for loop formation decrease linearly with increasing viscosity for all peptides (see Figure 4.31, B, D, F). The data could be fitted by equation 4.11

$$k_c = k_c^{H_2O} \cdot \left(\frac{\eta}{\eta_0}\right)^{\beta} \qquad (4.11)$$

η_0 is the reference solvent viscosity of water and k_c^{H2O} is the rate constant for loop formation at η_0. The empirical fractional exponent β reflects the sensitivity of the reaction to solvent viscosity. For β = -1, the dependence is linear, which means that the reaction is diffusion controlled. A β-value of 0 indicates that the reaction is independent of solvent viscosity.

The fitting revealed β-values for PG-peptides with the prolyl bond in *trans* configuration of -0.88 ± 0.02 for IPGI, -0.97 ± 0.03 for SPGS and -0.89 ± 0.02 for APGA. It was not possible to determine reliable β-values for peptides with the prolyl bond in *cis* configuration due to their low population. The relative amplitudes of the kinetic phases for IPGI, SPGS and APGA indicate that the β-turn content decreases with increasing glycerol concentration. The amplitude of the loop phase increases concomitantly.

Loop formation rate constants in water for PG-peptides are compared to loop formation kinetics of SSPS in Figure 4.32.

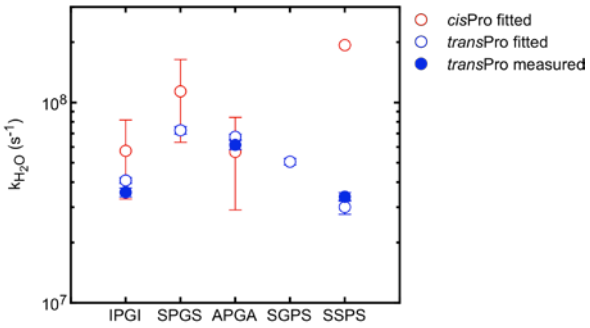


Figure 4.32 Loop formation rate constants in water for *cis* and *trans* isomers of β-turn model peptides compared to SSPS[223]. Open circles represent rate constants determined by fitting of viscosity dependences to equation 4.11, filled circles represent directly measured rate constants.

Loop formation in *trans* isomers of SPGS and APGA and SGPS is slightly faster compared to SSPS due to the high flexibility of the Gly residue and hence a lower activation energy. IPGI yields a loop formation rate constant similar to SSPS. Here, the accelerating effect of Gly

might be compensated by a lower Arrhenius pre-exponential factor due to a more restricted conformational space. Loop formation rate constants of *cis* isomers are slightly faster compared to *trans* isomers. However, the fitted values of the *cis* isomers are not reliable, due to the low population of this kinetic phase.

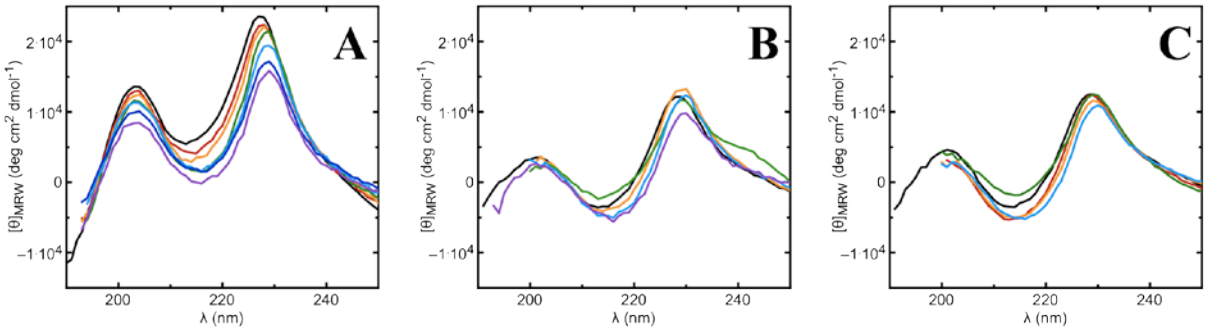


Figure 4.33 Far-UV CD spectra of (A) IPGI DA, (B) SPGS DA and (C) APGA DA at different solvent viscosities. The colors in the spectra correspond to the viscosities in Figure 4.30. The black spectrum is the reference measurement in PO_4 buffer.

The decrease in β-turn content is supported by the results from CD measurements at different solvent viscosities (see Figure 4.33). The CD bands corresponding to turn structure and excimer decrease with increasing glycerol concentration. The decrease is of similar magnitude for the excimer band and the β-turn type II band, indicating that β-turn structure in these peptides is associated with the formation of an excimer between the TTET labels Xan and Nal. The abolition of β-turn structure with increasing viscosity is surprising since glycerol is known to stabilize the folded state[250, 251]. However, glycerol molecules might electronically interact with the excimer and therefore destabilize the β-turn structure. The effect is more pronounced for IPGI compared to SPGS and APGA, probably due to a more restricted conformational space in IPGI because of the larger side chain of Ile.

For similar reasons, the addition of glycerol leads to a decrease of the turn content for SGPS as seen in the TTET measurements at different solvent viscosities (see Figure 4.34, A, B, C, D). The dead time amplitude decreases from ~90% to ~50% when the viscosity is increased. The additional *cis*Pro phase appears at viscosities ≥ 9 cP. However, from CD spectra (panel E) is seen that the minimum of ellipticity at 225 nm decreases with increasing glycerol concentration. This is surprising since an abolition of the excimer is expected to lead to an increase of the CD signal at ~230 nm.

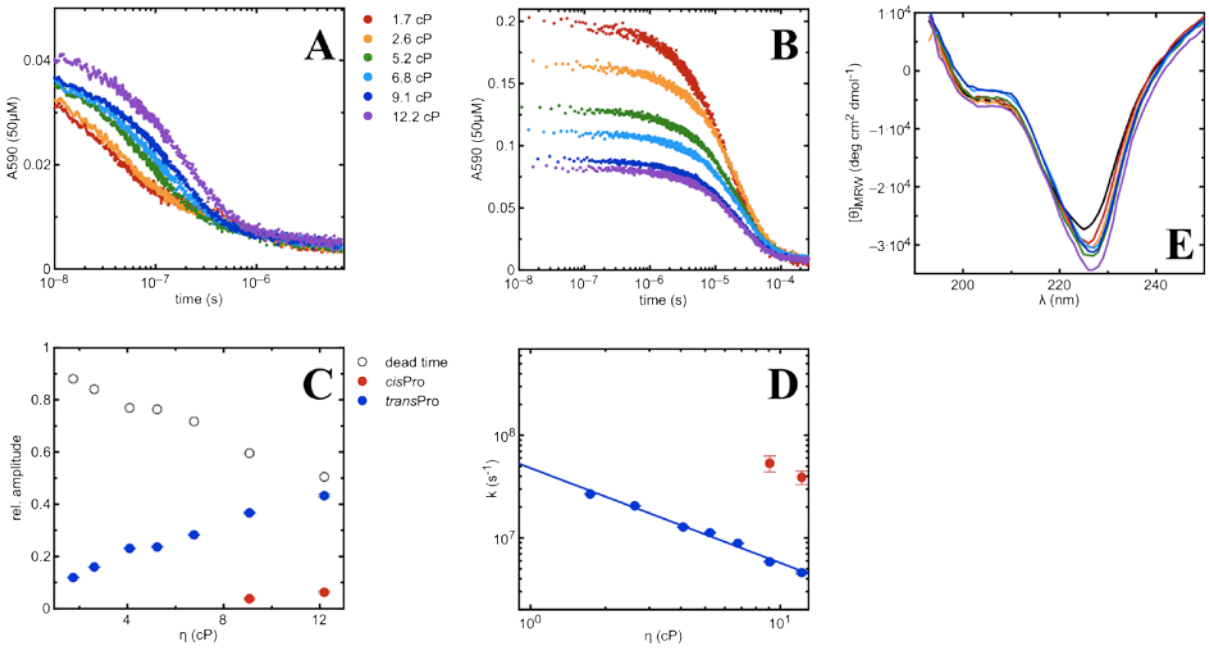


Figure 4.34 CD and TTET measurements of SGPS at different solvent viscosities. Absorbance change at 590 nm for the (A) DA variant and (B) DO variant. (C) relative amplitudes and (D) rate constants obtained from exponential fits. The colors of the rate constants and their relative amplitudes are equivalent. (E) Far-UV CD spectra; the colors of the spectra correspond to the glycerol concentrations in A. The black spectrum is the reference measurement in PO_4 buffer.

4.2.4 The effect of flanking amino acids on β-turn formation

It is controversially discussed whether β-turns play an active or passive role in protein folding. From an active point of view, turns may serve as nucleation sites for protein folding and are formed early in the folding process[149, 150]. However, turn formation might also be a consequence of the association of long-range interactions between flanking strands[151-153]. TTET and CD measurements of short PG model turn peptides revealed distinct β-turn structure in short peptides of 7 amino acids length. However, an excimer is formed between the flanking aromatic TTET chromophores that might stabilize the turn structure. Aromatic-aromatic (π-π) interactions are intermolecular forces involving π electron rich molecules and are known to play an important role in maintaining the overall structure of protein molecules and protein-DNA complexes. Formally, π-π interactions are defined as pairs of interacting aromatic residues whose centers are separated by a distance between 4.5 Å and 7 Å and the dihedral angles fall between 30º and 90º. The free energies of π-π formation lie between -0.6 and -1.3 kcal/mol[252]. We wanted to test for the requirement of stabilizing interactions for β-turn formation in PG model peptides. Therefore, we synthesized IPGI model peptides flanked by different aromatic and non-aromatic amino acids (see Table 4.9) and recorded far-UV CD spectra under folding conditions. ISGI serves as an unfolded reference peptide.

Peptide	sequence
Xan-IPGI-Nal	Xan-IPGI-Nal-SG-$NH3^+$
Xan-IPGI-Ala	Xan-IPGI-A-SG-$NH3^+$
Ac-IPGI-Nal	Ac-IPGI-Nal-SG-$NH3^+$
W-IPGI-GW	Ac-W-IPGI-GW-SG-$NH3^+$
W-IPGI-W	Ac-W-IPGI-W-SG-$NH3^+$
Xan-IPGI- F	Xan-IPGI-F-SG-$NH3^+$
F-IPGI-Nal	Ac-F-IPGI-Nal-SG-$NH3^+$
F-IPGI-G	Ac-F-IPGI-GY-SG-$NH3^+$
A-IPGI-G	Ac-A-IPGI-GY-SG-$NH3^+$
A-IPGI-A	Ac-A-IPGI-A-GY-SG-$NH3^+$
A-ISGI-G	Ac-A-ISGI-GY-SG-$NH3^+$

Table 4.9 Amino acid sequence of model β-turn peptides investigated by CD. The sequence IPGI, with high β-turn type II propensity is flanked by different aromatic and non-aromatic residues. Chain flexibility was increased in the proximity of the resin in the course of SPPS by Ser-Gly at the C-terminus.

The far-UV CD spectrum of A-ISGI-G shows the typical shape of an unfolded peptide in random coil conformation with a minimum of ellipticity around 200 nm. A-IPGI-G and A-IPGI-A reveal the same characteristic signal (see Figure 4.35, A). In these peptides, the IPGI turn sequence is flanked by residues that are not able to form stabilizing interactions.

Xan-IPGI-Nal displays the typical β-turn type II band at ~200 nm as well as the maximum of the ellipticity at ~230 nm, due to excimer formation between the TTET chromophores. The excimer signal is also found for W-IPGI-GW (see panel B), which gives evidence for a dense and rigid structure that brings the two tryptophans in close contact. However, due to the formation of small aggregates no CD signals could be recorded at wavelengths <199 nm and the presence of a β-turn type II band could not be confirmed. Furthermore, the C-terminal tryptophan is not adjacent to the IPGI sequence but a glycine residue is inserted. Hence, a closely packed configuration of the two Trps similar to the configuration of Xan and Nal in Xan-IPGI-Nal would result in a slightly different conformation.

As shown in panel C, the far-UV CD spectrum of Xan-IPGI-F exhibits the characteristic shape of β-turn type II structure[126, 129]. The aromatic interaction between xanthone and phenylalanine seems to be sufficient to induce β-turn formation. However, the single aromatic ring of Phe is not able to form an excimer with Xan. The spectrum of Xan-IPGI-A exhibits the same shape like Xan-IPGI-F with a smaller β-turn type II band. This is surprising, since Xan has no aromatic interaction partner at the C-terminus of the IPGI sequence. The CD spectra of F-IPGI-G and W-IPGI-W both show small maxima of ellipticity at 200 nm.

However, they lack the characteristic minimum at ~225 nm. It is therefore not clear if these peptides adopt β-turn structure.

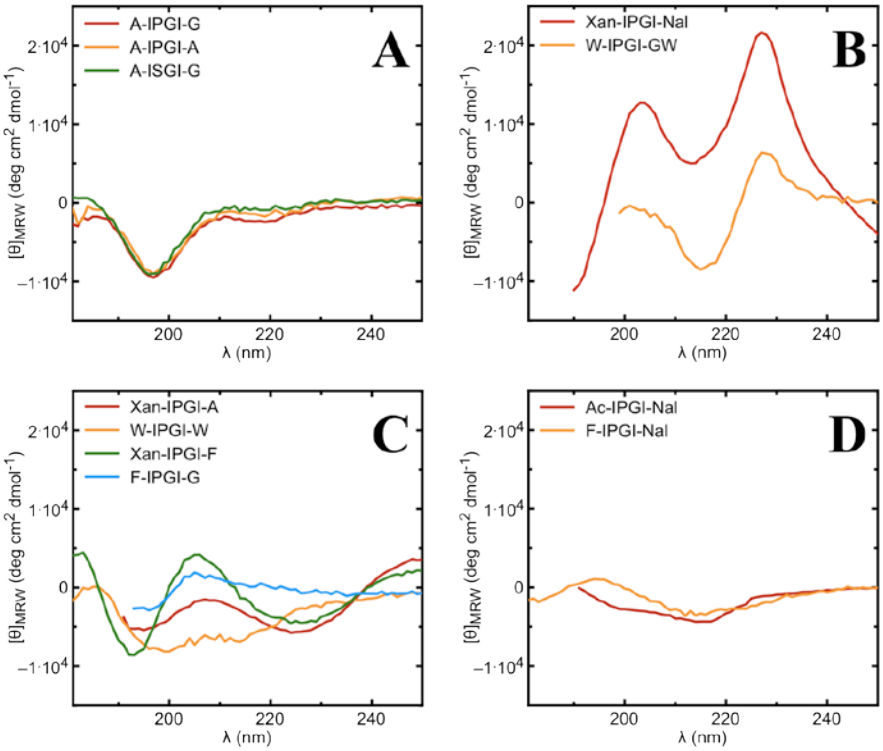


Figure 4.35 Far-UV CD spectra of IPGI derivative peptides. (A) random coil peptides, (B) peptides that show an excimer band at ~230 nm, (C) peptides that exhibit the β-turn type II band at ~200 nm but no excimer band, (D) peptides that could not be assigned to A, B or C.

For unknown reasons Ac-IPGI-Nal and F-IPGI-Nal (panel D) could not be assigned to random coil or turn structure. This might be due to an interfering CD signal that arises from the chromophore Nal.

In order to further characterize the conformations of IPGI derivative peptides we measured CD thermal transition curves at a wavelength of 204 nm, which corresponds to the maximum of ellipticity in the β-turn type II band (see Figure 4.36). Thermal melting was fully reversible for all peptides. For Xan-IPGI-Nal and Xan-IPGI-F we see a destabilization of the turn with increasing temperature (panels A, B). However, the ellipticity does not reach a limiting value, indicating the existence of residual turn structure even at 90ºC. For Xan-IPGI-A, we see an increase in signal with increasing temperature and for W-IPGI-W and F-IPGI-G, the CD signal stays on a constant value over the whole temperature range. This supports our presumption that the small maxima of ellipticity at ~200 nm in these peptides arise from conformations other than β-turn type II.

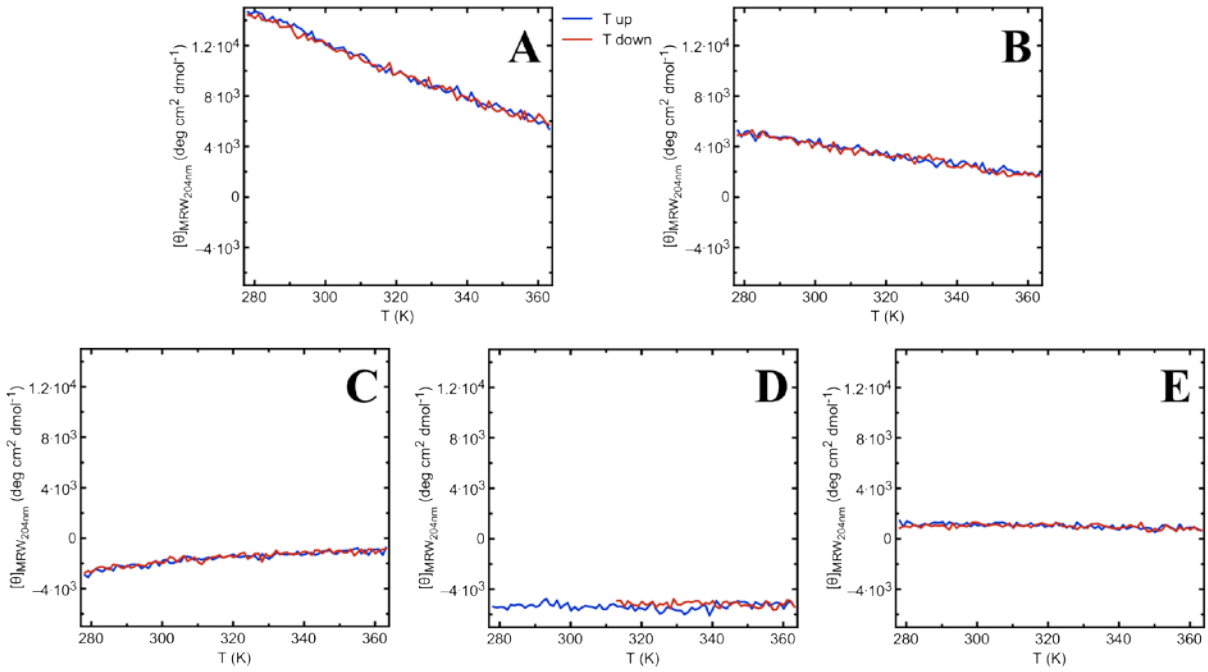


Figure 4.36 Thermal transition of (A) Xan-IPGI-Nal, (B) Xan-IPGI-F, (C) Xan-IPGI-A, (D) W-IPGI-W and (E) F-IPGI-G followed by CD spectroscopy at 204 nm.

Our findings indicate that an amino acid sequence with high turn propensity alone is not sufficient for β-turn formation but the β-turn structure can be stabilized by local interactions between aromatic residues. Stabilizing local interaction between non-hydrogen bonded tryptophans have been shown to frequently occur in β-hairpins[253]. These Trp–Trp pairs have emerged as a paradigm for the design of stable β-hairpins, such as the Trpzip peptides[241]. We looked for PG sequences in any conformation in a database of 6208 high resolution proteins (PISCES list[254]) and determined the normalized frequency of occurrence for every amino acid at the 3 positions before and 4 positions after the PG. We found 4994 PG sequences, the majority in β-turns but not all type II. However, we did not find any residue or amino acid group to be particularly preferred at one of the observed positions. Hence, from statistics, at this point we see no evidence for the need of local stabilizing interactions in PG-turns.

4.2.5 Hydrogen-bonding in β-turns examined by NMR measurements

It is controversially discussed whether a hydrogen-bonding requirement as an energetic criterion is needed to correctly calculate possible backbone dihedral angles of amino acid residues in protein structure[8, 255, 256]. If so, the N-H at Ile residue *i*+3 in a β-turn type II structure forming peptide of the sequence Ile-Pro-Gly-Ile could only be satisfied by an intramolecular hydrogen bond[8]. In order to examine hydrogen bonding in our PG-peptides we performed NMR measurements. The temperature dependence of amide-proton (HN) chemical

shifts has been shown to correlate with the presence of intramolecular hydrogen bonds. As temperature is raised, the hydrogen bond is weakened and the HN signal shifts upfield. The lengthening of average hydrogen bonds is greater for intermolecular hydrogen bonds such as those with water, than for intramolecular hydrogen bonds. Hence, amides hydrogen bonds to water will shift more upfield with temperature than internally hydrogen-bonded amides. If the temperature coefficient is more positive than -4.5 ppb/K, it is likely that the amide proton serves as an intrachain hydrogen bond donor[257, 258]. However, HN chemical shift temperature coefficients don't give information about the hydrogen bond acceptor. Fast exchanging protons give a positive temperature-coefficient when hydrogen-bonded, as observed in α-helices[259]. Hence, we assume that a positive HN chemical shift temperature coefficient in our PG-peptides indicates a stable and specific hydrogen bond that is needed for β-turn type II structure formation.

Table 4.10 summarizes residue-specific HN chemical shift temperature coefficients for five IPGI derivative peptides.

Peptide	**Residue**						
	Xaa (*i*-1)	**I (*i*)**	**P (*i*+1)**	**G (*i*+2)**	**I (*i*+3)**	**Yaa (*i*+4)**	**Zaa (*i*+5)**
Xan-IPGI-Nal				3.6			
Xan-IPGI-F		-13.1		0.7	-6		
W-IPGI-W	-10.2	-8.2		3.6	-5.1	-8.8	
W-IPGI-GW	-9.3	-6.3		-5.7	-4	-7.6	-7.1
Xan-IPGI-A		-9.4		-7.4	-6.9	-13	

Table 4.10 Residue-specific NMR HN chemical shift temperature coefficients (ppb/K) of IPGI derivative peptides under folding conditions. Xaa, Yaa and Zaa stand for the residues at position (*i*-1), (*i*+4) and (*i*+5), respectively. The amino acid sequences of the peptides are shown in Table 4.9.

We were not able to assign HN chemical shift temperature coefficients to every residue due to low solubility of the peptides and therefore weak NMR signals. Proline lacks a hydrogen on the backbone amide group because of its cyclic structure. Hence, it cannot serve as hydrogen bond donor. Furthermore, for the isoleucines it is not clear whether the temperature coefficients correspond to the N-terminal or the C-terminal residues.

For Xan-IPGI-Nal and Xan-IPGI-F, CD spectra confirmed β-turn type II structure. Surprisingly, HN chemical shift temperature coefficients ≤ -4 for the Ile residues in these peptides don't give evidence for *i* – *i*+3 hydrogen bonding, as predicted by Venkatachalam for β-turn structure[115] and expected if hydrogen-bonding requirement is applied[8]. However, in these peptides positive HN chemical shift temperature coefficients reveal that the amide of the

glycine (turn residue *i*+2) is involved in a stable hydrogen bond. The value is higher for Xan-IPGI-Nal compared to Xan-IPGI-F, which indicates that the hydrogen bond in Xan-IPGI-Nal is more populated, since the temperature coefficients reflect a population-weighted average over all conformations. The excimer forming interaction between Xan and Nal seems to be stronger than the aromatic interaction between Xan and Phe, which leads to an increased stabilization of the H-bonded structure.

Thus, our results show that the PG-peptides form tight β-turn structures, but they do not form the expected type-II β-turn. If we assume a β-turn structure, where the $i \rightarrow (i+3)$ NH is from the Gly the 4 aa turn sequence would be Xaa-IP-G. A Pro residue at the β-turn position (*i*+2) is rather uncommon[117]. However, β-turns type VI structure requires a *cis*Pro at position *i*+2. In these turns *trans*Pro at *i*+2 cause unallowed close contacts between C^{β}_{i+1} and C^{δ}_{i+2}. Only with a *cis*-prolyl peptide bond, C^{β}_{i+1} is removed far from the C^{δ}_{i+2} of the proline residue[260]. Notably, the CD curve of β-turns type VI look similar to β-turns type II[261]. However, it is also possible that the PG-peptides yield an *i* – (*i*+2) hydrogen bond, including the Ile and the Gly.

The glycine in W-IPGI-W exhibits a temperature coefficient of the same value as found for Xan-IPGI-Nal. This is surprising, because the CD spectra are very different in shape. We assume that this sequence can adopt a hydrogen-bonded structure other than Xan-IPGI-Nal and Xan-IPGI-F. For W-IPGI-GW and Xan-IPGI-A, we see no evidence for hydrogen bonding. By utilization of backbone dihedral ϕ-angles, obtained by analysis of NMR $^{3}J_{HN,H\alpha}$ coupling constants, we plan to perform computational simulations of possible conformations of the PG-peptides in order to gain information about their real structures.

Two of our IPGI model peptides (Xan-IPGI-Nal and Xan-IPGI-F) adopt hydrogen-bonded β-turn structure. However, the H-bond is not formed between the backbones C=O(*i*) and N-H(*i*+3) as predicted. Instead, the NH of Gly (*i*+2) is the H-bond donor. However, from NMR measurements we did not gain information about the H-bond acceptor and so we are not able to determine the exact conformation of the turn.

4.2.6 The Effect of Solvent Properties on β-Turn Formation

The co-solutes guanidinium chloride (GdmCl) and urea are frequently used to determine protein stability and to investigate kinetics and mechanisms of protein folding reactions. They serve as chemical denaturants and are known to unfold proteins by stabilizing the unfolded state compared to the native state[108, 262-264]. In poor solvents such as water, intramolecular interactions like hydrogen bonds are more favorable than peptide-solvent interactions and the

polypeptide chain folds. Denaturants provide a good solvent for peptides favoring peptide-solvent interactions. As a consequence, the polypeptide chain is more expanded and unfolds. In order to test the effect of the solvent on turn formation in the PG-peptides we performed TTET and far-UV CD measurements at different concentrations of urea and GdmCl. Additionally, we determined NMR HN chemical shift temperature coefficients for IPGI peptides under strong denaturing conditions (8M urea).

The presence of denaturants in buffer solutions is known to increase solvent viscosity, which affects chain diffusion[265]. Thus, the obtained rate constants (k_c) from TTET measurements have to be corrected against viscosity changes (η) using equation 4.12 and the β-values determined from TTET measurements at different glycerol concentrations (see Figure 4.31).

$$k_c' = k_c \cdot \left(\frac{\eta}{\eta_0}\right)^{\beta} \qquad (4.12)$$

The logarithm of the rate constants for loop formation in polypeptide chains was found to depend linearly on the molar denaturant concentrations of urea and GdmCl[23, 53]. The effect of denaturants on k_c can be described by equation 4.13

$$lnk_c = lnk_c^{H_2O} - m_c \cdot \frac{[D]}{RT} \qquad (4.13)$$

where k_c^{H2O} is the rate constant for loop formation in the absence of denaturants and m_c describes the strength of the influence of denaturants on loop formation.

Figure 4.37 shows the TTET kinetics measured by the decrease in xanthone absorbance at the indicated urea concentrations. A deceleration of the TTET kinetics of the DA peptides can be observed with increasing urea concentration, mainly due to the increased solvent viscosity. Simultaneously, the normalized amplitude increases. The kinetics for loop formation could be described by a single exponential fit. Again, a slower phase with lower amplitude was found corresponding to aggregated peptide. For the DO variants, urea has no quenching effect on the intrinsic lifetime of Xan but the normed amplitude is decreased with increasing urea concentration, as previously observed for glycerol. A substantial decrease of the dead time amplitude is observed, which indicates that the turn content decreases with increasing urea concentration. Relative amplitudes and viscosity corrected rate constants for loop formation in IPGI DA, SPGS and for APGA are shown in Figure 4.38.

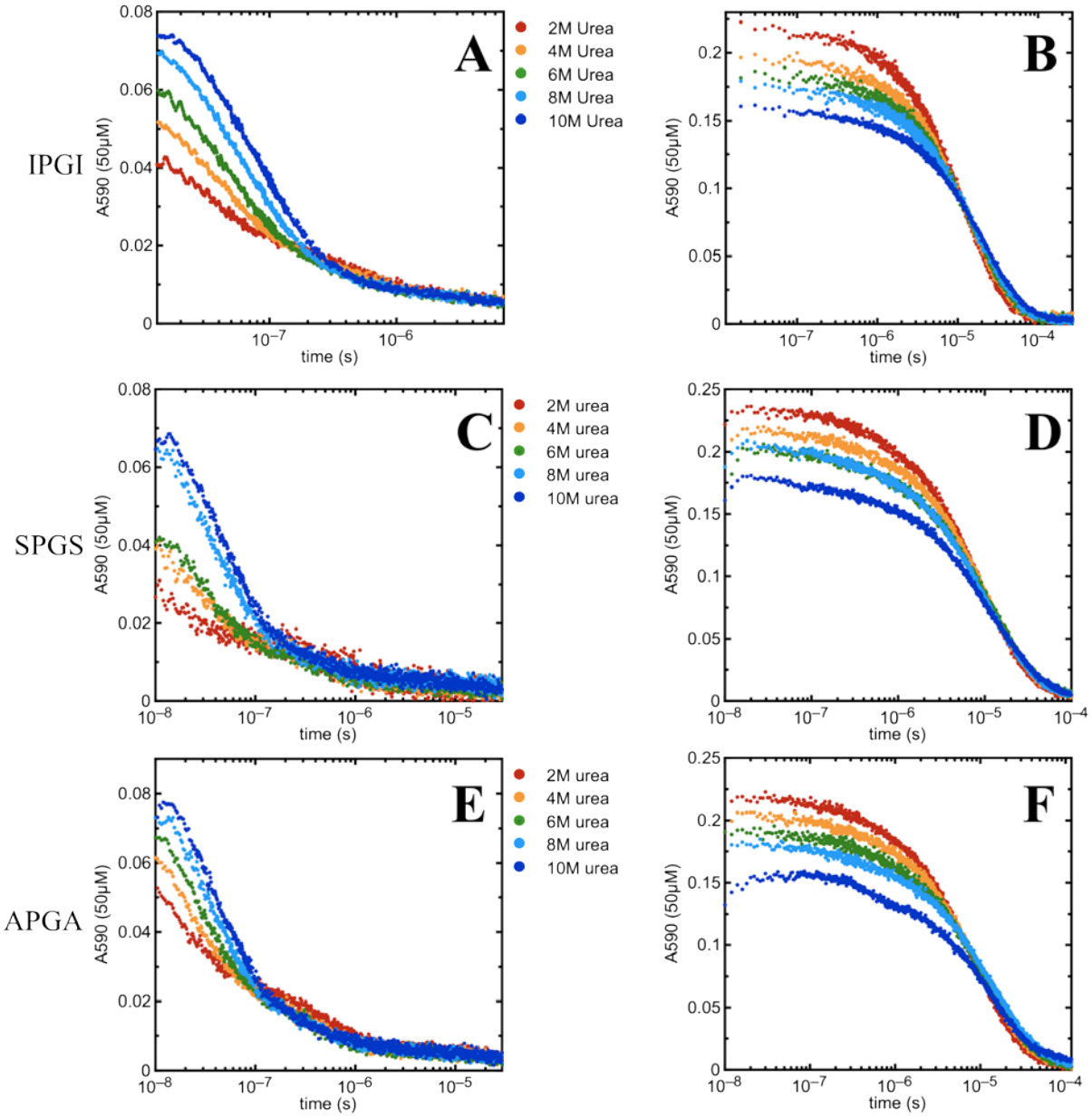


Figure 4.37 TTET measurements of PG turn model peptides at different urea concentrations. Absorbance decay at 590 nm for (A, C, E) DA variants and (B, D, F) DO peptides.

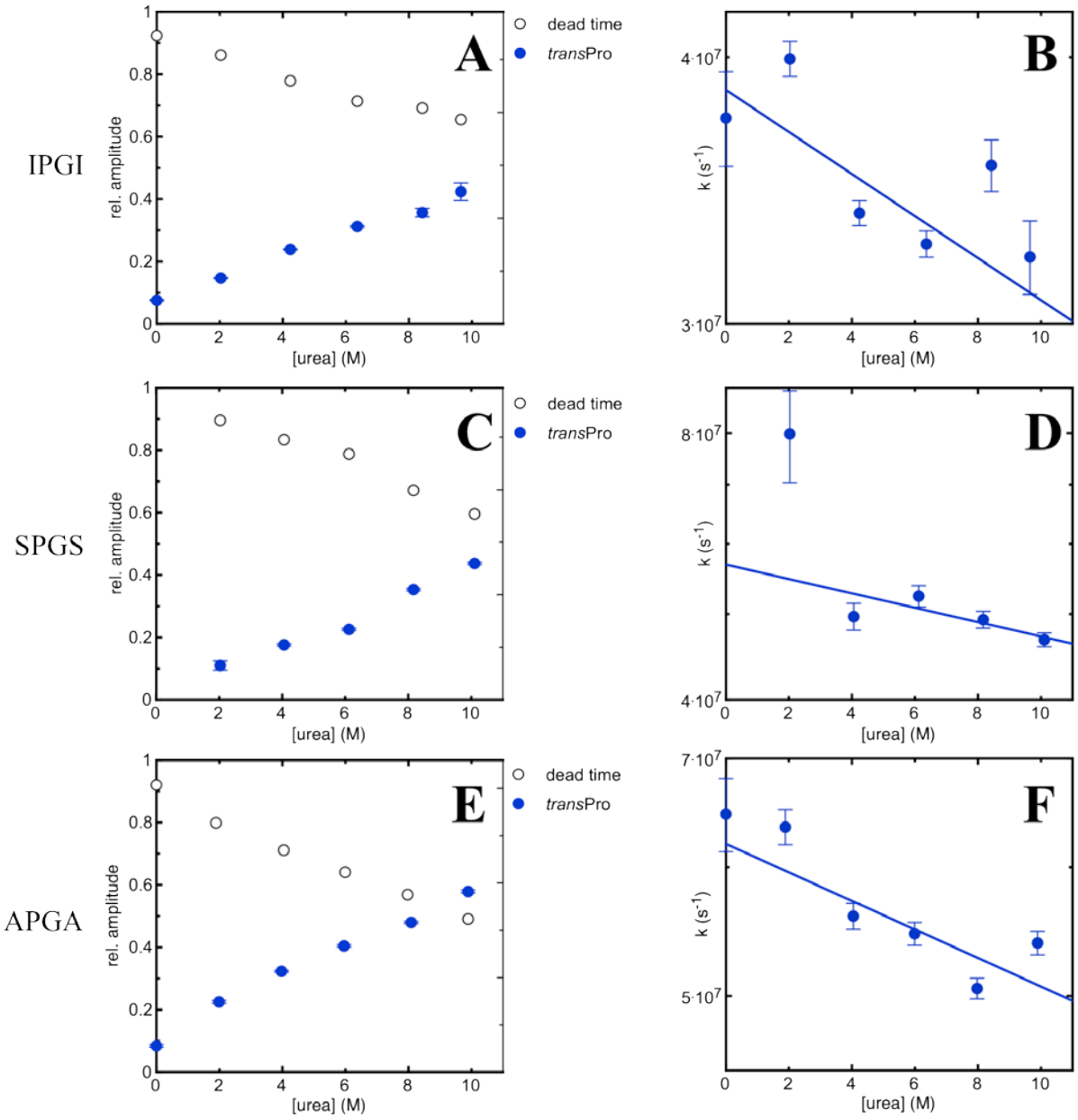


Figure 4.38 TTET measurements of PG turn model peptides at different urea concentrations. Relative amplitudes (A, C, E) and rate constants (B, D, F) obtained from exponential fits to the data.

In all peptides, the viscosity-corrected rate constants became slightly slower in the presence of urea. The data could be fitted by equation 4.13 resulting in m_c-values of 55.7 ± 9.3 J/mol/M for IPGI, 45.8 ± 13.3 J/mol/M for SPGS and 59.9 ± 13.7 J/mol/M for APGA. This indicates a rather small decelerating effect of urea on loop formation in PG turn model peptides. The fitted k_c^{H2O}-values equal (39.2 ± 0.8) · 10^6 s^{-1} for IPGI, (64.0 ± 2.9) · 10^6 s^{-1} for SPGS and (64.5 ± 1.3) · 10^6 s^{-1} for APGA and are in good accordance with the loop formation rate constants for *trans* isomers measured in buffer and fitted from the viscosity dependence (see Figure 4.32). For IPGI, the relative amplitude of the dead time phase decreases from ~95% in the absence of denaturant to ~66% in 10M urea. The amplitude of the loop phase increases concomitantly. For SPGS and APGA the dead time amplitude in 10M urea reaches values of ~60% and ~50%, respectively. Hence, even under very high concentrations of urea β-turn structure is not completely abolished but rather represents a significant part of the populated conformations. However, we assume that β-turn formation is not the only reaction, existent in the kinetics of PG turn model peptides, that occurs at a timescale which is not detectable by our TTET set up. The dead time phase may also include small contributions from kinetics due to peptides with the prolyl bond in *cis* conformation. Also, reactions that bring the TTET labels in contact through a few bond rotations must be considered[25]. The presence of a large fraction of β-turn structure even at high urea concentrations is also confirmed by the results from CD measurements (see Figure 4.39).

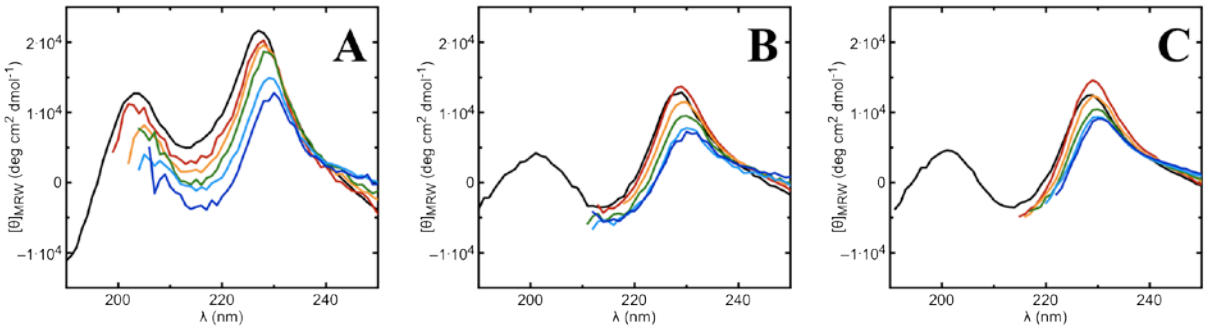


Figure 4.39 Far-UV CD spectra of (A) IPGI DA, (B) SPGS DA and (C) APGA DA at different urea concentrations. The colors of the spectra correspond to the urea concentrations in Figure 4.37. The black spectrum is the reference measurement in PO_4 buffer.

The maxima of the ellipticity, corresponding to turn structure and excimer, slightly decrease with increasing urea concentration but in 10M urea the curve still shows the characteristic shape. The decrease is of similar magnitude for the excimer band and the β-turn band, as shown for IPGI, indicating that turn structure in these peptides is associated with the formation of an excimer between the TTET labels Xan and Nal. The denaturing effect is more

pronounced for IPGI compared to SPGS and APGA, probably due to a more restricted conformational space because of the larger side chain of Ile compared to Ser and Ala.

We also examined the effect of urea on hydrogen bonding in turn structures. Therefore we determined NMR HN chemical shift temperature coefficients for the β-turn forming IPGI DA and for IPGI WW and IPGI WGW, where the adopted conformation is not clear, in 8M urea and compared the values to the coefficients in the absence of denaturant (see Table 4.11). From the CD spectrum we see a destabilization of the β-turn structure of IPGI DA upon the addition of 8M urea (see Figure 4.40, A). In buffer, a temperature coefficient of +3.6 ppb/K indicated a pronounced and stable hydrogen bond at the HN of Gly (turn residue *i*+2). Under denaturing conditions, the temperature coefficient decreases to a slightly negative value, which suggests that the intramolecular hydrogen bond is destabilized and may be partly replaced by a hydrogen bond to the solvent.

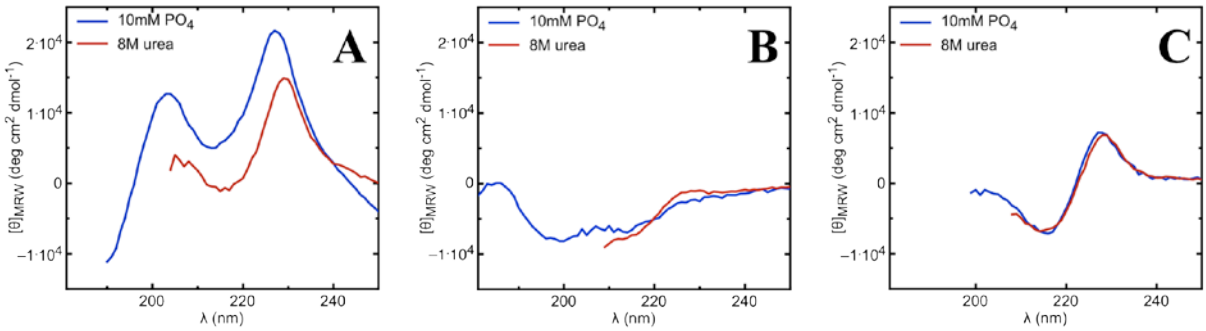


Figure 4.40 Far-UV CD spectra of IPGI derivative peptides under folding and unfolding conditions. (A) Xan-IPGI-Nal, (B) W-IPGI-W, (C) W-IPGI-GW.

For the Trp peptides (Figure 4.40, B,C), no significant change of the CD signal is observed in the presence of urea compared to native conditions. In buffer, for W-IPGI-W, the temperature coefficient decreases but is still positive, indicating the existence of a stable hydrogen bond even under strong denaturing conditions.

The structure adopted by W-IPGI-GW in buffer is not stabilized by an intramolecular hydrogen bond but by an excimer between the two Trp residues. The CD spectrum in 8M urea looks similar to the signal in the absence of denaturant, which suggests that the excimer is not affected by the addition of urea. The HN chemical shift temperature coefficients only slightly decrease under denaturing conditions.

Peptide	**Residue**						
	Xaa (*i*-1)	**I (*i*)**	**P (*i*+1)**	**G (*i*+2)**	**I (*i*+3)**	**Yaa (*i*+4)**	**Zaa (*i*+5)**
Xan-IPGI-Nal 10mM PO_4				3.6			
Xan-IPGI-Nal 8M urea		-12.7		-0.3	-5.7	-5.5	
W-IPGI-W 10mM PO_4	-10.2	-8.2		3.6	-5.1	-8.8	
W-IPGI-W 8M urea	-10.1	-8.5		2	-6.3	-7.9	
W-IPGI-GW 10mM PO_4	-9.3	-6.3		-5.7	-4.0	-7.6	-7.1
W-IPGI-GW 8M urea	-8.2	-7.0		-6.6	-5.3	-8.8	-7.2

Table 4.11 Residue-specific NMR HN chemical shift temperature coefficients (ppb/K) of IPGI derivative peptides under folding and denaturing conditions. Xaa, Yaa and Zaa stand for the residues at position (*i*-1), (*i*+4) and (*i*+5), respectively. The amino acid sequences of the peptides are shown in Table 4.9.

Besides urea, GdmCl is widely used as strong denaturant although the mechanisms of the denaturation are still incompletely understood. The main difference between the two molecules is that urea is neutral while GdmCl is ionic. As a consequence, its denaturing effect could partly be due to a salt effect[266]. Additionally, GdmCl is able to interact with aromatic residues by to cation-π interactions[267]. It is generally found that GdmCl is more effective as a denaturant than urea[70, 206] and we want to examine if the β-turn structure can be completely unfolded upon addition of high concentrations of GdmCl. Figure 4.41, shows the decrease in xanthone absorbance for IPGI, SPGS and APGA at the indicated GdmCl concentrations. For the DA peptides, a deceleration of the kinetics can be observed with increasing GdmCl concentration and the normalized amplitudes increase, simultaneously. Loop formation kinetics could be described by a single exponential fit. For the DO peptides, GdmCl has no quenching effect on the intrinsic lifetime of Xan but, similar to glycerol and urea, the normed amplitude is decreased with increasing denaturant concentration. Hence, we see a substantial decrease of the dead time amplitude, which indicates that the β-turn content decreases. Relative amplitudes and viscosity corrected rate constants for the detectable loop phase for IPGI, SPGS and APGA are shown in Figure 4.42.

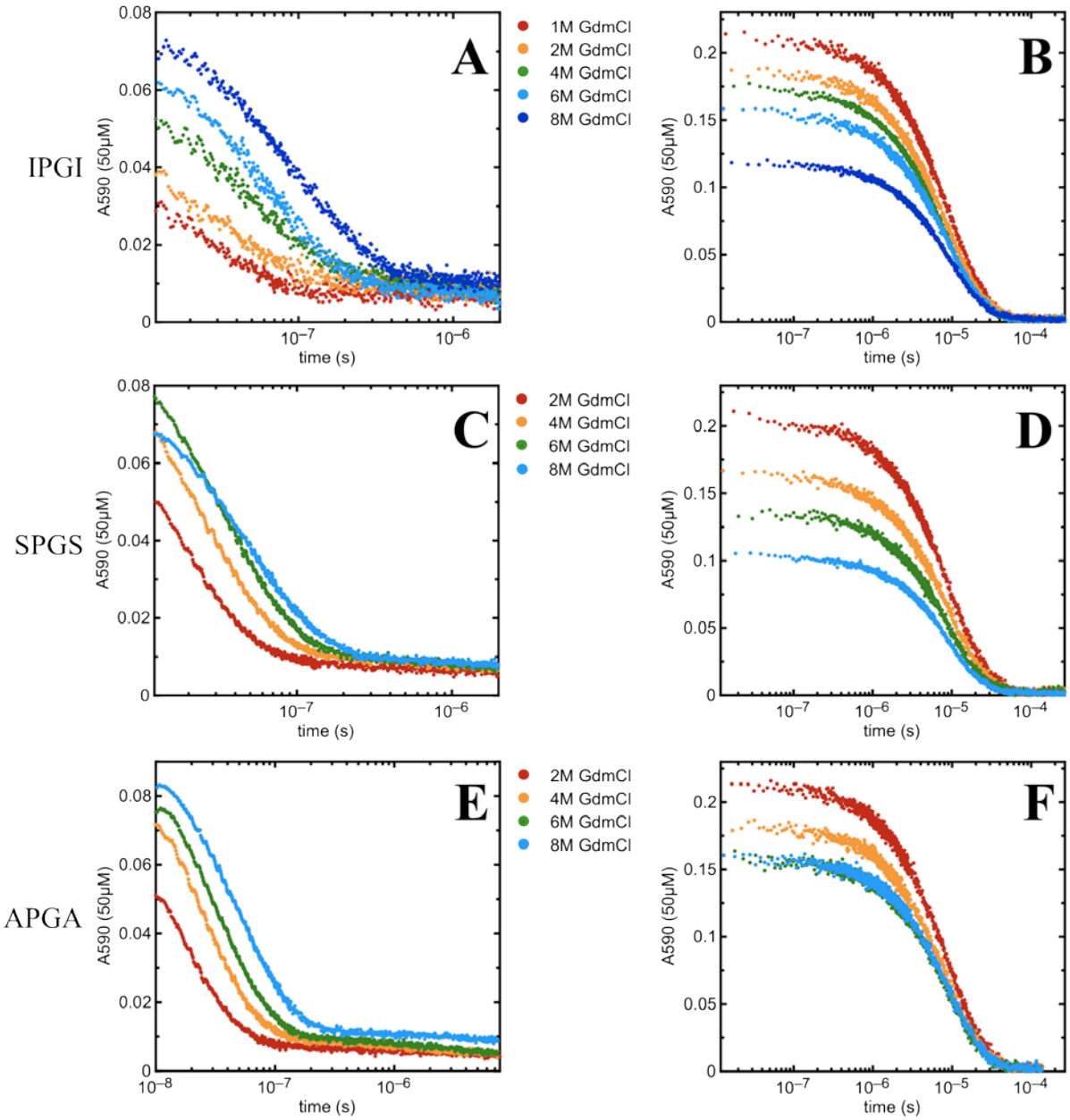


Figure 4.41 TTET measurements of PG turn model peptides at different GdmCl concentrations. Absorbance decay at 590 nm for (A, C, E) DA variants and (B, D, F) DO peptides.

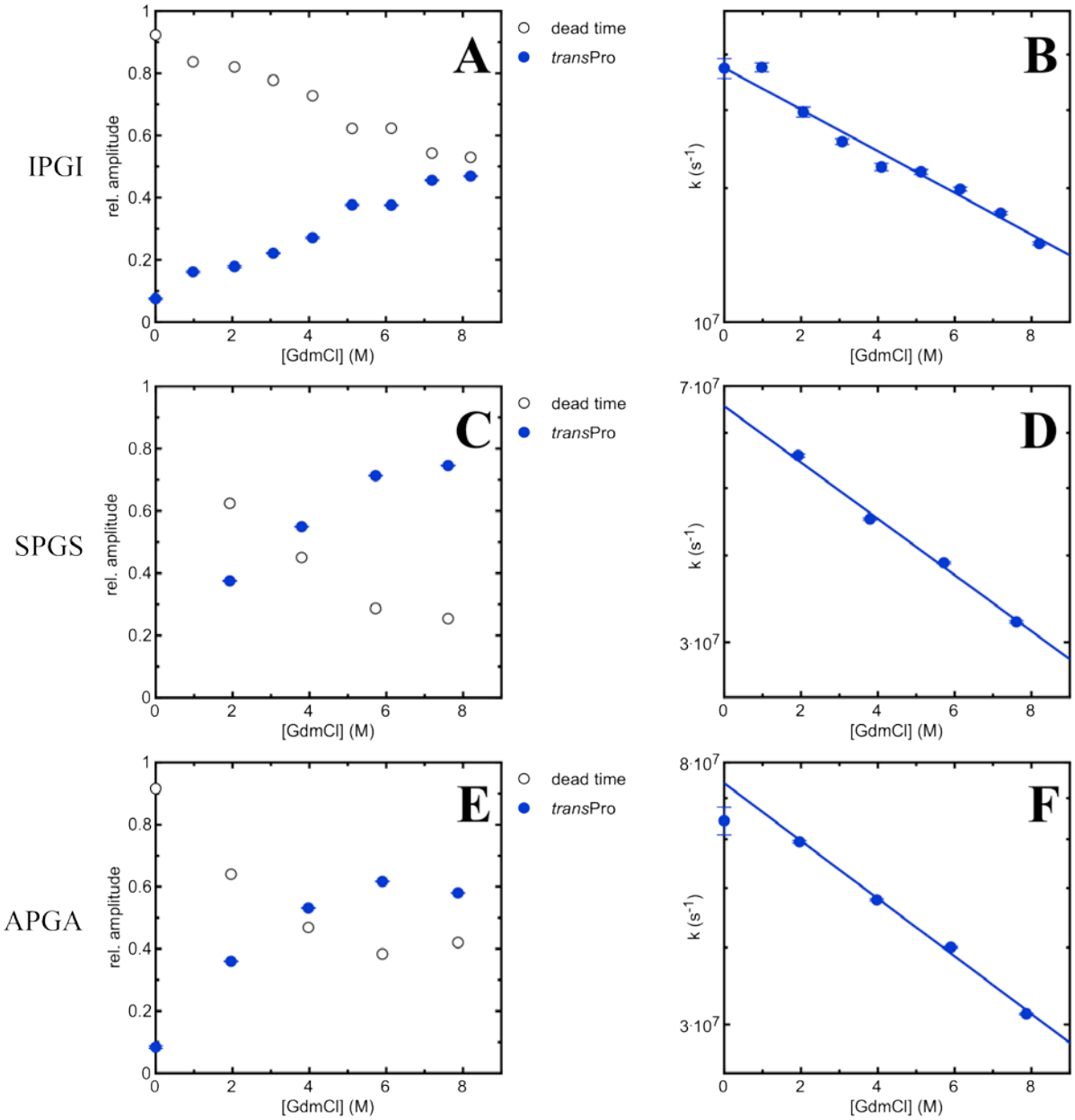


Figure 4.42 TTET measurements of PG turn model peptides at different GdmCl concentrations. Relative amplitudes (A, C, E) and rate constants (B, D, F) obtained from exponential fits to the data.

In all peptides, the viscosity-corrected rate constants linearly decrease with increasing GdmCl concentration. The data could be fitted by equation 4.13 resulting in k_c^{H2O}-values of (37.3 ± 0.5) · 10^6 s^{-1} for IPGI, (65.5 ± 0.3) · 10^6 s^{-1} for SPGS and (74.0 ± 0.4) · 10^6 s^{-1} for APGA. They are in good accordance with the loop formation rate constants for *trans* isomers measured in buffer and fitted from the viscosity dependence (see Figure 4.32).

The m_c-values equal 265.1 ± 5.1 J/mol/M for IPGI, 228.9 ± 2.6 J/mol/M for SPGS and 265.3 ± 2.2 J/mol/M for APGA and are ~5 times higher compared to the m_c-values in urea, which indicates a stronger decelerating effect of GdmCl on loop formation in PG turn model peptides. The ionic GdmCl might form cation-π interactions with the aromatic TTET chromophores and therefore destabilize the excimer. The same tendency is seen in the relative amplitudes of the different kinetic phases. For IPGI the relative amplitude of the dead time phase decreases from ~95% in the absence of denaturant to ~53% in 8M GdmCl. The amplitude of the loop phase increases concomitantly. For SPGS and APGA the dead time amplitude in 8M GdmCl reaches values of ~25% and ~42%, respectively. However, the remaining fraction of the dead time phase to the kinetics of the PG model peptides in 8M GdmCl still indicates the presence of β-turn structure. The stronger denaturing effect of GdmCl compared to urea is also confirmed by the results from CD measurements (see Figure 4.43). The strong excimer band at ~230 nm is drastically but still not completely decreased in 8M GdmCl.

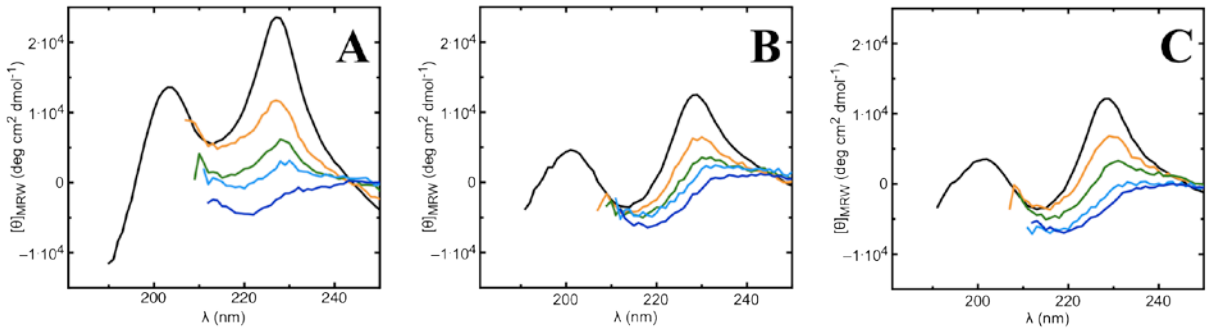


Figure 4.43 Far-UV CD spectra of (A) IPGI DA, (B) SPGS DA and (C) APGA DA at different GdmCl concentrations. The colors of the spectra correspond to the GdmCl concentrations in Figure 4.41. A, B. The black spectrum is the reference measurement in PO_4 buffer.

4.2.7 Conclusions and Outlook

We measured TTET and CD in model peptides of the sequence Xaa-PG-Yaa (PG-peptides) to gain insight into the formation of β-turns and their role in protein folding. It is not known whether turns can independently adopt their tight, chain reversing structure thereby serving as nucleation points for protein folding or if they are formed as a consequence of the formation of long-range interactions. We further wanted to use β-turns type II forming PG-peptides to answer the question whether a hydrogen-bonding requirement should be applied as an additional energetic criterion for the Ramachandran plot. If so, in β-turn type II structure, residues adjacent to the Gly in PG-peptides could only be satisfied by intramolecular H-bonds. In good solvents like high concentrations of urea or GdmCl, intramolecular hydrogen bonds are disfavored and thus the turn structure would be abolished[8].

In TTET measurements, we saw that β-turn structure is already formed in the dead time of our instrumental set up in a major population thereby eliminating ~90% of the amplitude of the signal. From CD spectroscopy we learned that our short model peptides are able to adopt β-turn structure when local stabilizing interactions are formed between flanking amino acids. We saw the highest turn content in PG-peptides including the TTET chromophores Xan and Nal. β-turn structure brings the two aromatic moieties in a close and rigid conformation thereby forming an excimer, which is represented in the CD spectrum as a maximum of the ellipticity at ~230 nm and a minimum at ~260 nm. Substitution of Nal to Phe also leads to β-turn structure but without the formation of an excimer. PG sequences flanked by residues, unable to form stabilizing interactions like Ala or Gly display typical random coil CD spectra.

We wanted to gain information about hydrogen-bonding in the PG-peptides by determining NMR HN chemical shift temperature coefficients. Positive temperature-coefficients suggest that the HN is involved in a stable and specific intramolecular hydrogen bond, as observed in α-helices[259]. We found positive temperature coefficients in two β-turn forming model peptides but not at position *i*+3 as predicted. In contrast, the HN of the Gly itself was found to be an H-bond donor. Thus, our results show that the PG peptides form tight β-turn structures, but they do not form the expected type-II β-turn. Unfortunately, we can not identify the hydrogen bonding acceptor of the Gly NH, and so we are not able to determine the exact conformation of the turn.

In order to test if the excimer that stabilizes β-turn structure in PG model peptides can also be formed between natural aromatic amino acids we synthesized PG-peptides flanked by Trp residues (W-IPGI-W and W-IPGI-GW). The results from CD and NMR measurements of the Trp peptides are inconsistent with earlier findings. In W-IPGI-W the TTET chromophores are

substituted by tryptophans and a positive NMR HN chemical shift temperature coefficient indicates the presence of a stable H-bond at the backbone HN of the Gly but the CD spectrum showed neither β-turn nor excimer formation. In W-IPGI-GW, a Gly is inserted between the turn and the Trp. This peptide adopts a structure that is dense enough that the two Trps form an excimer. However, NMR measurements did not give evidence for a H-bond and the presence of β-turn structure could also not be confirmed by CD spectroscopy.

From TTET and CD measurements we saw that the β-turn content in PG-peptides decreases with increasing denaturant concentration. However, even at high concentrations of denaturant the turn structure is not completely depleted. The denaturing effect of GdmCl on the β-turns is stronger compared to urea probably due to interactions of GdmCl molecules with the aromatic side chains involved in the stabilizing excimer. The NMR HN temperature coefficient in 8M urea is decreased to a slightly negative value indicating that the intramolecular hydrogen bond is partly replaced by a H-bond to the solvent.
However, for W-IPGI-W peptide a positive HN chemical shift temperature coefficient was still observed in 8M urea indicating a highly stable hydrogen bond..

Summarized, our results suggest an active role for β-turns in protein folding since they can be stabilized locally by interactions between aromatic residues and do not need long-range interactions. However, from first statistic examinations of PG sequences we did not find evidence for the preference of aromatic residues as flanking amino acids. We plan to break these statistics down to the different turn types for PG motifs.
Furthermore, our results indicate that the hydrogen-bond pattern in β-turns is more complex than predicted by Venkatachalam[115]. In order to gain more information about conformations possibly adopted by our PG-peptides we plan to perform computational simulations based on backbone dihedral ϕ-angles obtained by analysis of NMR $^{3}J_{HN,H\alpha}$ coupling constants.

4.3 Using TTET to investigate the influence of long-range interactions on the dynamics in an unfolded protein

Many insights into protein folding could be obtained by investigation of short peptides. TTET was successfully used to gain knowledge about dynamics of loop formation[268], folding and unfolding of α-helices[98] or the formation of turns (see chapter 4.2). Secondary structure elements in the observed short peptides are mainly determined by local interactions between near neighbor amino acids along the main chain. However, the native structure in full-length proteins is further stabilized by long-range interactions between amino acid residues that are far apart in the polypeptide chain. It was shown that in the native state, ~85% of residues are involved in long-range contacts (> ± 4 residues)[269]. It is not known whether some long-range interactions are already formed in the unfolded state. The calcium-binding protein parvalbumin constitutes remarkable resistance to heat and denaturants when Ca^{2+} is bound[171]. Calcium removal causes a decrease in α-helical content and the lack of fixed tertiary structure due to destabilization of the charge-charge interactions[179, 270]. We wanted to investigate whether the existence of residual helical structure in the disordered apo-state of parvalbumin is due to long-range interactions that are already formed in the unfolded state. Therefore we planned to perform TTET measurements in isolated peptide fragments and compare them with the full-length protein.

In order to use the TTET method to investigate the influence of long-range interactions on the dynamics in the unfolded state, xanthone and naphthalene have to be introduced in full-length proteins or independently folding protein domains. This has been done for the 35-residue villin headpiece subdomain by Fmoc SPPS[110, 112] but SPPS is restricted to proteins of 70 - 100 aa length and the yields decrease with increasing peptide size. We therefore wanted to establish a method to site-specifically introduce TTET labels during the expression of larger proteins, using non-natural amino acids and bioorthogonal chemistry.

4.3.1 A method to site-specifically introduce TTET labels into carp β-parvalbumin

Carp β-parvalbumin ps WT (PV) serves as a suitable model protein for TTET measurements due to the absence of xanthone triplet-quenching amino acids.

Our strategy was to introduce xanthone by click chemistry to the side chain of a residue in the protein sequence (see Figure 1.16) and naphthylalanine by stop codon suppression as shown in Figure 1.15. Therefore, we used an orthogonal tRNA/aaRS pair (kind gift of P.G.

Schultz[194]) that allows for the incorporation of the unnatural amino acid naphthylalanine at the amber stop codon UAG.

To be able to introduce Xan by click chemistry, the group of Prof. Marcel Mayor (Universität Basel) synthesized a xanthone derivative bearing an alkyne moiety (see Figure 4.44).

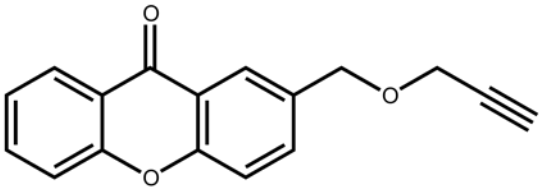


Figure 4.44 Xanthone derivative for click chemistry (synthesized by the group of Prof. Marcel Mayor, Universität Basel)

This xanthone-alkyne can be attached to an azide moiety in the protein. Therefore, we wanted to incorporate the methionine analogue azidohomoalanine (Aha, see Figure 1.14) at the start codon AUG via selective pressure incorporation (SPI). As Met is always the first amino acid of a protein, Met analogues incorporated at the N-terminal position have to be cleaved off co-translationally by the Met-aminopeptidase (MetAP). Small amino acids on the second position like Ala, which is the case in PV, facilitate this process. We chose the azide bearing Aha over the alkyne bearing Met analogue homopropargylglycine (Hpg) to be incorporated into the protein, because Hpg is less efficiently cleaved off the N-terminus[189].

4.3.2 Testing the suitability of the CuAAC-click reaction for the introduction of xanthone into proteins for TTET measurements

We wanted to use the Copper(I)-catalyzed azide-alkyne Huisgen cycloaddition (CuAAC) to attach a Xanthone-alkyne derivative to Aha, which is incorporated in the protein (see Figure 4.45, B). Hereby, a triazole ring is formed that might have a quenching effect on the triplet state of xanthone. Additionally, in comparison to the usually applied coupling of xanthonic acid to diaminopropionic acid (Dpr) (see Figure 4.45, A), CuAAC of the xanthone-alkyne derivative to Aha leads to a linker that is 5 bonds longer.

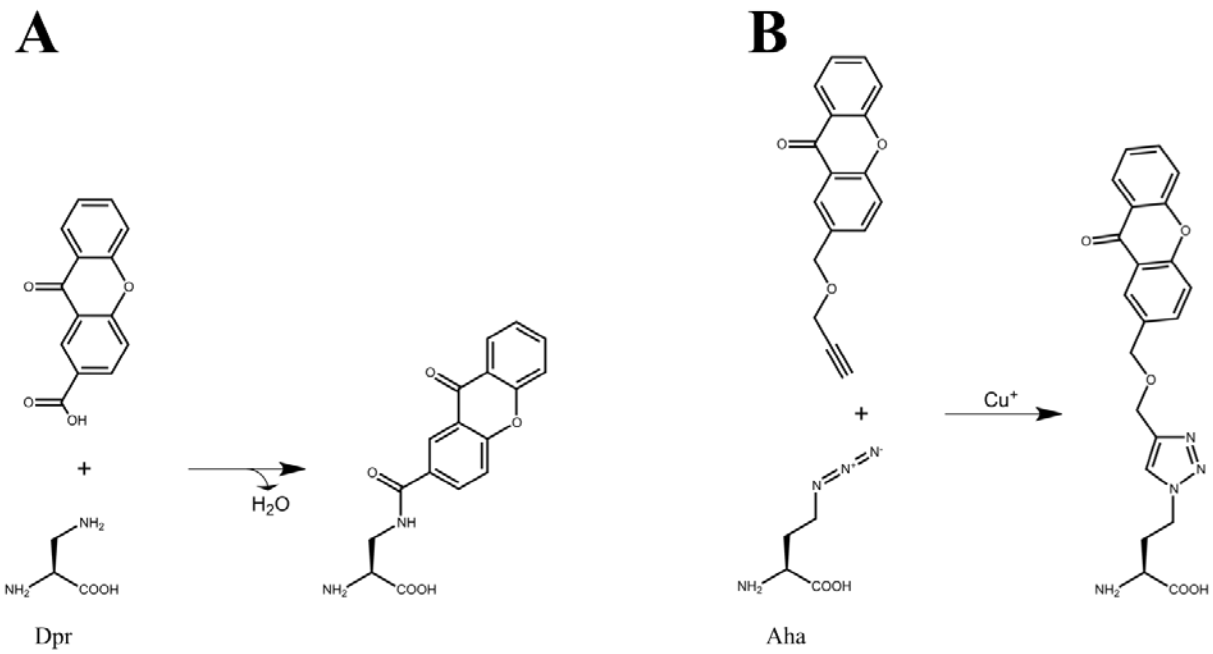


Figure 4.45 Comparison of chemical reactions suitable to attach xanthone to a polypeptide chain. (A) Coupling of xanthonic acid to diaminopropionic acid (Dpr). (B) Xanthone-alkyne derivative attached to Aha via CuAAC.

In order to test the influence of the triazole ring and the longer linker on the TTET reaction, we synthesized $(GS)_x$ model peptides with the sequence Ac-GS-Aha-$(GS)_x$-Nal-RG-NH_2 with x = 2, 4 via SPPS. In addition, a donor-only peptide of $(GS)_2$ was synthesized where Nal was replaced by Ala.

The click reaction was performed as described in chapter 3.3.1 and the product was purified using HPLC (for $(GS)_4$ see Figure 4.46, A).

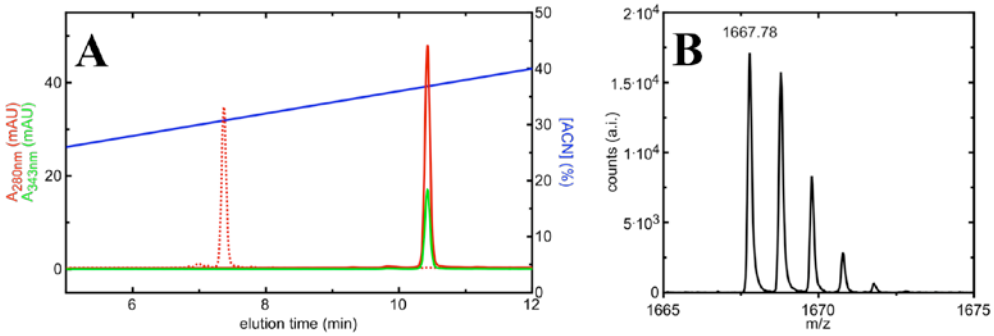


Figure 4.46 (A) Analytical HPLC of $(GS)_4$-click. Dashed line represents signal before the click-reaction, solid line shows signal of the clicked peptide. Absorption signal at 280 nm represented in red and at 343 nm represented in green. Acetonitrile (ACN) concentration is shown in blue. (B) Mass spectral analysis of purified GS_4-click. The detected mass of 1667 Da corresponds to the calculated molecular weight of the peptide with the sequence Ac-GS-Aha(Xan)-$(GS)_4$-Nal-RG-NH_2.

The clicked peptide eluted later than the educt due to increased hydrophobicity by the attached xanthone. In addition, xanthone was detected by its specific absorption signal at 343 nm. The success of the reaction was further verified by mass spectrometry (see Figure 4.46, B).

In order to test for quenching effects of the triazole ring, the time based absorbance change of the donor-only peptide ($(GS)_2$ DO) in degassed PO_4 buffer at 22.5 °C was compared to that of pure xanthonic acid and to a 21-amino acid Ala-based α-helix where Xan was attached to the side chain of the non-natural amino acid diaminopropionic acid (Dpr) (see Figure 4.47, A). The data could be described by a double exponential function. All three data sets show a slow main phase with ~85% amplitude. This phase corresponds to the intrinsic lifetime of the xanthone triplet state. For xanthonic acid free in solution, the intrinsic lifetime is about ~10 µs. When coupled to Dpr the intrinsic lifetime of Xan can be increased to ~35 µs[271]. For $(GS)_2$ DO, it has a slightly faster rate constant, indicating a small quenching effect of the triazole ring. However, the time constant for $(GS)_2$ DO still exceeds 8 µs which allows for measurements of absolute rate constants in the range of 10 ps to 8 µs. Additionally, a faster phase with lower amplitude (~15%) was found, corresponding to quenching residual oxygen in the solution.

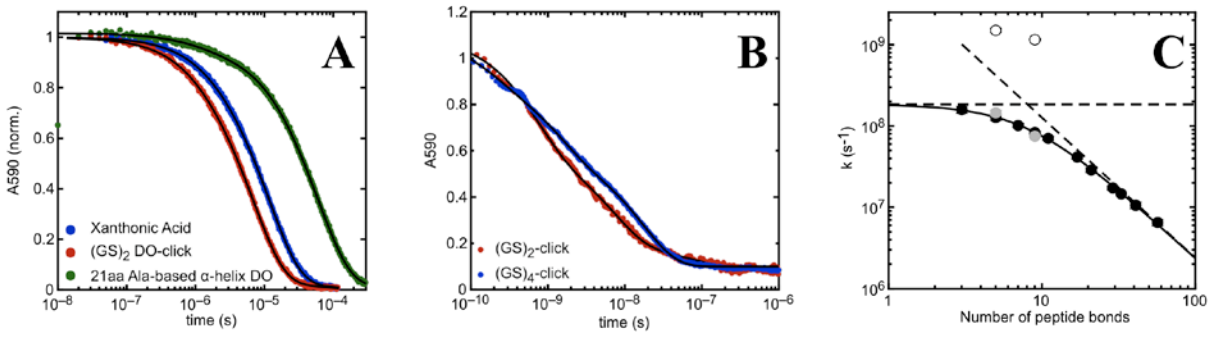


Figure 4.47 (A) Time based change in absorbance at 590 nm of xanthonic acid in blue, the 21 aa poly(Ala) α-helix in green and $(GS)_2$ DO in red. Black lines represent double-exponential fits. (B) TTET measurements of the DA variants of $(GS)_2$-click and $(GS)_4$-click (C) Loop formation rate constants (k_c) plotted versus loop size. Black circles represent k_c of $(GS)_x$ peptides, where xanthonic acid was coupled to the N-terminus. The two measured rate constants of loop formation in $(GS)_x$-click peptides are shown in grey and as open circles, respectively.

The dynamics of loop formation in $(GS)_x$-peptides were extensively studied using TTET. The rate constant for loop formation k_c decreases with increasing chain length (see Figure 4.47 B, black circles) and can be described by equation 4.14[53].

$$k_c = \frac{1}{1/k_0 + (k_l \cdot N^{\gamma})} \qquad (4.14)$$

In addition, combination of femtosecond and nanosecond laserflash experiments revealed complex kinetics of loop formation on the sub-nanosecond time scale. These kinetics reflect reactions that only require a few bond rotations to form a loop without crossing significant free energy barriers[25].

Experiments on $(GS)_x$-peptides, in which xanthonic acid was coupled to the N-terminus (depicted as black circles in Figure 4.47, B), were restricted to the nanosecond time scale and peptide concentrations >20 µM because of the experimental system applied. Due to low solubility of the $(GS)_x$-click peptides, TTET measurements were performed on a setup where lower peptide concentrations are sufficient. Here, xanthone was excited by a 120 ps laser flash that allows for determination of rate constants in the pico- to nanosecond range.

Loop formation in $(GS)_x$-click peptides could be described by a single exponential fit and revealed rate constants of $(14.5 \pm 1.0) \cdot 10^7$ s^{-1} for $(GS)_2$-click and $(7.5 \pm 0.2) \cdot 10^7$ s^{-1} for $(GS_)4$-click. They are in good agreement with the rate constants obtained for $(GS)_x$-peptides with xanthonic acid coupled to the N-terminus (see Figure 4.47, B, grey circles). A second phase with large similar exhibits rate constants of $(15.0 \pm 0.8) \cdot 10^8$ s^{-1} for $(GS)_2$-click and $(11.5 \pm 0.2) \cdot 10^8$ s^{-1} for $(GS)_4$-click and is due to TTET in a population of chains that only requires a few bond rotations to form the loop[25]. An additional slow phase with small amplitude represents the intrinsic donor lifetime and presumably arises from small peptide aggregates in which the labels are unable to form contact. This phase has been observed before in all investigated peptides.

Altogether, the CuAAC reaction with the above-described moieties is well suited to label proteins with xanthone. TTET measurements allow for determination of rate constants up to 8 µs and loop formation dynamics in $(GS)_n$-click peptides were essentially identical to those in $(GS)_n$-peptides where xanthonic acid was attached to the N-terminus.

4.3.3 Selection of labeling sites

The CD-spectrum of apo-PV reveals residual helical structure. In order to test for helix formation in the disordered state we planned to introduce TTET labels into a helical sequence of PV and compare the results from the isolated fragment to TTET between the same positions in the full-length protein.

The N-terminal A-helix of PV (residues 4 – 20, sequence: GVLNDADIAAALEASKA) is rich in alanine and possesses 4 branched amino acids. AGADIR predicts about 9% helical content for the isolated fragment under physiological conditions at 22.5ºC[272-276]. Xanthone should be attached on position 7, substituting the Asn and naphthylalanine should be incorporated on position 16 instead of Glu (see Figure 4.48). Consequently, the labels are placed in *i*,*i*+9 spacing and point to opposing sides in the folded structure. Therefore, the labels cannot interact in the helical state and TTET is prevented. However, in an unfolded or partially unfolded state, loop formation can occur and the triplet energy is transferred via van der Waals contact (see Figure 1.4, B).

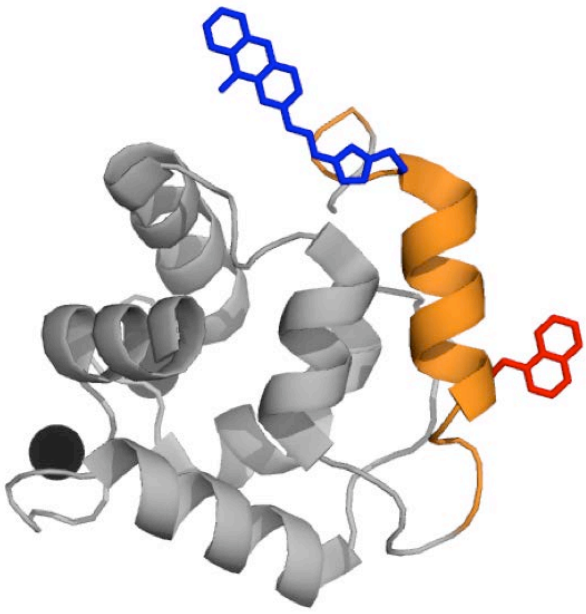


Figure 4.48 PV variant for TTET measurements. Xanthone, attached at position 7 is shown in blue and naphthalene at position 16 in red. The A-Helix is represented in orange.

4.3.4 Production of PV-constructs for TTET measurements

We synthesized the isolated PV A-helix fragment (residues 4 – 20) by SPPS with Aha on position 7 and Nal on position 16. Xanthone-alkyne was clicked to Aha via CuAAC. The N- and the C-termini were protected by acetylation and amidation, respectively.

We incorporated the methionine-analogue azidohomoalanine into the full-length β-carp Parvalbumin pseudo WT (PV) via selective pressure incorporation. Therefore, we introduced the methionine-encoding start codon ATG at position 7 by site-directed mutagenesis. The Met-auxotrophic *E.coli* strain B834(DE3) was co-transformed with two plasmids: the high expression plasmid pQE16 bears the DNA sequence of PV-N7-Aha fused N-terminally to

thioredoxin for better solubility[277] and a His-tag for purification. A thrombin cleavage site was introduced between PV and thioredoxin in order to remove the tags after purification. The plasmid pRep4 carries the *lacI* gene that encodes for the *lac* operon, which is the most commonly used system for regulation of expression of recombinant proteins in *E.coli*. The cells were grown in lysogeny broth (LB) medium until they reached the logarithmic growth phase. Subsequently, the cells were washed with sterile H_2O to get rid of residual Met present in the LB-medium and were resuspended in minimal medium lacking Met. Protein expression was induced by 0.1% lactose and Aha was added simultaneously. Usually, the allolactose mimic isopropyl β-D-1-thiogalactopyranoside (IPTG) is used to trigger transcription of the *lac* operon. Unlike allolactose, it is not hydrolysable by β-galactosidase and its concentration remains constant in the cells. However, for unknown reasons, former experiments showed that PV expression gives higher yields when induced with excessive lactose.

During protein expression, Aha which was added in excess, is recognized by the Met aminoacyl-tRNA synthetase and incorporated at every codon ATG. Aha in the first position is co-translationally cleaved off by the MetAP.

SDS-PAGE analysis of the bacterial lysate before and after expression showed a strong band appearing at ~25 kDa, corresponding to the calculated weight of the PV-thioredoxin-His-tag construct (see Figure 4.49, first two lanes).

For purification, the harvested cells were resuspended in a buffer containing a strong denaturant (8M Urea). After sonication, the lysate was cleared by high-speed centrifugation. None of the expressed protein remained in the cells as confirmed by SDS-PAGE of the resuspended cell pellet (see Figure 4.49, P).

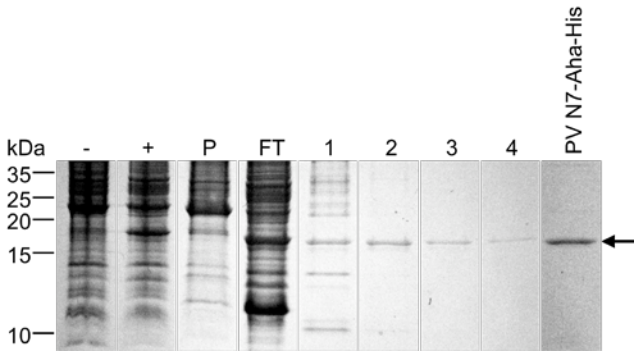


Figure 4.49 SDS-PAGE analysis of the expression and purification of PV N7-Aha-His. (-) bacterial lysate of non-induced control, (+) bacterial lysate after expression, (P) pelleted cells, (FT) flow-through, (1 – 4) collected fractions from the Ni-NTA column containing the desired product, (PV N7-Aha-His) pure fractions pooled and concentrated. PV N7-Aha-His is indicated by the arrow.

The protein was purified from the lysate by Ni-NTA chromatography. Unfortunately, most of the construct has not bound to the column and was found in the flow-through (see Figure 4.49, FT). The reason might be that other proteins sticking to PV shield the His-tag. It seems that PV tends to build unspecific complexes with other proteins present in the cell lysate. It was shown that these complexes could also not be separated by size exclusion chromatography of PV ps WT (see Figure 4.50). Surprisingly, neither the addition of a strong denaturant like urea, nor the use of detergents led to a separation of PV from the cell lysate. For the wild type protein, this was only achieved by addition of high concentrations of ammonium sulfate. Under these conditions, cell lysate proteins precipitate but PV remains soluble[278]. Unfortunately, upon incorporation of Aha, PV is destabilized and already precipitates at low concentrations of ammonium sulfate.

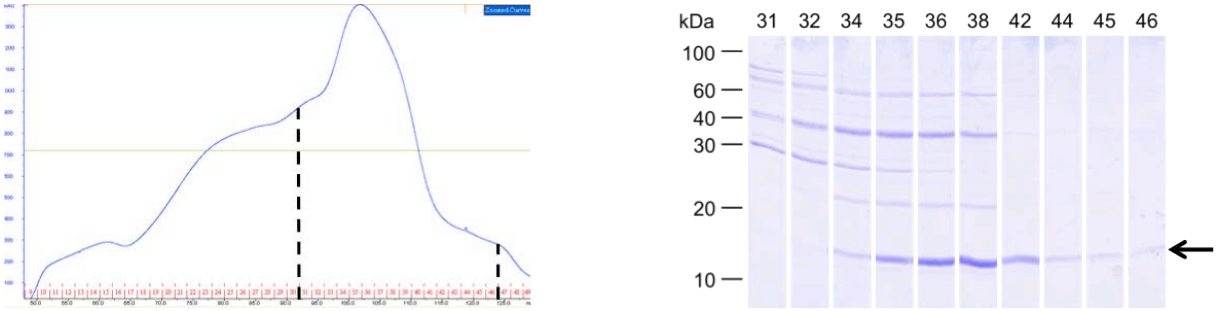


Figure 4.50 Size exclusion chromatography (left panel) and SDS-PAGE analysis (right panel) of PV ps WT. The over-expressed protein eluted in fractions between the dashed lines. PV ps WT is indicated by the arrow.

However, a small amount of the PV-thioredoxin-His-tag construct was bound to the Ni-NTA column and pure protein could be eluted in several fractions (see Figure 4.49, 1 – 4).
Subsequently, the C-terminal thioredoxin- and His-tag was removed by addition of thrombin. Figure 4.51, A shows the cleavage reaction progress over 48 hours. The construct band at ~25 kDa vanished over time and two bands at ~14 kDa (thioredoxin + His-tag) and ~11 kDa (PV N7-Aha) appeared. Cleaved PV was purified from the cleavage solution by a second Ni-NTA chromatography step where thrombin and the thioredoxin-His tag bound to the column and PV was collected in the flow-through.

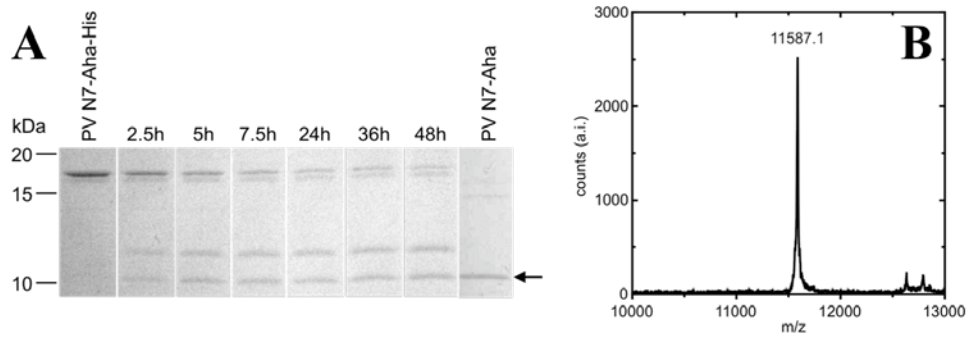


Figure 4.51 (A) SDS-PAGE analysis of thrombin cleavage; full-length PV with incorporated Aha is indicated by the arrow. (B) Mass spectral analysis of purified PV N7-Aha.

The successful incorporation of Aha was checked by mass spectrometry. A mass of 11587 Da was measured, identical with the expected mass for PV ps WT with N7 replaced by Aha (see Figure 4.51, B).

For attachment of xanthone to PV N7-Aha via CuAAC, the protein was treated as described above. The HPLC profile (see Figure 4.52, A) showed a small peak of unclicked protein after the reaction. It might arise from proteins where Aha was not accessible for the xanthone-alkyne. Since the reaction was performed in 100% DMSO under strong denaturing conditions, the formation of protein structures that shield Aha from the solvent can be excluded. Since the solubility of PV in DMSO is low, it is more likely that the polypeptide chains form aggregates where the residue on position 7 is buried inaccessible to the solvent.

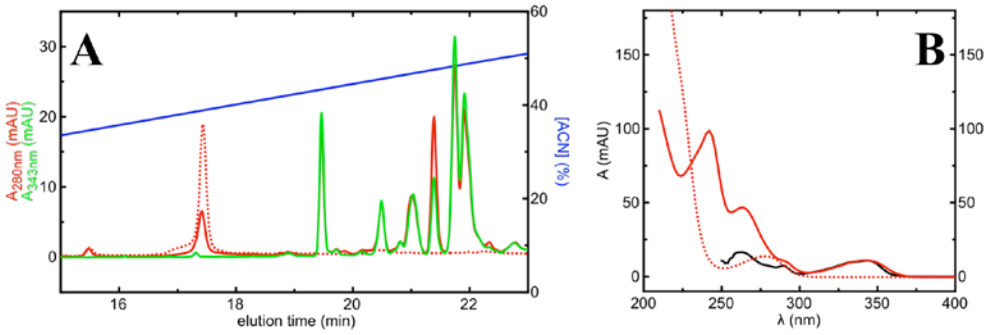


Figure 4.52 (A) Analytical HPLC of CuAAC of PV-N7-Aha. Dashed line represents signal before the click-reaction, solid line shows signal of the clicked protein. Absorption signal at 280 nm is represented in red, 343 nm absorption is represented in green. Acetonitrile (ACN) concentration is shown in blue. (B) Absorption spectra of HPLC peaks. Dashed line represents peak of the protein before the click-reaction, red solid line represents one of the HPLC peaks after the click reaction, black solid line represents the pure xanthon-alkyne derivative.

The reaction product eluted in several peaks from the HPLC column. Since the absorption spectrum for each peak looked similar (see Figure 4.52, B, red, solid line), we were not able to identify the right product. It is not clear whether the peaks arise from PV or from excessive xanthone-alkyne label. The calculated mass for PV N7-Xan of 11850 Da could also not be

found by mass spectral analysis. This might be due to degradation of PV during the CuAAC reaction.

Naphthylalanine can be incorporated into proteins during expression via stop codon suppression. Therefore an orthogonal tRNA/aaRS pair was used that recognizes naphthylalanine and the amber stop codon. In order to generate a PV variant carrying both the TTET donor Xan and the TTET acceptor Nal, the residue E16 in the mutated cDNA of PV N7-Aha-His was chosen to introduce the amber codon TAG.

For the double-incorporation, we transformed Met-auxotrophic B834(DE3) cells with the high expression plasmid pQE80_L that carries the PV-thioredoxin-His-tag construct and the *lacI* gene for regulation of expression. Simultaneously, we introduced the plasmid pEVOL_NapA, which encodes for an aaRS that is specific for naphthylalanine, and its cognate tRNA that recognizes the amber stop codon ($tRNA_{CUA}$)[194]. Additionally, it bears the *araC* repressor gene that allows for the induction of protein expression by arabinose.

Co-transformed *E.coli* cells were treated as described above. For induction of protein expression we added Aha, Nal, lactose and arabinose. SDS-PAGE analysis of the expression and purification of the PV double mutant is represented in Figure 4.53.

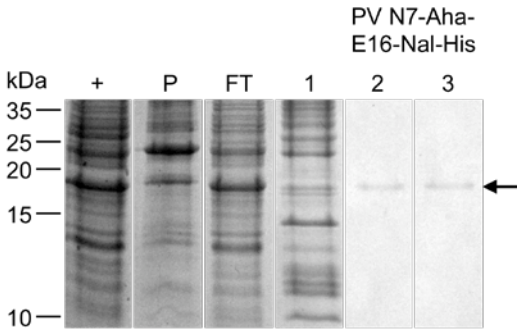


Figure 4.53 SDS-PAGE analysis of the expression and purification of PV N7-Aha-E16-Nal-His. (+) crude bacterial lysate after expression, (P) pelleted cells, (FT) flow-through, (1 – 3) collected fractions from the Ni-NTA column containing the desired product. PV N7-Aha-E16-Nal-His is indicated by the arrow.

In the crude bacterial lysate, a distinct band appeared at ~25 kDa after expression, representing the double-labeled PV-thioredoxin-His-tag construct (see Figure 4.53, first lane). This band indicates that Nal was successfully incorporated, as translation would have stopped otherwise at the amber stop codon at position 16. The orthogonality of the aaRS/$tRNA_{CUA}$ pair ensures that Nal and no other natural amino acid was incorporated.

The protein was purified via Ni-NTA chromatography under denaturing conditions as described above for the donor-only protein. Again, only small amounts of the protein were

bound to the column and the main fraction was found in the flow-through, probably due to the formation of aggregates or unspecific complexes with bacterial proteins and subsequent shielding of the His-tag from the solvent. The resulting yield of pure PV N7-Aha-E16-Nal-His (see Figure 4.53, 2+3) was not sufficient for subsequent modification steps.

4.3.5 Characterization of the isolated A-helix of PV

In order to examine the effect of long-range interactions on helix formation in disordered PV, data obtained from measurements of the full-length protein should be compared to the kinetics determined by TTET between the same positions in the isolated A-helix fragment.

The far-UV CD spectrum of the isolated A-helix in phosphate (PO_4) buffer at 22.5ºC displays typical helical bands with a maximum of the ellipticity at 190 nm and minima at 208 nm and 222 nm (see Figure 4.54, A). A quantitative analysis of the helix content using the signal at 222 nm gives a value of about 17% helix content, twice the number predicted by AGADIR. However, this value has some uncertainty due to CD bands of the TTET labels in the far UV region.

The helix content can be increased by addition of 2,2,2-trifluorethanol (TFE)[279]. TFE increases the helix propensity of all amino acids[280]. However, it is not clear if this is due to shielding of the CO and NH groups of the backbone from the aqueous solvent in the unfolded state[281] or clustering around nonpolar side chains as well as disruption of H-bonds formed between polar side chains in the folded state[282].

In the presence of 40% TFE, the CD spectrum of the PV A-helix shows lower minima at 208 nm and 222 nm. The calculated helix content increases to 35%.

TTET kinetics were monitored by the decay of the xanthone triplet absorbance band at 590 nm (see Figure 4.54, B). TTET data in pure buffer or supplemented with TFE exhibit double exponential kinetics, indicating that both the helical state and the coil state are populated to detectable amounts at equilibrium. The relative amplitude of the slower phase corresponds to the calculated helical content from the CD spectra and increases with the addition of TFE (see Figure 4.54, C). A loss in TTET amplitude was found for the dynamics in the donor-acceptor peptide compared to the donor-only peptide due to a fast reaction occurring in the dead-time of the kinetic measurement. This phase is attributed to fast loop formation in a subset of conformations in the coil state and decreases when the helical content is increased by addition of TFE.

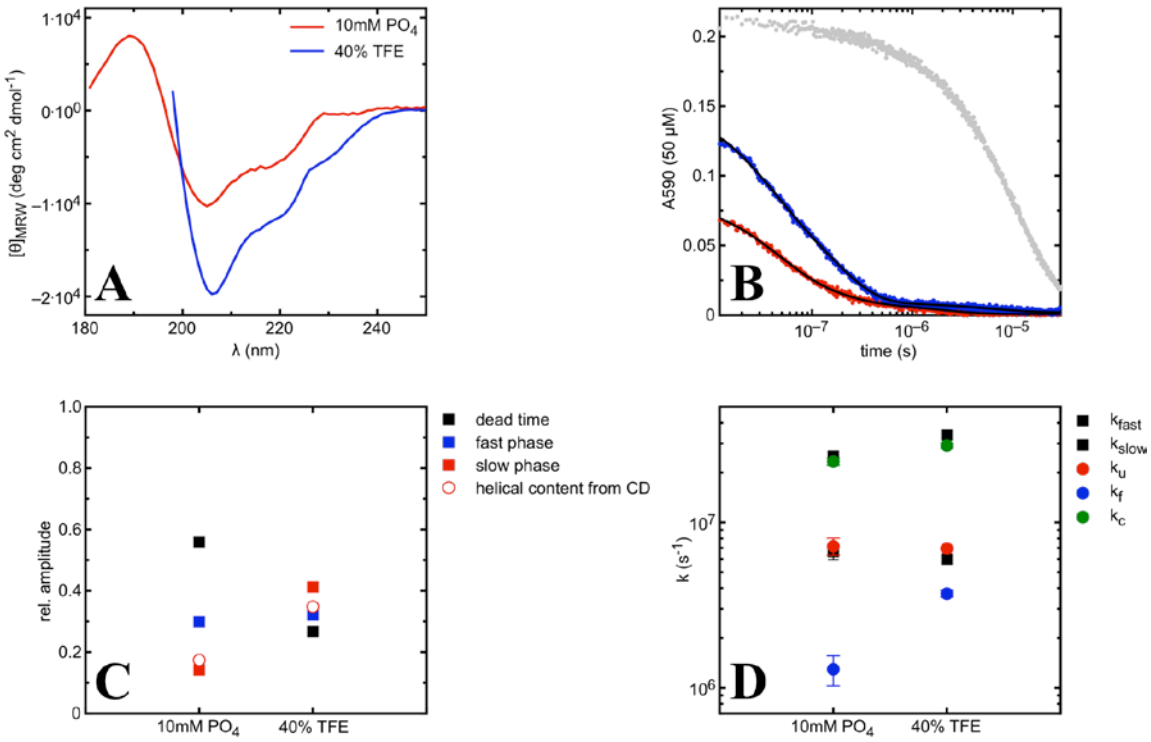


Figure 4.54 (A) Far-UV CD spectra of PV A-helix at 22.5ºC. Measurement in 10mM potassium PO_4 buffer and in 40% 2,2,2-trifluorethanol (TFE). (B) Triplet decay curves of xanthone monitored by the change in absorbance at 590 nm. The color-coding corresponds to panel A. In addition, the grey line represents the triplet decay for a donor-only peptide as a reference. Black lines represent exponential fits to the kinetics. (C) Obtained amplitudes are plotted as filled squares. The open circles represent calculated helical content from the CD spectra. (D) Macroscopic rate constants are represented as black squares, microscopic rate constants obtained from a fit according to equation 4.3 are plotted as colored circles.

Since TTET is an irreversible process, the microscopic rate constants for folding (k_f), unfolding (k_u) and loop formation (k_c) can be determined from the two observable rate constants λ_1 and λ_2 and their corresponding amplitudes A_1 and A_2 according to equation 4.3. The dead-time amplitude was added to the amplitude of the fast TTET process (A_2), which mainly represents TTET from the coil state[98, 109].

$$\lambda_{1/2} = \frac{k_u + k_f + k_c \pm \sqrt{(k_u + k_f + k_c)^2 - 4k_u k_c}}{2} \qquad (4.3a)$$

$$A_1 = \frac{1}{\lambda_1(\lambda_1 - \lambda_2)}([C]_0 \cdot k_c \cdot (k_u - \lambda_1) + [H]_0 \cdot k_u \cdot k_c) \qquad (4.3b)$$

$$A_2 = \frac{1}{\lambda_2(\lambda_1 - \lambda_2)}([C]_0 \cdot k_c \cdot (\lambda_2 - k_u) + [H]_0 \cdot k_u \cdot k_c) \qquad (4.3c)$$

Figure 4.54, D shows the experimentally measured rate constants and the microscopic rate constants obtained from equation 4.3 in PO_4 buffer and TFE. In PO_4 buffer, the microscopic rate constant for loop formation (k_c) is ~2.2-fold slower than the end-to-end loop formation

rate constant in a poly(Ser) chain of similar length. The decelerating effect arises from the extension of the N-terminal end by 3 residues and C-terminal end by 4 residues, which lead to decreased flexibility in the chain center. The effect is similar to the effects of end extensions on unfolded natural sequences (see chapter 4.1). The microscopic rate constant for folding (k_f) yields (1.3 ± 0.3) · 10^6 s^{-1} and is comparable to k_f in poly(Ala) model helices[98, 109] while the microscopic rate constants for unfolding (k_u = (7.2 ± 0.9) · 10^6 s^{-1}) is significantly higher, probably due to contributions from residues with lower helix propensity, which destabilizes the helix. The microscopic rate constants for unfolding (k_u) and loop formation (k_c) are only slightly affected by the addition of TFE. The rate constant for helix formation (k_f) is significantly increased, which reflects the helix favoring effect of TFE. In addition, the equilibrium constant (K_{eq}) for helix formation can be calculated from $K_{eq} = k_f / k_u$. K_{eq} equals about 0.18 ± 1.5 in PO_4 buffer consistent with the helical content calculated from the CD signal at 222 nm. In TFE, K_{eq} is increased reaching a value of about 0.53 ± 29.1. However, this value has a large error and might be smaller.

The results show that some helical structure is already formed in the isolated A-helix fragment without the influence of long-range interactions. The helical content can be further increased by the addition of TFE. In the future we plan to produce TTET labeled full-length PV to examine whether the helical content in the disordered state is increased due to long-range interactions.

4.3.6 Conclusions and Outlook

We aimed to investigate the influence of long-range interactions on the dynamics in an unfolded protein. Therefore we measured TTET between two sites near the ends of an amino acid sequence corresponding to the A-helix of the calcium binding protein carp β-parvalbumin (PV) and wanted to compare the results to TTET between the same positions in the full-length PV. In the isolated A-helix, some helical structure is already formed without the influence of long-range interactions. The helical content could be further increased by addition of TFE.

The combination of selective pressure incorporation, stop codon suppression and click chemistry is well suited for the introduction of triplet labels into proteins of larger size that are inaccessible by SPPS. TTET measurements in $(GS)_2$- and $(GS)_4$-click peptides showed that the triazole ring, which is formed in the course of the click reaction, does only have a small quenching effect on the lifetime of the Xan triplet state. Furthermore, loop formation dynamics in $(GS)_n$-click peptides were identical to those in $(GS)_n$-peptides where xanthonic

acid was attached to the N-terminus. Experiments showed the successful incorporation of Aha and Nal into carp β-parvalbumin and good protein production yields. However, the purification of PV remains challenging due to its tendency to stick to proteins from the bacterial lysate even under strong denaturing conditions or in the presence of detergents. A new purification strategy for PV has to be developed. A larger affinity tag such as the 26 kDa protein glutathione-S-transferase (GST) binding to immobilized glutathione might be less shielded by proteins from the bacterial lysate. Since parvalbumin can be efficiently refolded, another opportunity could be the use of fusion partners specifically designed to accumulate in inclusion bodies[283]. However, prior to the use of new tags, it must be ensured that their functionality is not affected by the substitution of Met by Aha.

For short peptides the click reaction in DMSO worked well. The reaction yield was nearly 100% and no side products were detected. However, DMSO is a bad solvent for PV, resulting in low reaction yields and probably degradation of the protein. A new solvent has to be found that meets the following conditions: good solvent for xanthone-alkyne and PV, suitable for click reaction and volatile, dialyzable or compatible with HPLC columns.

Besides PV, the here established method is suitable to label every protein that can be studied by TTET. Since methionine is an efficient xanthon triplet quencher, it must be absent or mutated in the studied protein anyway. However, it has to be ensured that the N-terminal Met is co-translationally cleaved off. One possible system is the interaction of the phosphorylated Kinase-inducible domain (pKID) and the KID interacting domain (KIX) studied by TTET. Here, Nal could be introduced into pKID via SPPS and KIX can be labeled with xanthone using SPI and click chemistry[284].

5. Summary

In order to understand how proteins fold into their native, biologically active state it is important to investigate the interplay of protein structure and protein dynamics. In this work, we characterized the dynamics in both unfolded and folded states of proteins and peptides using the method of triplet triplet energy transfer (TTET). TTET represents a two electron transfer process between an electronically excited triplet donor and a triplet acceptor, which have to come into van der Waals contact for transfer to occur. We used xanthone (Xan) as triplet donor and naphthalene (Nal) as acceptor. Bimolecular TTET experiments between these chromophores showed that the transfer process is diffusion controlled with an inverse viscosity-dependence and zero activation energy. Further, the bimolecular transfer rate constant is consistent with the predicted value for a diffusion-controlled reaction according to the Einstein-Smoluchowski equation. Introducing Xan and Nal into a polypeptide chain thus yields absolute rate constants for site-specific intrachain loop formation on a timescale from a few of picoseconds up to tens of microseconds.

Intrachain diffusion in unfolded polypeptide chains

During protein folding, the polypeptide chain explores the conformational space in the search for energetically favorable interactions. Thus, the investigation of site-specific contact formation between two sites in an unfolded polypeptide chain (loop formation) provides insight into the elementary steps of protein folding. In previous experiments, TTET was used to study dynamics in unfolded polypeptide chains using poly(Gly-Ser) and poly(Ser) model peptides as well as natural sequences derived from unstructured loops or intrinsically disordered proteins. For poly(GS) homopolymers, it has been shown that different chemical processes limit the diffusion over long and short distances. Longer peptides behave like statistical chains, while short chains are controlled by contributions from chain stiffness, as indicated by high activation energies (E_A) and weaker than $1/\eta$ viscosity dependencies. Chain dynamics in poly(Ser) chains and natural loop sequences from carp β-parvalbumin (PV) were shown to be slower and exhibit higher activation energies than the more flexible poly(GS) chains. Loop formation in a fragment from the intrinsically disordered DNA-binding Brinker domain (Brk_{66-89}) is significantly slowed down due to repulsive interactions between positively charged side chains. In this thesis, we investigated the effect of increased chain length and chain stiffness on end-to-end loop formation in natural sequences by measuring

TTET in three fragments derived from the intrinsically disordered antitoxin phd (AT phd 1, 2 and 3). Among the investigated peptides, AT phd 3 is the stiffest chain as it does not contain any Gly residues. Loop formation kinetics in AT phd 1 and AT phd 3 are comparable to those observed in poly(Ser) chains of similar length. Temperature dependence of loop formation revealed that increased chain stiffness in the AT phd 3 fragment leads to a higher enthalpical barrier as indicated by higher E_A-values and to a lower entropic barrier. These effects compensate and result in similar end-to-end loop formation rate constants for the two fragments. Intrachain diffusion in AT phd 2 is significantly slower compared to the other AT phd fragments and similar to the Brk_{66-89} fragment. Loop formation in AT phd 2 is enthalpically disfavored, due to residual structure or due to a prolyl bond in *trans* configuration near the center of the peptide, Loop formation in the Brk_{66-89} fragment, in contrast, is entropically disfavored, probably due to repulsive interactions, which lead to more extended chains.

During protein folding, end-to-end loop formation is rather uncommon. A more relevant case is interior contact formation between residues within the polypeptide chain. Earlier TTET measurements on homopolypeptide chains revealed slower loop formation rate constants for interior-to-interior loop formation compared to end-to-end loop formation. The effect of the additional tails was ascribed to decreased chain flexibility in the regions of the labels. This finding is in accordance with NMR experiments, which showed that polypeptide chains are more flexible at the ends compared to the chain center. In order to examine the effect of additional tails on the dynamics in unfolded peptides derived from natural sequences, we extended both ends of various peptides that differ in their chain stiffness. We therefore extended both ends of two parvalbumin fragments, the AT phd 3 fragment and the Brk_{66-89} fragment by 6 amino acids corresponding to the natural protein sequences and measured the effect of the dynamics of loop formation by TTET. The results were compared to a $(GS)_4$ peptide with N- and C-terminal extensions of 8 amino acids. The AT phd 3 peptide and the parvalbumin fragments with end extensions exhibit 2.2 – 2.5 times slower loop formation rate constants than the corresponding sequences without end extensions. The temperature-dependence of the TTET kinetics showed a higher activation energy for interior-to-interior loop formation compared to end-to-end loops. This effect becomes more pronounced with increasing chain stiffness in the loop region. Entropically, loop formation is favored by the attachment of tails, which may be explained by a restriction of conformational space due to the additional tails and hence a decreased loss in entropy upon loop formation. These

opposing effects sum up to the observed 2.2 – 2.5-fold decrease in the observed loop formation rate constants between end-to-end and interior-to-interior loop formation independent of chain length or flexibility. Dynamics of loop formation in the Brk_{66-89} fragment were only slowed down by a factor of 1.3 by the end-extensions despite a large unfavorable activation enthalpy, which is, however, nearly compensated for by favorable entropic contributions.

It was proposed that the dynamics of intrachain loop formation in polypeptides depend on solvent viscosity (η) and are further influenced by contributions from a viscosity-independent internal friction arising from intra-chain interactions or steric effects between the amino acid side chains. It was suggested that internal friction could be described as an additive term to solvent friction. In this case, the determination of loop formation time constants at different solvent viscosities would allow for the separation of internal friction from solvent friction through an extrapolation to zero solvent viscosity. Reports in literature found evidence for additive contributions from internal friction to chain dynamics assuming a perfect 1/η viscosity dependence for chain dynamics. However, for all peptides investigated in this thesis, a weaker than 1/η viscosity-dependence of the kinetics of loop formation was observed. As a result, extrapolating the loop formation time constants to zero viscosity reveals the absence of additional contribution from internal friction. Rather than contributing additively to solvent viscosity, internal friction seems to increase the activation energy and to modulate the effect of solvent viscosity on the dynamics of loop formation.

Formation of turns in model polypeptides

Turns are protein secondary structures that allow proteins to adopt compact globular states. It is still under debate, whether turns play an active role in protein folding, serving as initiation points for the folding process or whether they form as a consequence of the formation of long-range interactions between α-helices or β-strands. In order to gain insight into the structural and dynamic properties of short turn sequences we synthesized model peptides of the canonical sequence Xaa-PG-Yaa (PG-peptides) flanked by the TTET chromophores Xan and Nal. From the statistical analysis of known protein structures, it is known that Pro occurs most frequently at position *i*+1 in type-II β-turns, often followed by a glycine residue on position *i*+2. We wanted to use type II β-turn forming PG-peptides to test for the existence of a disfavored bridge in the Ramachandran plot, which was proposed by Rose and co-workers. In the disfavored bridge region the polypeptide chain can neither form local intramolecular

hydrogen bonds nor peptide solvent hydrogen bonds, which leads to a large energetic penalty. However, structure formation including formation of long-range hydrogen bonds can stabilize conformations in the disfavored bridge region. Formation of a type II β-turn in the PG-peptides would require the ϕ, ψ-angles of the Gly residue to be located in the disfavored bridge region of the Ramachandran diagram. As a result, this region should not be populated under good solvent conditions, like in the presence of high concentrations of denaturants, which destabilizes long-range intrachain interactions. In poor solvent conditions, however, the type-II β-turn should be formed due to formation of hydrogen bonds with a C=O(*i*) – N-H(*i*+3) H-bond. The CD spectra of all PG peptides in water display two maxima of the ellipticity at ~200 nm and ~230 nm as well as a minimum at ~260 nm. The CD band at ~200 nm is in agreement with a type-II β-turn conformation. The bands at 230 and 260 nm are due to the formation of an excimer between the TTET labels that may stabilize the turn structure and indicate high rigidity of the turn. TTET measurements on the PG peptides in water revealed fast TTET occurring in the dead-time of our experimental setup (~ 10 ns) for about 90 % of the molecules indicating that a turn is formed in a major fraction of the peptides in water, in agreement with the CD spectra. To further characterize the structure of the turn formed by the PG-peptides, we performed NMR experiments at different temperatures. Positive temperature-coefficients of the chemical shifts of NH protons are indicative for hydrogen bond formation and thus allow to identify hydrogen bonding donors. A positive NMR HN chemical shift temperature coefficient was observed only for Gly(*i*+2) in the PG-peptides indicating that its amide proton is involved a hydrogen bond. However, formation of the expected type-II β turn would result in hydrogen bond formation involving the adjacent residue at position (*i*+3). Thus, our results show that the PG-peptides form tight β-turn structures, but they do not form the expected type-II β-turn. Unfortunately, we can not identify the hydrogen bonding acceptor of the Gly(*i*+2) NH, and so we are not able to determine the exact conformation of the turn.

In order to test for the requirement of stabilizing interactions for β-turn formation in PG model peptides, we synthesized Ile-PG-Ile (IPGI) peptides flanked by different aromatic and non-aromatic amino acids. PG sequences flanked by non-aromatic residues revealed random coil structure as judged from CD spectroscopy. For IPGI flanked by Xan and Phe the CD spectroscopy indicated a β-turn signal but no signature of excimer formation was observed. Introducing Trp residues at the ends of the turn region (W-IPGI-W) also revealed the presence of a stable H-bond at the backbone HN of Gly(*i*+2). However, the Trp residues neither

induced β-turn structure, as judged by CD, nor the formation of an excimer. Interestingly, insertion of a Gly residue adjacent to the IPGI sequence (W-IPGI-GW) leads to excimer formation between two Trp residues and gives rise to a CD spectrum that closely resembles that of a Trp-zipper β-turn peptide. However, NMR experiments did not give any evidence for an H-bond involving the NH of Gly(*i*+2).

To test the effect of the solvent on turn formation in the PG-peptides, we performed TTET and CD measurements in the presence of various denaturants. Increasing concentrations of urea or GdmCl lead to a decrease in the fraction of excimers and in β-turn content. However, even at high concentrations of denaturant the turn structure is not completely depleted. The NMR HN chemical shift temperature coefficient is decreased to a slightly negative value in the presence of 8M urea, in accordance with a decreased fraction of turn conformations. However, for the W-IPGI-W peptide a positive HN chemical shift temperature coefficient was still observed in 8M urea indicating a highly stable hydrogen bond.

In summary, our results suggest an active role for β-turns in protein folding since they can be stabilized locally by interactions between aromatic residues and do not require long-range interactions between elements of secondary structure. Even short peptides can form rigid and highly stable turns that even persist in high concentrations of denaturants. Furthermore, our results indicate that the hydrogen-bonding pattern in β-turns is more complex than expected. It is essential to gain more information about the conformations that are possibly adopted by the PG-peptides. This could be done by computational simulations based on backbone dihedral ϕ-angles obtained by analysis of NMR $^{3}J_{HN,H\alpha}$ coupling constants.

Dynamics in the unfolded state of a full-length protein

In the last part of this thesis we wanted to investigate whether long-range interactions already exist in unfolded or partially folded proteins thereby inducing residual structure. We measured TTET between two sites near the ends of a peptide fragment corresponding to the A-helix of the calcium binding protein carp β-parvalbumin (PV) and wanted to compare the results to TTET between the same positions in the full-length PV. PV was chosen for this study since it is intrinsically disordered in the absence of calcium but shows some residual helical structure. In addition, PV lacks amino acids that interfere with the triplet state of xanthone. CD measurements on the isolated A-helix fragment showed the presence of about 17% helical structure. In accordance with the presence of helical structure, the TTET

experiments revealed biphasic transfer kinetics, as previously observed for helical, Ala-based peptides. The microscopic rate constant for loop formation (k_c) is ~2.2-fold slower than the end-to-end loop formation rate constant in a poly(Ser) chain of similar length, probably due to the extended ends that lead to increased stiffness in the chain center. The microscopic rate constant for helix formation (k_f) is comparable to k_f in poly-alanine model helices whereas the microscopic rate constants for unfolding (k_u) is significantly higher, probably due to contributions from residues with lower helix propensity, which destabilize the helix. The helical content in the A-helix fragment could be further increased by addition of 2,2,2-trifluoroethanol as revealed by CD spectroscopy. TTET experiments showed that the microscopic rate constants for unfolding (k_u) and loop formation (k_c) are only slightly affected by the addition of TFE. However, the rate constant for helix formation (k_f) is significantly increased, which shows that the transition state for helix formation already experiences the helix-stabilizing effect of TFE.

To be able to measure TTET in the 108 amino acid protein PV, which is not accessible to solid phase peptide synthesis, we established a method to site-specifically introduce the TTET chromophores during the expression of larger proteins using non-natural amino acids and bio-orthogonal chemistry. Our strategy was to introduce a xanthone-derivative bearing an alkyne moiety into the protein by Copper(I)-catalyzed azide-alkyne Huisgen cycloaddition (CuAAC, click reaction) to the non-natural amino acid azidohomoalanine (Aha). TTET measurements in $(GS)_2$- and $(GS)_4$-click peptides showed that the triazole ring, which is formed in the course of the click reaction, does only slightly decrease the lifetime of the xanthone triplet state. Furthermore, loop formation dynamics in $(GS)_n$-click peptides were essentially identical to those in $(GS)_n$-peptides where xanthonic acid was attached to the N-terminus. These results show that click chemistry is compatible with our TTET experiments, since the product does only slightly quench the xanthone triplet state and does not influence chain dynamics.

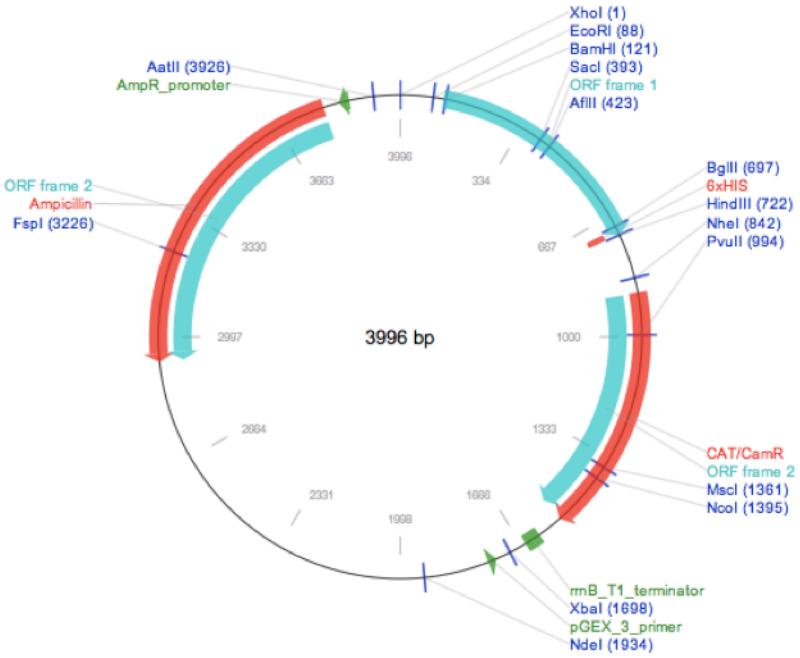
To be able to perform click-chemistry on full-length PV we introduced azidohomoalanine (Aha), which is a Met-analogue, successfully via selective pressure incorporation. Additionally, the non-natural amino acid naphthylalanine was incorporated into PV via stop codon suppression. The incorporation of the non-natural amino acids and attachment of xanthone to Aha via click-chemistry worked well, but the purification of PV was challenging due to its tendency to stick to proteins from the bacterial lysate resulting in very poor yields of pure protein. As a result we could not produce sufficient amounts of double-labeled PV for

TTET measurements. It is therefore essential for future work to develop a new purification strategy for PV.

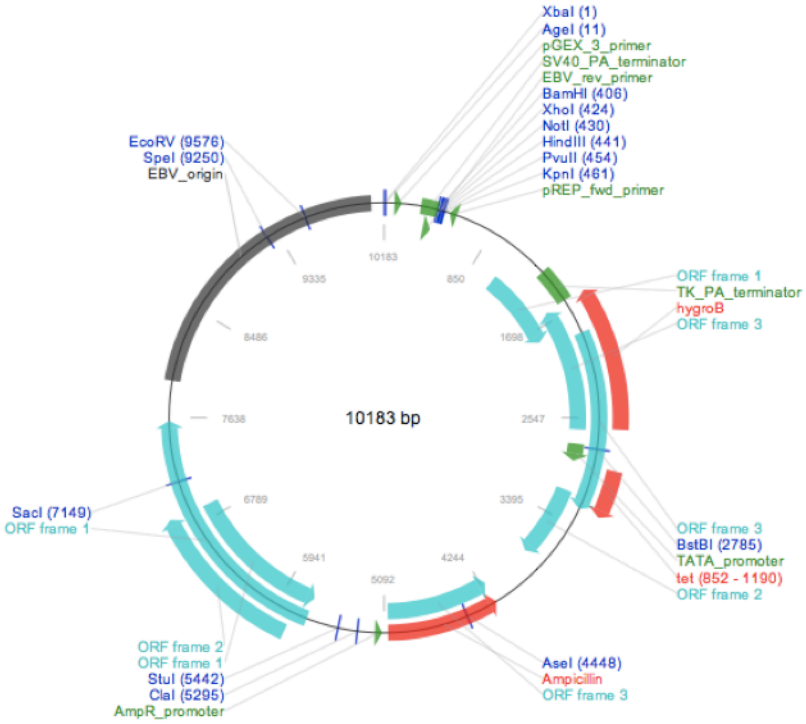
Appendix

Plasmid Maps

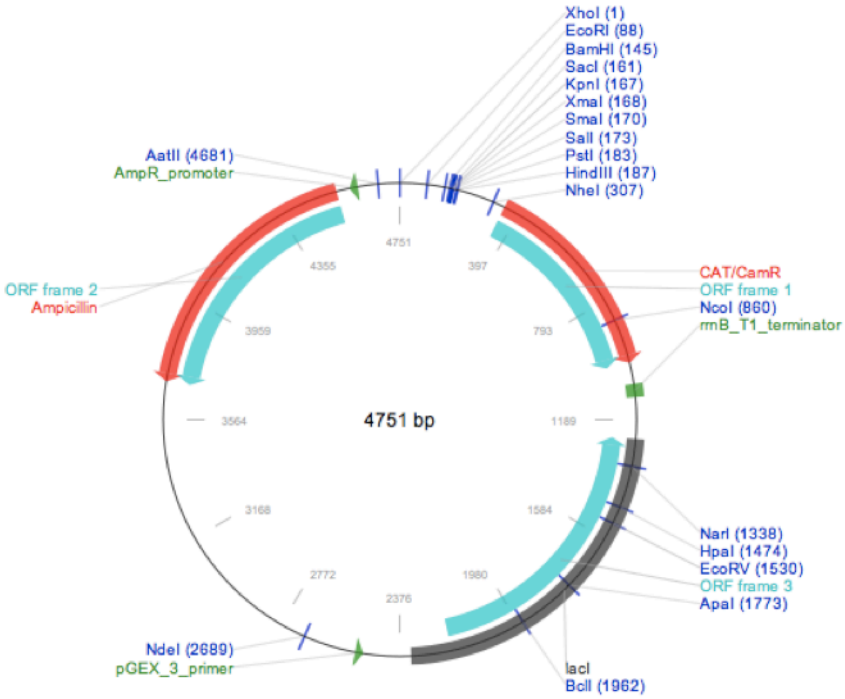
Name:	**pQE16**
Source:	Qiagen
Plasmid Type:	bacterial expression
Promotor:	T5
Size:	3996 bp
Bacterial Resistance:	Ampicillin
Notes:	His-tag removed



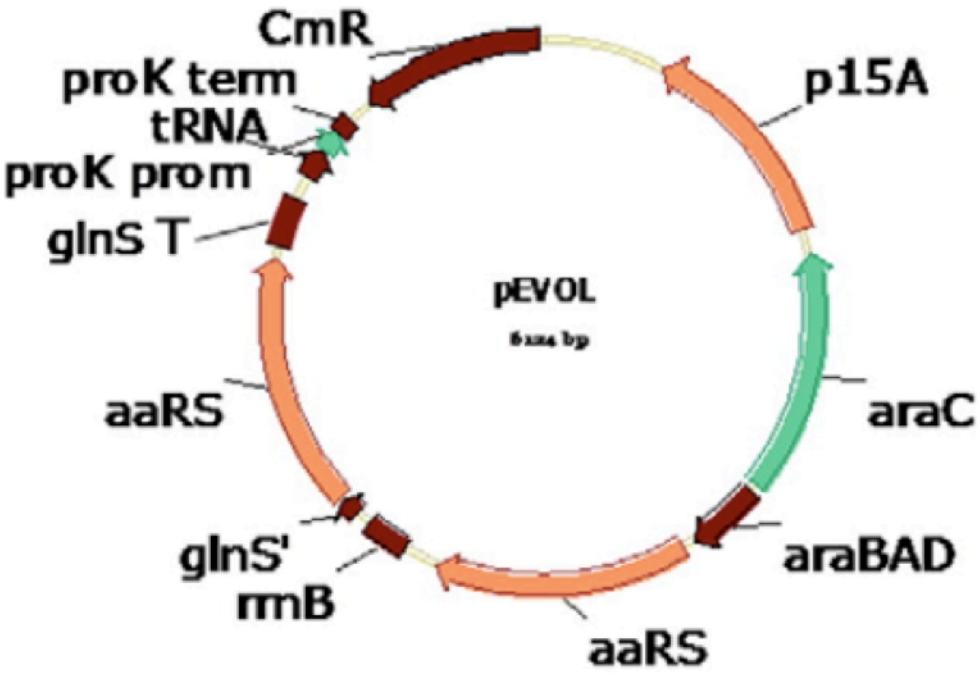
Name:	**pRep4**
Source:	Qiagen
Plasmid Type:	-
Promotor:	TATA
Size:	10183 bp
Bacterial Resistance:	Ampicillin
Notes:	Bacterial plasmid with p15A origin and lacI gene for regulating expression from pQE vectors.



Name:	**pQE80_L**
Source:	Qiagen
Plasmid Type:	Bacterial Expression
Promotor:	T5
Size:	4751 bp
Bacterial Resistance:	Ampicillin
Notes:	His-tag removed



Name:	**pEVOL**
Source:	Peter G. Schultz, The Scripps Research Institute
Plasmid Type:	Incorporation of unnatural amino acids at the amber stop codon
Promotor:	araBAD (aaRS gene), proK (tRNA)
Size:	6124 bp
Bacterial Resistance:	Chloramphenicol
Notes:	pEVOL-NapA for the incorporation of naphthylalanine



List of Figures

List of Tables

List of Abbreviations

AA	Amino acid
aaRS	Aminoacyl-tRNA synthetase
Ac_2O	Acetic anhydride
ACN	Acetonitrile
Aha	Azidohomoalanine
Ala	Alanine
Amp	Ampicillin
Arg	Arginine
Asn	Asparagine
Asp	Aspartic acid
AT	Antitoxin
Brk	Brinker domain
CAM	Chloramphenicol
Cpa	Cyclopropylalanine
Cu	Copper
CuAAC	Copper-mediated Huisgen 1,2-dipolar cycloaddition
Cys	Cysteine
DA	Donor-acceptor
DCM	Dichloromethane
DIPEA	N,N-diisopropylethylamine
DMF	Dimethylformamid
DMG	Dry molten globule
DO	Donor-only
Dpr	Diaminopropionic acid
E_A	Activation energy
ESI	Electrospray ionization
Fmoc	Flourenylmethoxycarbonyl
FRET	Förster resonance energy transfer
GdmCl	Guanidinium chloride
Gln	Glutamine
Glu	Glutamic acid
Gly	Glycine
GST	Gluthatione-S-transferase
H-bond	Hydrogen bond
HATU	O-(7-azabenzotriazol-1-yl)-N,N,N’,N’-tetramethyluroniumhexafluorophosphate
HBTU	O-(benzotriazol-1-yl)-N,N,N’,N’-tetramethyluronium hexafluorophosphate
His	Histidine
Hpg	Homopropargylglycine
HPLC	High-performance liquid chromatography
Ile	Isoleucine

IPTG	Isopropyl β-D-1-thiogalactopyranoside
ISC	Intersystem crossing
Kan	Kanamycin
KIX	KID interacting domain
LB	Lysogeny broth
Leu	Leucine
Lys	Lysine
MALDI-TOF	Matrix-assisted laser desorption/ionization-time of flight
Met	Methionine
MetAP	Methionine aminopeptidase
mRNA	Messenger RNA
N	Native state
Nal	Naphthalene
NCL	Native chemical ligation
NMM	N-methylmorpholine
NMR	Nuclear magnetic resonance
PCR	Polymerase chain reaction
Phe	Phenylalanine
pKID	Phosphorylated Kinase-inducible domain
PO_4	Phosphate
ppII	Polyproline II
Pro	Proline
PV	Carp β-parvalbumin
REFERs	Rate equilibrium free energy relationships
RNA	Ribonucleic acid
RT	Room temperature
SASA	Solvent accessible surface area
SDS-PAGE	Sodium dodecylsulfate polyacrylamide gel electrophoresis
Ser	Serine
SPI	Selective pressure incorporation
SPPS	Solid phase peptide synthesis
TES	Triethylsilane
TFA	Trifluoroacetic acid
TFE	2,2,2-triflourethanol
Thr	Threonine
TIPS	Triisopropylsilane
tRNA	Transfer RNA
Trp	Tryptophan
TS	Transition state
TTET	Triplet triplet energy transfer
Tyr	Tyrosine
U	Unfolded state
UAA	Unnatural amino acid
Val	Valine
WT	Wildtype
Xan	Xanthone

Acknowledgment

This work was carried out from June 2010 until February 2015 in the Group for Biophysical Chemistry at the department of chemistry, Technische Universität München.

At first, I want to thank Prof. Dr. Thomas Kiefhaber for giving me the great possibility to work in his group, for his excellent supervision and his support during the past years.
And I want to thank Prof. George Rose (Ph.D.) for the idea to test his theory on β-turns, the nice collaboration and valuable discussions of the results.

Many thanks to all current and former members of the group for the fantastic atmosphere in the lab: Tobias Aumüller, Annett Bachmann, Matthias Berg, Maren Büttner, Peter Enke, Susanne Halbritter, Martin Himmelreich, Michael Hösl, Kristine Steen Jensen, Peter Kämmerer, Richard Kil, Sabine Kullick, Sabine Neumaier, Christian Nyffenegger, Alexander Ogrodnik, Tobias Schümmer, Lena Schwarzer, Ursula Seidel, Jeremy Sloan, Karin Stecher, Matthias Stecher, Traudl Wenger, Stefan Wicht, Daniel Winter, Philipp Zimmermann and Ursula Zinth. Especially, I want to thank Christian for answering all my scientific and technical questions and helping me get settled in Munich, in- and outside the lab.
Thank you to my students Franziska Tippel, Beatrice Ramm, Carolin Gottmann and Anja Wurzer, who have contributed to the progress of my projects.

Thank you to Dr. Martin Haslbeck from the Chair of Biotechnology (Prof. Dr. Buchner), TUM for giving me the possibility to use his lab equipment and CD set up after our lab moved to Halle.
Many thanks to Dr. Gerd Gemmecker from the Chair of Biomolecular NMR-Spectroscopy, TUM for performing the NMR measurements and for the analysis of the NMR data.

I want to express my gratitude to my family for their extraordinary support and advice not only during this thesis but through my whole life.
And a very special thanks to Daniel for making me happy, encouraging me and supporting me on so many levels.

Bibliography

[1] Cohen, F. E.: Protein misfolding and prion diseases, Journal of molecular biology 1999, 293:313-20

[2] Lang, K., Schmid, F. X., Fischer, G.: Catalysis of protein folding by prolyl isomerases, Nature 1987, 329:268-70

[3] Eyring, H.: The activated complex in chemical reactions., J Chem Phys 1935, 3:107-15

[4] Anfinsen, C. B., Haber, E., Sela, M., White, F. H.: The kinetics of formation of native ribonuclease during oxidation of the reduced polypeptide chain, Proc Natl Acad Sci USA 1961, 47:1309-14

[5] Tanford, C.: Protein denaturation, Advan Prot Chem 1968, 23:121-282

[6] Tanford, C.: Protein Denaturation. Part C. Theoretical models for the mechanism of denaturation, Adv Prot Chem 1970, 24:1-95

[7] Ramachandran, G. N., Ramakrishnan, C., Sasisekharan, V.: Stereochemistry of Polypeptide Chain Configurations, J Mol Biol 1963, 7:95-9

[8] Porter, L. L., Rose, G. D.: Redrawing the Ramachandran plot after inclusion of hydrogen-bonding constraints, Proceedings of the National Academy of Sciences of the United States of America 2011, 108:109-13

[9] Flory, P. J.: Statistical Mechanics of Chain Molecules. Edited by Munich, Hanser Publishers, 1969,

[10] Kuhn, W.: Über die Gestalt fadenförmiger Moleküle in Lösungen., Kolloid-Z 1934, 52:269

[11] Kuhn, W.: Beziehungen zwischen Molekülgrösse, statistischer Molekülgestalt und elastischen Eigenschaften hochpolymerer Stoffe, Kolloid-Z 1936, 76:258

[12] Flory, P. J.: Principles of polymer chemistry. Edited by Ithaca, Cornell University Press, 1953,

[13] Rouse, P. E.: A theory of the linear viscoelastic properties of dilute solutions of coiling polymers, J Chem Phys 1953, 21:1272-80

[14] Zimm, B.: Dynamics of polymer molecules in dilute solutions: viscoelasticity, flow bifringence and dielectric loss, J Chem Phys 1956, 24:269-78

[15] Szabo, A., Schulten, K., Schulten, Z.: First passage time approach to diffusion controlled reactions, J Chem Phys 1980, 72:4350-7

[16] Tanford, C.: Protein denaturation. Part A. Characterization of the denatured state., Adv Prot Chem 1968, 23:121-217

[17] Brown, J. E., Klee, W. A.: Helix-coil transition of the isolated amino terminus of ribonuclease., Biochemistry 1971, 10:470-6

[18] Blanco, F. J., Jimenez, M. A., Rico, M., Santoro, J., Herranz, J., Nieto, J. L.: Tendamistat (12-26) fragment: NMR characterization of isolated β-turn folding intermediates., Eur J Biochem 1991, 200:345-51

[19] Alexandrescu, A. T., Abeygunawardana, C., Shortle, D.: Structure and dynamics of a denatured 131-residue fragment of staphylococcal nuclease: a heteronuclear NMR study, Biochemistry 1994, 33:1063-72

[20] Krieger, F., Fierz, B., Axthelm, F., Joder, K., Meyer, D., Kiefhaber, T.: Intrachain diffusion in a protein loop fragment from carp parvalbumin, Chem Phys 2004, 307:209-15

[21] Bieri, O., Wirz, J., Hellrung, B., Schutkowski, M., Drewello, M., Kiefhaber, T.: The speed limit for protein folding measured by triplet-triplet energy transfer, Proc Natl Acad Sci USA 1999, 96:9597-601

[22] Krieger, F.: Dynamics in Unfolded Polypeptide Chains as Model for Elementary Steps in Protein Folding. Edited by Basel, 2004, p.

[23] Möglich, A., Krieger, F., Kiefhaber, T.: Molecular basis for the effect of urea and guanidinium chloride on the dynamics of unfolded polypeptide chains, J Mol Biol 2005, 345:153-62

[24] Möglich, A., Joder, K., Kiefhaber, T.: End-to-end distance distributions and intrachain diffusion constants in unfolded polypeptide chains indicate intramolecular hydrogen bond formation, Proc Natl Acad Sci USA 2006, 103:12394-9

[25] Fierz, B., Satzger, H., Root, C., Gilch, P., Zinth, W., Kiefhaber, T.: Loop formation in unfolded polypeptide chains on the picoseconds to microseconds time scale, Proc Natl Acad Sci USA 2007, 104:2163-8

[26] Religa, T. L., Markson, J. S., Mayor, U., Freund, S. M. V., Fersht, A. R.: Solution structure of a protein denatured state and folding intermediate, Nature 2005, 437:1053-6

[27] Pashley, C. L., Morgan, G. J., Kalverda, A. P., Thompson, G. S., Kleanthous, C., Radford, S. E.: Conformational properties of the unfolded state of Im7 in nondenaturing conditions, Journal of molecular biology 2012, 416:300-18

[28] Uversky, V. N.: Intrinsically disordered proteins and their environment: Effects of strong denaturants, temperature, pH, counter ions, membranes, binding partners, osmolytes, and macromolecular crowding., Protein J 2009, 28:305-25

[29] Hofmann, H., Soranno, A., Borgia, A., Gast, K., Nettels, D., Schuler, B.: Polymer scaling laws of unfolded and intrinsically disordered proteins quantified with single-molecule spectroscopy, Proceedings of the National Academy of Sciences of the United States of America 2012, 109:16155-60

[30] Soranno, A., Koenig, I., Borgia, M. B., Hofmann, H., Zosel, F., Nettels, D., Schuler, B.: Single-molecule spectroscopy reveals polymer effects of disordered proteins in crowded environments, Proceedings of the National Academy of Sciences of the United States of America 2014, 111:4874-9

[31] Anfinsen, C. B.: Principles that govern the folding of protein chains, Science 1973, 181:223-30

[32] Pauling, L., Corey, R. B.: Configurations of polypeptide chains with favored orientations around single bonds: two new pleated structures, Proc Natl Acad Sci USA 1951, 37:729-40

[33] Pauling, L., Corey, R. B.: Atomic coordinates and structure factors for two helical configurations of polypeptide chains, Proc Natl Acad Sci USA 1951, 37:235-40

[34] Kauzmann, W.: Some factors in the interpretation of protein denaturation., Adv Prot Chem 1959, 14:1-63

[35] Baldwin, R. L.: Energetics of protein folding, Journal of molecular biology 2007, 371:283-301

[36] Anderson, D. E., Becktel, W. J., Dahlquist, F. W.: pH-induced denaturation of proteins: a single salt bridge contributes 4-5 kcal/mol to the free energy of folding of T4 lysozyme, Biochemistry 1990, 29:2403-8

[37] Perl, D., Mueller, U., Heinemann, U., Schmid, F. X.: Two exposed amino acid residues confer thermostability on a cold shock protein, Nat Struct Biol 2000, 7:380-3

[38] Wetlaufer, D. B.: Nucleation, rapid folding, and globular intrachain regions in proteins, Proc Natl Acad Sci USA 1973, 70:697-701

[39] Perutz, M. F., al., e.: Structure of hemoglobin. A three-dimensional Fourier synthesis at 5.5 Å resolution, obtained by x-ray analysis., Nature 1960, 185:416-22

[40] Wuthrich, K.: Protein structure determination in solution by nuclear magnetic resonance spectroscopy, Science 1989, 243:45-50

[41] Modig, K., Liepinsh, E., Otting, G., Halle, B.: Dynamics of protein and peptide hydration, J Am Chem Soc 2004, 126:102-14

[42] Palmer III, A. G.: NMR characterization of the dynamics of biomacromolecules, Chem Rev 2004, 104:3623-40

[43] Mulder, F. A., Mittermaier, A., Hon, B., Dahlquist, F. W., Kay, L. E.: Studying excited states of proteins by NMR spectroscopy, Nat Struct Biol 2001, 8:932-5

[44] Lange, O. F., Lakomek, N.-A., Farès, C., Schröder, G. F., Walter, K. F. A., Becker, S., Meiler, J., Grubmüller, H., Griesinger, C., de Groot, B. L.: Recognition dynamics up to microseconds revealed from an RDC-derived ubiquitin ensemble in solution., Science 2008, 320:1471-5

[45] Boehr, D. D., Nussinov, R., Wright, P. E.: The role of dynamic conformational ensembles in biomolecular recognition, Nat Chem Biol 2009, 5:789-96

[46] Jackson, S. E.: How do small single-domain proteins fold?, Folding & Design 1998, 3:R81-R91

[47] Gibbs, J. W.: On the Equilibrium of Heterogeneous Substances, Transactions of the Connecticut Academy 1875, III:108-248

[48] Lim, W. K., Rosgen, J., Englander, S. W.: Urea, but not guanidinium, destabilizes proteins by forming hydrogen bonds to the peptide group, Proceedings of the National Academy of Sciences of the United States of America 2009, 106:2595-600

[49] Levinthal, C.: How to fold graciously. Edited by Allerton House, Monticello, Ill., 1969, p. pp. 22-4

[50] Kubelka, J., Chiu, T. K., Davies, D. R., Eaton, W. A., Hofrichter, J.: Sub-microsecond protein folding, J Mol Biol 2006, 359:546-53

[51] Jaenicke, R.: Folding and association of proteins, Progr Biophys Mol Biol 1987, 49:117-237

[52] Kiefhaber, T., Quaas, R., Hahn, U., Schmid, F. X.: Folding of ribonuclease T1. 1. Existence of multiple unfolded states created by proline isomerization, Biochemistry 1990, 29:3053-61

[53] Krieger, F., Fierz, B., Bieri, O., Drewello, M., Kiefhaber, T.: Dynamics of unfolded polypeptide chains as model for the earliest steps in protein folding, J Mol Biol 2003, 332:265-74

[54] Fierz, B., Kiefhaber, T.: End-to-end vs interior loop formation kinetics in unfolded polypeptide chains, J Am Chem Soc 2007, 129:672-9

[55] Jaenicke, R.: Stability and folding of domain proteins, Progr Biophys Mol Biol 1999, 71:155-241

[56] Creighton, T. E.: Protein folding coupled to disulphide bond formation., Biol Chem 1997, 378:731-44

[57] Balbach, J., Schmid, F. X.: Proline isomerization and its catalysis in protein folding. Edited by Pain, R. Oxford, Oxford University Press, 2000, p.

[58] Brandts, J. F., Halvorson, H. R., Brennan, M.: Consideration of the possibility that the slow step in protein denaturation reactions is due to *cis-trans* isomerism of proline residues, Biochemistry 1975, 14:4953-63

[59] Schmid, F. X.: Prolyl isomerases: enzymatic catalysis of slow protein-folding reactions, Annu Rev Biomol Struct 1993, 22:123-43

[60] Stein, R. L.: Mechanism of enzymatic and nonenzymatic prolyl cis-trans isomerization, Adv Prot Chem 1993, 44:1-24

[61] Bergmann, L. W., Kuehl, W. M.: Formation of an intrachain disulfide bond on nascent chain immunoglobulin light chains., J Biol Chem 1979, 254:8869-76

[62] Braakmann, I., Hoover-Litty, H., Wagner, K. R., Helenius, A.: Folding of influenza hemagglutinin in the endoplasmic reticulum, J Cell Biol 1991, 114:401-11

[63] Creighton, T. E.: Intermediates in the refolding of reduced ribonuclease A, Journal of molecular biology 1979, 129:411-31

[64] Schwaller, M., Wilkinson, B., Gilbert, H. F.: Reduction-reoxidation cycles contribute to catalysis of disulfide isomerization by protein-disulfide isomerase, J Biol Chem 2003, 278:7154-9

[65] Leffler, J. E.: Parameters for the description of transition states, Science 1953, 117:340-1

[66] Pappenberger, G., Saudan, C., Becker, M., Merbach, A. E., Kiefhaber, T.: Denaturant-induced movement of the transition state of protein folding revealed by high pressure stopped-flow measurements, Proc Natl Acad Sci USA 2000, 97:17-22

[67] Pohl, F. M.: Temperature-dependence of the kinetics of folding of chymotrypsinogen A., FEBS Lett 1976, 65:293-6

[68] Schönbrunner, N., Pappenberger, G., Scharf, M., Engels, J., Kiefhaber, T.: Effect of Pre-Formed Correct Tertiary Interactions on Rapid Two-State Tendamistat Folding: Evidence for Hairpins as Initiation Sites for β-Sheet Formation., Biochemistry 1997, 36:9057-65

[69] Schätzle, M., Kiefhaber, T.: Shape of free energy barriers for protein folding probed by multiple perturbation analysis, J Mol Biol 2006, 357:655-64

[70] Myers, J. K., Pace, C. N., Scholtz, J. M.: Denaturant m values and heat capacity changes: relation to changes in accessible surface areas of protein unfolding., Protein Sci 1995, 4:2138-48

[71] Matouschek, A., Kellis, J. J., Serrano, L., Fersht, A. R.: Mapping the transition state and pathway of protein folding by protein engineering, Nature 1989, 340:122-6

[72] Fersht, A. R., Matouschek, A., Serrano, L.: The folding of an enzyme. I. Theory of protein engineering analysis of stability and pathway of protein folding, J Mol Biol 1992, 224:771-82

[73] Sánchez, I. E., Kiefhaber, T.: Origin of unusual phi-values in protein folding: Evidence against specific nucleation sites, J Mol Biol 2003, 334:1077-85

[74] Hammond, G. S.: A correlation of reaction rates., J Am Chem Soc 1955, 77:334-8

[75] Sánchez, I. E., Kiefhaber, T.: Non-linear rate-equilibrium free energy relationships and Hammond behavior in protein folding, Biophys Chem 2003, 100:397-407

[76] Jencks, W. P.: A primer for the Bema Hapothle. An empirical approach to the characterization of changing transition-state structures, Chem Rev 1985, 85:511-27

[77] Leffler, J. E., Grunwald, E.: Rates and equilibria of organic reactions. Edited by New York, Dover, 1963, p.pp. 458

[78] Sánchez, I. E., Kiefhaber, T.: Hammond behavior versus ground state effects in protein folding: evidence for narrow free energy barriers and residual structure in unfolded states., J Mol Biol 2003, 327:867-84

[79] Farcasiu, D.: The use and misuse of the Hammond postulate, J Chem Ed 1975, 52:76-9

[80] Bachmann, A., Kiefhaber, T.: Apparent two-state tendamistat folding is a sequential process along a defined route, J Mol Biol 2001, 306:375-86

[81] Kramers, H. A.: Brownian motion in a field of force and the diffusion model of chemical reactions., Physica 1940, 4:284-304

[82] Fierz, B., Kiefhaber, T.: Dynamics of unfolded polypeptide chains. Edited by Buchner, J., Kiefhaber, T. Weinheim, WILEY-VCH, 2005, p. pp. 805-51

[83] Zhu, W., Gisser, D. J., Ediger, M. D.: C-13 Nmr Measurements of Polybutadiene Local Dynamics in Dilute-Solution - Further Evidence for Non-Kramers Behavior, Journal of Polymer Science Part B-Polymer Physics 1994, 32:2251-62

[84] Zhu, W., Ediger, M. D.: Viscosity Dependence of Polystyrene Local Dynamics in Dilute Solutions, Macromolecules 1997, 30:1205-10

[85] Tylianakis, E. I., Dais, P., Heatley, F.: Non-Kramers' Behavior of the Chain Local Dynamics of PVC in Dilute Solution. Carbon-13 NMR Relaxation Study, J Polym Sci: Part B 1997, 35:317-29

[86] Beece, D., Eisenstein, L., Frauenfelder, H., Good, D., Marden, M. C., Reinisch, L., Reynolds, A. H., Sorensen, L. B., Yue, K. T.: Solvent Viscosity and Protein Dynamics, Biochemistry 1980, 19:5147-57

[87] Yedgar, S., Tetreau, C., Gavish, B., Lavalette, D.: Viscosity dependence of O2 escape from respiratory proteins as a function of cosolvent molecular weight, Biophys J 1995, 68:665-70

[88] Kleinert, T., Doster, W., Leyser, H., Petry, W., Schwarz, V., Settles, M.: Solvent Composition and Viscosity Effects on the Kinetics of CO Binding to Horse Myoglobin, Biochemistry 1998, 37:717-33

[89] Ansari, A., Jones, C. M., Henry, E., Hofrichter, J., Eaton, W. A.: The role of solvent viscosity in the dynamics of protein conformational changes, Science 1992, 256:1796-8

[90] Qiu, L., Hagen, S. J.: A limiting speed for protein folding at low solvent viscosity, J Am Chem Soc 2004, 126:3398-9

[91] Pabit, S. A., Roder, H., Hagen, S. J.: Internal friction controls the speed of protein folding from a compact configuration, Biochemistry 2004, 43:12532-8

[92] Jacob, M., Schindler, T., Balbach, J., Schmid, F. X.: Diffusion control in an elementary protein folding reaction., Proc Natl Acad Sci USA 1997, 94:5622-7

[93] Haas, E. K.-K., E; Steinberg, IZ: Brownian Motion of the Ends of Oligopeptide Chains in Solution as Estimated by Energy Transfer Between the Chain Ends, Biopolymers 1978, 17:11-3

[94] Qiu, L. H., S.J.: Internal friction in the ultrafast folding of the tryptophan cage. , Chemical Physics 2004, 312:327-33

[95] Jacob, M., Schmid, F. X.: Protein folding as a diffusional process, Biochemistry 1999, 38:13773-9

[96] Jacob, M., Geeves, M., Holtermann, G., Schmid, F. X.: Diffusional barrier crossing in a two-state protein folding reaction, Nat Struct Biol 1999, 6:923-6

[97] Chung, H. S., Eaton, W. A.: Single-molecule fluorescence probes dynamics of barrier crossing, Nature 2013, 502:685-8

[98] Fierz, B., Reiner, A., Kiefhaber, T.: Local conformational fluctuations in α-helices measured by fast triplet transfer., Proc Natl Acad Sci USA 2009, 106:1057-62

[99] Satzger, H., Schmidt, B., Root, C., Zinth, W., Fierz, B., Krieger, F., Kiefhaber, T., Gilch, P.: Ultrafast quenching of the xanthone triplet by energy transfer: new insight into the intersystem crossing kinetics, J Phys Chem A 2004, 108:10072-9

[100] Heinz, B., B., S., Root, C., Satzger, H., Milota, F., Fierz, B., Kiefhaber, T., Zinth, W., Gilch, P.: On the unusual fluorescence properties of xanthone in water., Phys Chem Chem Phys 2006, 8:3432 - 9

[101] Dexter, D. L.: A Theory of Sensitized Luminescence in Solids, J Chem Phys 1953, 21:836-50

[102] von Smoluchowski, M.: Zur kinetischen Theorie der Brownschen Molekularbewegung und der Suspensionen, Ann d Phys 1906, 21:756-80

[103] Closs, G. L., Johnson, M. D., Miller, J. R., Piotrowiak, P.: A connection between intramolecular long-range electron, hole and triplet energy transfer, J Am Chem Soc 1989, 111:3751-3

[104] Wagner, P. J., Klán, P.: Intramolecular triplet energy transfer in flexible molecules: electronic, dynamic, and structural aspects., J Am Chem Soc 1999, 121:9626-35

[105] Nyffenegger, C.: Effect of Osmolytes and other Co-Solutes on the Dynamics of Loop Formation in Proteins and Peptides, PhD thesis, Technische Universität München 2012,

[106] Joder, K.: Intramolecular and intermolecular diffusion processes in protein folding and assembly, PhD thesis, Technische Universität München 2011,

[107] Schellman, J. A.: Selective binding and solvent denaturation, Biopolymers 1987, 26:549-59

[108] Tanford, C.: Isothermal unfolding of globular proteins in aqueous urea solutions., J Am Chem Soc 1964, 86:2050-9

[109] Neumaier, S., Reiner, A., Buttner, M., Fierz, B., Kiefhaber, T.: Testing the diffusing boundary model for the helix-coil transition in peptides, Proceedings of the National Academy of Sciences of the United States of America 2013, 110:12905-10

[110] Reiner, A., Henklein, P., Kiefhaber, T.: An Unlocking/Relocking Barrier in Conformational Fluctuations of Villin Headpiece Subdomain, Proc Natl Acad Sci U S A 2010, 107:4955-60

[111] Shaknovich, E. I., Finkelstein, A. V.: Theory of cooperative transition in protein folding. I. Why denaturation of globular protein is a first-order phase transition, Biopolymers 1989, 28:1667-80

[112] Neumaier, S., Kiefhaber, T.: Redefining the dry molten globule state of proteins, Journal of molecular biology 2014, 426:2520-8

[113] Richardson, J. S.: The anatomy and taxonomy of protein structure, Adv Protein Chem 1981, 34:167-339

[114] Chou, K. C.: Prediction of tight turns and their types in proteins, Anal Biochem 2000, 286:1-16

[115] Venkatachalam, C. M.: Stereochemical criteria for polypeptides and proteins. V. Conformation of a system of three linked peptide units, Biopolymers 1968, 6:1425-36

[116] Lewis, P. N., Momany, F. A., Scheraga, H. A.: Chain reversals in proteins, Biochim Biophys Acta 1973, 303:211-29

[117] Hutchinson, E. G., Thornton, J. M.: A revised set of potentials for beta-turn formation in proteins, Protein science : a publication of the Protein Society 1994, 3:2207-16

[118] Karle, I. L.: Variability in the backbone conformation of cyclic pentapeptides. Crystal structure of cyclic(Gly-L-Pro-D-Phe-Gly-L-Ala), Int J Pept Protein Res 1986, 28:420-7

[119] Karle, I. L., Gibson, J. W., Karle, J.: The conformation and crystal structure of the cyclic polypeptide -gly-gly-D-ala-D-ala-gly-gly .3H2O, J Am Chem Soc 1970, 92:3755-60

[120] Karle, I. L., Urry, D. W.: Crystal structure of cyclic (APGVGV)2, an analog of elastin, and a suggested mechanism for elongation/contraction of the molecule, Biopolymers 2005, 77:198-204

[121] Torchia, D. A., Wong, S. C., Deber, C. M., Blout, E. R.: Cyclic peptides. 3. Solution conformations of cyclo(serylprolylglycylserylprolylglycyl) from nuclear magnetic resonance, J Am Chem Soc 1972, 94:616-20

[122] Torchia, D. A., Di Corato, A., Wong, S. C., Deber, C. M., Blout, E. R.: Cyclic peptides. II. Solution conformations of cyclo(ProlyLserylglycylprolylserylglycyl) from nuclear magnetic resonance, J Am Chem Soc 1972, 94:609-15

[123] Pease, L. G., Deber, C. M., Blout, E. R.: Cyclic peptides. V. 1 H and 13 C nuclear magnetic resonance determination of the preferred beta conformation for proline-containing cyclic hexapeptides, J Am Chem Soc 1973, 95:258-60

[124] Gierasch, L. M., Deber, C. M., Madison, V., Niu, C. H., Blout, E. R.: Conformations of (X-L-Pro-Y)2 cyclic hexapeptides. Preferred beta-turn conformers and implications for beta turns in proteins, Biochemistry 1981, 20:4730-8

[125] Madison, V.: Conformational energy and circular dichroism computed for cyclo-(pro-gly)3, Biopolymers 1973, 12:1837-52

[126] Bush, C. A., Sarkar, S. K., Kopple, K. D.: Circular dichroism of beta turns in peptides and proteins, Biochemistry 1978, 17:4951-4

[127] Brahms, S., Brahms, J., Spach, G., Brack, A.: Identification of beta,beta-turns and unordered conformations in polypeptide chains by vacuum ultraviolet circular dichroism, Proceedings of the National Academy of Sciences of the United States of America 1977, 74:3208-12

[128] Perczel, A., Hollosi, M., Tusnady, G., Fasman, G. D.: Convex constraint analysis: a natural deconvolution of circular dichroism curves of proteins, Protein Eng 1991, 4:669-79

[129] Perczel, A., Fasman, G. D.: Quantitative analysis of cyclic beta-turn models, Protein science : a publication of the Protein Society 1992, 1:378-95

[130] Rose, G. D.: Prediction of chain turns in globular proteins on a hydrophobic basis, Nature 1978, 272:586-90

[131] Guruprasad, K., Rajkumar, S.: Beta-and gamma-turns in proteins revisited: a new set of amino acid turn-type dependent positional preferences and potentials, J Biosci 2000, 25:143-56

[132] Hsu, H. J., Chang, H. J., Peng, H. P., Huang, S. S., Lin, M. Y., Yang, A. S.: Assessing computational amino acid beta-turn propensities with a phage-displayed combinatorial library and directed evolution, Structure 2006, 14:1499-510

[133] Trevino, S. R., Schaefer, S., Scholtz, J. M., Pace, C. N.: Increasing protein conformational stability by optimizing beta-turn sequence, Journal of molecular biology 2007, 373:211-8

[134] Fuchs, P. F., Alix, A. J.: High accuracy prediction of beta-turns and their types using propensities and multiple alignments, Proteins 2005, 59:828-39

[135] Kaur, H., Raghava, G. P.: A neural network method for prediction of beta-turn types in proteins using evolutionary information, Bioinformatics 2004, 20:2751-8

[136] Liu, L., Fang, Y., Li, M., Wang, C.: Prediction of beta-turn in protein using E-SSpred and support vector machine, Protein J 2009, 28:175-81

[137] McGuffin, L. J., Bryson, K., Jones, D. T.: The PSIPRED protein structure prediction server, Bioinformatics 2000, 16:404-5

[138] Petersen, B., Lundegaard, C., Petersen, T. N.: NetTurnP--neural network prediction of beta-turns by use of evolutionary information and predicted protein sequence features, PLoS One 2010, 5:e15079

[139] McGregor, M. J., Flores, T. P., Sternberg, M. J.: Prediction of beta-turns in proteins using neural networks, Protein Eng 1989, 2:521-6

[140] Matthews, B. W.: Comparison of the predicted and observed secondary structure of T4 phage lysozyme, Biochim Biophys Acta 1975, 405:442-51

[141] Madan, B., Seo, S. Y., Lee, S. G.: Structural and sequence features of two residue turns in beta-hairpins, Proteins 2014, 82:1721-33

[142] Marcelino, A. M., Gierasch, L. M.: Roles of beta-turns in protein folding: from peptide models to protein engineering, Biopolymers 2008, 89:380-91

[143] Kiehna, S. E., Waters, M. L.: Sequence dependence of beta-hairpin structure: comparison of a salt bridge and an aromatic interaction, Protein science : a publication of the Protein Society 2003, 12:2657-67

[144] Hughes, R. M., Waters, M. L.: Effects of lysine acetylation in a beta-hairpin peptide: comparison of an amide-pi and a cation-pi interaction, J Am Chem Soc 2006, 128:13586-91

[145] Ciani, B., Jourdan, M., Searle, M. S.: Stabilization of beta-hairpin peptides by salt bridges: role of preorganization in the energetic contribution of weak interactions, J Am Chem Soc 2003, 125:9038-47

[146] Blanco, F., Ramirez-Alvarado, M., Serrano, L.: Formation and stability of beta-hairpin structures in polypeptides, Curr Opin Struct Biol 1998, 8:107-11

[147] Munoz, V., Henry, E. R., Hofrichter, J., Eaton, W. A.: A statistical mechanical model for beta-hairpin kinetics, Proceedings of the National Academy of Sciences of the United States of America 1998, 95:5872-9

[148] Dinner, A. R., Lazaridis, T., Karplus, M.: Understanding beta-hairpin formation, Proceedings of the National Academy of Sciences of the United States of America 1999, 96:9068-73

[149] Lewis, P. N., Momany, F. A., Scheraga, H. A.: Folding of polypeptide chains in proteins: a proposed mechanism for folding, Proceedings of the National Academy of Sciences of the United States of America 1971, 68:2293-7

[150] Zimmerman, S. S., Scheraga, H. A.: Local interactions in bends of proteins, Proceedings of the National Academy of Sciences of the United States of America 1977, 74:4126-9

[151] Levitt, M.: A simplified representation of protein conformations for rapid simulation of protein folding, Journal of molecular biology 1976, 104:59-107

[152] Rose, G. D., Winters, R. H., Wetlaufer, D. B.: A testable model for protein folding, FEBS Lett 1976, 63:10-6

[153] Lesk, A. M., Rose, G. D.: Folding units in globular proteins, Proc Natl Acad Sci U S A 1981, 78:4304-8

[154] Dyson, H. J., Rance, M., Houghten, R. A., Lerner, R. A., Wright, P. E.: Folding of immunogenic peptide fragments of proteins in water solution. I. Sequence requirements for the formation of a reverse turn, Journal of molecular biology 1988, 201:161-200

[155] Shukla, R. T., Sasidhar, Y. U.: Energetics of beta-turn formation in a mutant peptide YPGDV from influenza hemagglutinin: an MD simulation study, Phys Chem Chem Phys 2013, 15:18571-83

[156] Yi, Q., Bystroff, C., Rajagopal, P., Klevit, R. E., Baker, D.: Prediction and structural characterization of an independently folding substructure in the src SH3 domain., J Mol Biol 1998, 283:293-300

[157] McCallister, E. L., Alm, E., Baker, D.: Critical role of beta-hairpin formation in protein G folding, Nat Struct Biol 2000, 7:669-73

[158] Kim, D. E., Fisher, C., Baker, D.: A breakdown of symmetry in the folding transition state of protein L, J Mol Biol 2000, 298:971-84

[159] Chen, P. Y., Gopalacushina, B. G., Yang, C. C., Chan, S. I., Evans, P. A.: The role of a beta-bulge in the folding of the beta-hairpin structure in ubiquitin, Protein science : a publication of the Protein Society 2001, 10:2063-74

[160] Jäger, M., Nguyen, H., Crane, J. C., Kelly, J. W., Gruebele, M.: The folding mechanism of a beta-sheet: the WW domain, J Mol Biol 2001, 311:373-93

[161] Vu, D. M., Brewer, S. H., Dyer, R. B.: Early turn formation and chain collapse drive fast folding of the major cold shock protein CspA of Escherichia coli, Biochemistry 2012, 51:9104-11

[162] Fowler, S. B., Clarke, J.: Mapping the folding pathway of an immunoglobulin domain: structural detail from Phi value analysis and movement of the transition state, Structure (Camb) 2001, 9:355-66

[163] Deuticke, H. J.: Über die Sedimentationskonstante von Muskelproteinen. , Hoppe-Seyler's Zeitschrift für physiologische Chemie 1934, 216-28

[164] Lehky, P., Blum, H. E., Stein, E. A., Fischer, E. H.: Isolation and characterization of parvalbumins from the skeletal muscle of higher vertebrates, J Biol Chem 1974, 249:4332-4

[165] Blum, H. E., Lehky, P., Kohler, L., Stein, E. A., Fischer, E. H.: Comparative properties of vertebrate parvalbumins, J Biol Chem 1977, 252:2834-8

[166] Ushio, H., Watabe, S.: Carp parvalbumin binds to and directly interacts with the sarcoplasmic reticulum for Ca2+ translocation, Biochem Biophys Res Commun 1994, 199:56-62

[167] Jiang, Y., Johnson, J. D., Rall, J. A.: Parvalbumin relaxes frog skeletal muscle when sarcoplasmic reticulum Ca(2+)-ATPase is inhibited, Am J Physiol 1996, 270:C411-7

[168] Schwaller, B., Dick, J., Dhoot, G., Carroll, S., Vrbova, G., Nicotera, P., Pette, D., Wyss, A., Bluethmann, H., Hunziker, W., Celio, M. R.: Prolonged contraction-relaxation cycle of fast-twitch muscles in parvalbumin knockout mice, Am J Physiol 1999, 276:C395-403

[169] Muntener, M., Kaser, L., Weber, J., Berchtold, M. W.: Increase of skeletal muscle relaxation speed by direct injection of parvalbumin cDNA, Proceedings of the National Academy of Sciences of the United States of America 1995, 92:6504-8

[170] Griesmeier, U., Vazquez-Cortes, S., Bublin, M., Radauer, C., Ma, Y., Briza, P., Fernandez-Rivas, M., Breiteneder, H.: Expression levels of parvalbumins determine allergenicity of fish species, Allergy 2010, 65:191-8

[171] Elsayed, S., Aas, K.: Characterization of a major allergen (cod). Observations on effect of denaturation on the allergenic activity, J Allergy 1971, 47:283-91

[172] Mori A, M. H., Ishizaki S, Shiomi K: Importance of conformation for the IgE reactivity of sarcoplasmic calcium-binding protein from the black tiger shrimp Panaeus monodon. , Eur Food Res Technol 2013, 165-70

[173] Tomura S, I. S., Nagashima Y, Shiomi K: Reduction in the IgE reactivity of Pacific mackerel parvalbumin by mutations at Ca2+-binding sites. , Fisheries Sci 2008, 411-7

[174] Goodman, M., Pechere, J. F.: The evolution of muscular parvalbumins investigated by the maximum parsimony method, J Mol Evol 1977, 9:131-58

[175] Kretsinger, R. H., Nockolds, C. E.: Carp muscle calcium-binding protein. II. Structure determination and general description, J Biol Chem 1973, 248:3313-26

[176] Berchtold, M. W.: Structure and expression of genes encoding the three-domain Ca2+-binding proteins parvalbumin and oncomodulin, Biochim Biophys Acta 1989, 1009:201-15

[177] Henzl, M. T., Agah, S., Larson, J. D.: Association of the AB and CD-EF domains from rat alpha- and beta-parvalbumin, Biochemistry 2004, 43:10906-17

[178] Arif, S. H.: A Ca(2+)-binding protein with numerous roles and uses: parvalbumin in molecular biology and physiology, Bioessays 2009, 31:410-21

[179] Kuwajima, K., Sakuraoka, A., Fueki, S., Yoneyama, M., Sugai, S.: Folding of Carp Parvalbumin Studied by Equilibrium and Kinetic Circular Dichroism Spectra, Biochemistry 1988, 27:7419-28

[180] Coffee, C. J., Bradshaw, R. A.: Carp muscle calcium-binding protein. I. Characterization of the tryptic peptides and the complete amino acid sequence of component B, J Biol Chem 1973, 248:3305-12

[181] Kumar, V. D., Lee, L., Edwards, B. F.: Refined crystal structure of calcium-liganded carp parvalbumin 4.25 at 1.5 Å resolution., Biochemistry 1990, 1404-12

[182] Bernstein, F. C., Koetzle, T. F., Williams, G. J., Meyer, E. F., Jr., Brice, M. D., Rodgers, J. R., Kennard, O., Shimanouchi, T., Tasumi, M.: The Protein Data Bank: a computer-based archival file for macromolecular structures, J Mol Biol 1977, 112:535-42

[183] Nyffenegger, C.: Dynamics in Different States of Parvalbumin Measured by Triplet-Triplet-Energy-Transfer, Master's Thesis, Universität Basel 2007,

[184] Crick, F. H.: On protein synthesis, Symp Soc Exp Biol 1958, 12:138-63

[185] Crick, F. H., Barnett, L., Brenner, S., Watts-Tobin, R. J.: General nature of the genetic code for proteins, Nature 1961, 192:1227-32

[186] Mann, M., Jensen, O. N.: Proteomic analysis of post-translational modifications, Nat Biotechnol 2003, 21:255-61

[187] Brawerman, G., Ycas, M.: Incorporation of the amino acid analog tryptazan into the protein of Escherichia coli, Arch Biochem Biophys 1957, 68:112-7

[188] Hendrickson, T. L., de Crecy-Lagard, V., Schimmel, P.: Incorporation of nonnatural amino acids into proteins, Annu Rev Biochem 2004, 73:147-76

[189] Wiltschi, B., Merkel, L., Budisa, N.: Fine tuning the N-terminal residue excision with methionine analogues, Chembiochem 2009, 10:217-20

[190] Wang, L. M., TJ. Liu, DR. Schultz, PG.: A new functional suppressor tRNA/aminoacyl-tRNA synthetase pair for the in vivo incorporation on unnatural amino acids into proteins. , J Am Chem Soc 2000, 122:5010-1

[191] Liu, C. C., Schultz, P. G.: Adding new chemistries to the genetic code, Annu Rev Biochem 2010, 79:413-44

[192] Wu, N., Deiters, A., Cropp, T. A., King, D., Schultz, P. G.: A genetically encoded photocaged amino acid, J Am Chem Soc 2004, 126:14306-7

[193] Zhang, Z., Alfonta, L., Tian, F., Bursulaya, B., Uryu, S., King, D. S., Schultz, P. G.: Selective incorporation of 5-hydroxytryptophan into proteins in mammalian cells, Proceedings of the National Academy of Sciences of the United States of America 2004, 101:8882-7

[194] Young, T. S., Ahmad, I., Yin, J. A., Schultz, P. G.: An enhanced system for unnatural amino acid mutagenesis in E. coli, Journal of molecular biology 2010, 395:361-74

[195] Anderson, J. C., Wu, N., Santoro, S. W., Lakshman, V., King, D. S., Schultz, P. G.: An expanded genetic code with a functional quadruplet codon, Proceedings of the National Academy of Sciences of the United States of America 2004, 101:7566-71

[196] Neumann, H., Peak-Chew, S. Y., Chin, J. W.: Genetically encoding N(epsilon)-acetyllysine in recombinant proteins, Nat Chem Biol 2008, 4:232-4

[197] Chatterjee, A., Xiao, H., Schultz, P. G.: Evolution of multiple, mutually orthogonal prolyl-tRNA synthetase/tRNA pairs for unnatural amino acid mutagenesis in Escherichia coli, Proceedings of the National Academy of Sciences of the United States of America 2012, 109:14841-6

[198] Lepthien, S., Merkel, L., Budisa, N.: In vivo double and triple labeling of proteins using synthetic amino acids, Angew Chem Int Ed Engl 2010, 49:5446-50

[199] Wan, W., Huang, Y., Wang, Z., Russell, W. K., Pai, P. J., Russell, D. H., Liu, W. R.: A facile system for genetic incorporation of two different noncanonical amino acids into one protein in Escherichia coli, Angew Chem Int Ed Engl 2010, 49:3211-4

[200] Wang, Q., Parrish, A. R., Wang, L.: Expanding the genetic code for biological studies, Chem Biol 2009, 16:323-36

[201] Nilsson, B. L., Kiessling, L. L., Raines, R. T.: Staudinger ligation: a peptide from a thioester and azide, Org Lett 2000, 2:1939-41

[202] Devaraj, N. K., Weissleder, R., Hilderbrand, S. A.: Tetrazine-based cycloadditions: application to pretargeted live cell imaging, Bioconjug Chem 2008, 19:2297-9

[203] Rostovtsev, V. V., Green, L. G., Fokin, V. V., Sharpless, K. B.: A stepwise huisgen cycloaddition process: copper(I)-catalyzed regioselective "ligation" of azides and terminal alkynes, Angew Chem Int Ed Engl 2002, 41:2596-9

[204] Tron, G. C., Pirali, T., Billington, R. A., Canonico, P. L., Sorba, G., Genazzani, A. A.: Click chemistry reactions in medicinal chemistry: applications of the 1,3-dipolar cycloaddition between azides and alkynes, Med Res Rev 2008, 28:278-308

[205] Graham, R., Lewis, J. R.: Synthesis of 9-oxoxanthen-2-carboxylic acids., J Chem Soc Perkin Trans 1 1978, 876-81

[206] Pace, C. N.: Determination and analysis of urea and guanidine hydrochloride denaturation curves, Meth Enzymol 1986, 131:266-80

[207] Luo, P., Baldwin, R. L.: Mechanism of helix induction by trifluoroethanol: a framework for extrapolating the helix-forming properties of peptides from trifluoroethanol/water mixtures back to water, Biochemistry 1997, 36:8413-21

[208] Findeisen, M., Brand, T., Berger, S.: A 1H-NMR thermometer suitable for cryoprobes, Magnetic resonance in chemistry : MRC 2007, 45:175-8

[209] Liu, M. M., X.; He, C.; Huang, H.; Nicholson, J.K.; Lindon, J.C.: Improved WATERGATE Pulse Sequences for Solvent Suppression in NMR Spectroscopy, J Magn Reson 1998, 132:125-9

[210] Bax, A. D., D.G.: MLEV-17-based two-dimensional homonuclear magnetization transfer spectroscopy, J Magn Reson 1985, 65:355-60

[211] Davis, A. L. K., J; Laue, E.D.; Moskau, D.: Experiments for recording pure-absorption heteronuclear correlation spectra using pulsed field gradients, J Magn Reson 1992, 98:207-16

[212] Gong, H., Rose, G. D.: Assessing the solvent-dependent surface area of unfolded proteins using an ensemble model, Proc Natl Acad Sci U S A 2008, 105:3321-6

[213] Saiki, R. K., Gelfand, D. H., Stoffel, S., Scharf, S. J., Higuchi, R., Horn, G. T., Mullis, K. B., Erlich, H. A.: Primer-directed enzymatic amplification of DNA with a thermostable DNA polymerase, Science 1988, 239:487-91

[214] Hösl, M.: PhD. thesis, Technische Universität München 2011,

[215] Laemmli, U. K.: Cleavage of structural proteins during the assembly of the head of bacteriophage T4, Nature 1970, 227:680-5

[216] Flory, P. J.: Moments of End-to-End Vector of a Chain Molecule, Its Persistence and Distribution, Proceedings of the National Academy of Sciences of the United States of America 1973, 70:1819-23

[217] Schwalbe, H., Fiebig, K. M., Buck, M., Jones, J. A., Grimshaw, S. B., Spencer, A., Glaser, S. J., Smith, L. J., Dobson, C. M.: Structural and Dynamical Properties of a Denatured Protein. Heteronuclear 3D NMR Experiments and Theoretical Simulations of Lysozyme in 8 M Urea, Biochemistry 1997, 36:8977-91

[218] Klein-Seetharaman, J., Oikawa, M., Grimshaw, S. B., Wirmer, J., Duchardt, E., Ueda, T., Imoto, T., Smith, L. J., Dobson, C. M., Schwalbe, H.: Long-range interactions within a nonnative protein, Science 2002, 295:1719-22

[219] Perico, A., Beggiato, M.: Intramolecular Diffusion-Controlled Reactions in Polymers in the Optimized Rouse Zimm Approach .1. The Effects of Chain Stiffness, Reactive Site Positions, and Site Numbers, Macromolecules 1990, 23:797-803

[220] Cordier, F., Hartmann, B., Rogowski, M., Affolter, M., Grzesiek, S.: DNA recognition by the brinker repressor--an extreme case of coupling between binding and folding, J Mol Biol 2006, 361:659-72

[221] Garcia-Pino, A., Balasubramanian, S., Wyns, L., Gazit, E., De Greve, H., Magnuson, R. D., Charlier, D., van Nuland, N. A., Loris, R.: Allostery and intrinsic disorder mediate transcription regulation by conditional cooperativity, Cell 2010, 142:101-11

[222] Gromiha, M. M., Selvaraj, S.: Influence of medium and long range interactions in protein folding, Prep Biochem Biotechnol 1999, 29:339-51

[223] Krieger, F., Möglich, A., Kiefhaber, T.: Effect of proline and glycine residues on dynamics and barriers of loop formation in polypeptide chains, J Am Chem Soc 2005, 127:3346-52

[224] Friedman, B., O'Shaughnessy, B.: Theory of Intramolecular Reactions in Polymeric Liquids, Macromolecules 1993, 26:4888-98

[225] Ortiz-Repiso, M., Rey, A.: Intramolecular Reaction Rates of Flexible Polymers. 1. Simulation Results and the Classical Theory, Macromolecules 1998, 31:8356-62

[226] Ortiz-Repiso, M., Rey, A.: Intramolecular Reaction Rates of Flexible Polymers. 2. Comparison with the Renormalization Group Theory, Macromolecules 1998, 31:8363-9

[227] Oberer, M., Zangger, K., Gruber, K., Keller, W.: The solution structure of ParD, the antidote of the ParDE toxin antitoxin module, provides the structural basis for DNA and toxin binding, Protein science : a publication of the Protein Society 2007, 16:1676-88

[228] Drobnak, I., De Jonge, N., Haesaerts, S., Vesnaver, G., Loris, R., Lah, J.: Energetic basis of uncoupling folding from binding for an intrinsically disordered protein, J Am Chem Soc 2013, 135:1288-94

[229] Arrhenius, S.: Über die Reaktionsgeschwindigkeit bei der Inversion von Rohrzucker durch Säuren., Z Phys Chem 1889, 4:226-48

[230] Handbook of Chemistry and Physics. Edited by Weast, R. C. Cleveland, OH, CRC, 1972,

[231] Bieri, O., Kiefhaber, T.: Kinetic models in protein folding. Edited by Pain, R. Oxford, Oxford University Press, 2000, p. pp. 34-64

[232] Hagen, S. J., Eaton, W. A.: Two-state expansion and collapse of a polypeptide, Journal of molecular biology 2000, 301:1019-27

[233] Klimov, V., Thirumalai, D.: Viscosity dependence of the folding rates of proteins, Phys Rev Lett 1997, 79:317-20

[234] Bieri, O., Kiefhaber, T.: Elementary steps in protein folding, Biol Chem 1999, 380:923-9

[235] Haas, E., Wilchek, M., Katchalski-Katzir, E., Steinberg, I. Z.: Distribution of end-to-end distances of oligopeptides in solution as estimated by energy transfer, Proceedings of the National Academy of Sciences of the United States of America 1975, 72:1807-11

[236] Kuhn, W. K., H.: Bedeutung beschränkt freier Drehbarkeit für die Strömungsdoppelbrechung von Fadenmolekellösungen I. , Helvetica Chimica Acta 1945, 28:1533-79

[237] Jas, G. S. E., W. A.; Hofrichter, J.: Effect of Viscosity on the Kinetics of alpha-Helix and beta-Hairpin Formation., J Phys Chem B 2012, 105:261-72

[238] Schulz, J. C., Schmidt, L., Best, R. B., Dzubiella, J., Netz, R. R.: Peptide chain dynamics in light and heavy water: zooming in on internal friction, J Am Chem Soc 2012, 134:6273-9

[239] Soranno, A., Buchli, B., Nettels, D., Cheng, R. R., Muller-Spath, S., Pfeil, S. H., Hoffmann, A., Lipman, E. A., Makarov, D. E., Schuler, B.: Quantifying internal friction in unfolded and intrinsically disordered proteins with single-molecule spectroscopy, Proc Natl Acad Sci U S A 2012, 109:17800-6

[240] Birks, J. B.: Excimers, Rep Prog Phys 1975, 38:903-74

[241] Cochran, A. G., Skelton, N. J., Starovasnik, M. A.: Tryptophan zippers: stable, monomeric beta -hairpins, Proc Natl Acad Sci U S A 2001, 98:5578-83

[242] Andersen, N. H., Olsen, K. A., Fesinmeyer, R. M., Tan, X., Hudson, F. M., Eidenschink, L. A., Farazi, S. R.: Minimization and optimization of designed beta-hairpin folds, J Am Chem Soc 2006, 128:6101-10

[243] Eidenschink, L., Crabbe, E., Andersen, N. H.: Terminal sidechain packing of a designed beta-hairpin influences conformation and stability, Biopolymers 2009, 91:557-64

[244] Wu, L., McElheny, D., Huang, R., Keiderling, T. A.: Role of tryptophan-tryptophan interactions in Trpzip beta-hairpin formation, structure, and stability, Biochemistry 2009, 48:10362-71

[245] Eidenschink, L., Kier, B. L., Huggins, K. N., Andersen, N. H.: Very short peptides with stable folds: building on the interrelationship of Trp/Trp, Trp/cation, and Trp/backbone-amide interaction geometries, Proteins 2009, 75:308-22

[246] Mirassou, Y., Santiveri, C. M., Perez de Vega, M. J., Gonzalez-Muniz, R., Jimenez, M. A.: Disulfide bonds versus TrpTrp pairs in irregular beta-hairpins: NMR structure of vammin loop 3-derived peptides as a case study, Chembiochem 2009, 10:902-10

[247] Hatfield, M. P., Murphy, R. F., Lovas, S.: VCD spectroscopic properties of the beta-hairpin forming miniprotein CLN025 in various solvents, Biopolymers 2010, 93:442-50

[248] Fierz, B.: Dynamics of Unfolded and alpha-Helical Polypeptide Chains., PhD thesis, Universität Basel 2006,

[249] Sekhar, A., Latham, M. P., Vallurupalli, P., Kay, L. E.: Viscosity-dependent kinetics of protein conformational exchange: microviscosity effects and the need for a small viscogen, J Phys Chem B 2014, 118:4546-51

[250] Gekko, K., Timasheff, S. N.: Mechanism of protein stabilization by glycerol: preferential hydration in glycerol-water mixtures, Biochemistry 1981, 20:4667-76

[251] Gekko, K., Timasheff, S. N.: Thermodynamic and kinetic examination of protein stabilization by glycerol, Biochemistry 1981, 20:4677-86

[252] Anjana, R., Vaishnavi, M. K., Sherlin, D., Kumar, S. P., Naveen, K., Kanth, P. S., Sekar, K.: Aromatic-aromatic interactions in structures of proteins and protein-DNA complexes: a study based on orientation and distance, Bioinformation 2012, 8:1220-4

[253] Santiveri, C. M., Jimenez, M. A.: Tryptophan residues: scarce in proteins but strong stabilizers of beta-hairpin peptides, Biopolymers 2010, 94:779-90

[254] Wang, G., Dunbrack, R. L., Jr.: PISCES: a protein sequence culling server, Bioinformatics 2003, 19:1589-91

[255] Porter, L. L., Rose, G. D.: Comment on "Revisiting the Ramachandran plot from a new angle", Protein science : a publication of the Protein Society 2011, 20:1771-3; author reply 4

[256] Zhou, A. Q., O'Hern, C. S., Regan, L.: Revisiting the Ramachandran plot from a new angle, Protein Sci 2011, 20:1166-71

[257] Cierpicki, T., Otlewski, J.: Amide proton temperature coefficients as hydrogen bond indicators in proteins, J Biomol NMR 2001, 21:249-61

[258] Baxter, N. J., Williamson, M. P.: Temperature dependence of ^{1}H chemical shifts in proteins, J Biomolec NMR 1997, 9:359-69

[259] Reiner, A., Wildemann, D., Fischer, G., Kiefhaber, T.: Effect of thioxopeptide bonds on alpha-helix structure and stability, J Am Chem Soc 2008, 130:8079-84

[260] Huber, R., Steigemann, W.: Two cis-prolines in the Bence-Jones protein Rei and the cis-pro-bend, FEBS Lett 1974, 48:235-7

[261] Halab, L., Lubell, W. D.: Use of Steric Interactions To Control Peptide Turn Geometry. Synthesis of Type VI beta-Turn Mimics with 5-tert-Butylproline, J Org Chem 1999, 64:3312-21

[262] Simpson, R. B., Kauzman, W.: The kinetics of protein denaturation. I. The behavior of the optical rotation of ovalbumin in urea solutions, J Am Chem Soc 1953, 75:5139-52

[263] Frensdorff, K. H., Watson, M. T., Kauzman, W.: The kinetics of protein denaturation. V. The viscosity of urea solutions of serum albumin, J Am Chem Soc 1953, 75:5167-72

[264] Makhatadze, G. I.: Thermodynamics of protein interactions with urea and guanidinium hydrochloride, J Phys Chem 1999, 103:4781-5

[265] Perl, D., Jacob, M., Bánó, M., Stupâk, M., Antalík, M., Schmid, F. X.: Thermodynamics of a diffusional protein folding reaction, Biophys Chem 2002, 2-3:173-90

[266] Smith, J. S., Scholtz, J. M.: Guanidine hydrochloride unfolding of peptide helices: separation of denaturant and salt effects, Biochemistry 1996, 35:7292-7

[267] Mason, P. E., Dempsey, C. E., Neilson, G. W., Kline, S. R., Brady, J. W.: Preferential interactions of guanidinum ions with aromatic groups over aliphatic groups, J Am Chem Soc 2009, 131:16689-96

[268] Fierz, B., Joder, K., Krieger, F., Kiefhaber, T.: Using triplet-tiplet energy transfer to measure conformational dynamics in polypeptide chains. Edited by Bai, Y., Nussinov, R. Humana Press, 2006, p.

[269] Gromiha, M. M., Selvaraj, S.: Importance of long-range interactions in protein folding, Biophys Chem 1999, 77:49-68

[270] Permyakov, S. E., Bakunts, A. G., Denesyuk, A. I., Knyazeva, E. L., Uversky, V. N., Permyakov, E. A.: Apo-parvalbumin as an intrinsically disordered protein, Proteins 2008, 72:822-36

[271] Reiner, A.: Conformational Dynamics and Stability of Structured Peptides and Small Proteins, PhD thesis, Universität Basel 2007,

[272] Munoz, V., Serrano, L.: Elucidating the folding problem of helical peptides using empirical parameters, Nature Struct Biol 1994, 1:399-409

[273] Munoz, V., Serrano, L.: Elucidating the folding problem of helical peptides using empirical paramters. II. Helix macrodipole effects and rational modification of the helical content of natural peptides, J Mol Biol 1995, 245:275-96

[274] Munoz, V., Serrano, L.: Elucidating the folding problem of helical peptides using empirical parameters. III. Temperature and pH dependence, J Mol Biol 1995, 245:297-308

[275] Munoz, V., Serrano, L.: Development of the multiple sequence approximation within the AGADIR model of alpha-helix formation: comparison with Zimm-Bragg and Lifson-Roig formalisms, Biopolymers 1997, 41:495-509

[276] Lacroix, E., Viguera, A. R., Serrano, L.: Elucidating the folding problem of alpha-helices: local motifs, long-range electrostatics, ionic-strength dependence and prediction of NMR parameters, J Mol Biol 1998, 284:173-91

[277] LaVallie, E. R., DiBlasio, E. A., Kovacic, S., Grant, K. L., Schendel, P. F., McCoy, J. M.: A thioredoxin gene fusion expression system that circumvents inclusion body formation in the E. coli cytoplasm, Biotechnology (N Y) 1993, 11:187-93

[278] Moncrieffe, M. C., Juranic, N., Kemple, M. D., Potter, J. D., Macura, S., Prendergast, F. G.: Structure-fluorescence correlations in a single tryptophan mutant of carp parvalbumin: solution structure, backbone and side-chain dynamics, J Mol Biol 2000, 297:147-63

[279] Rohl, C. A., Chakrabartty, A., Baldwin, R. L.: Helix propagation and N-cap propensities of the amino acids measured in alanine-based peptides in 40 volume percent trifluoroethanol, Protein Sci 1996, 5:2623-37

[280] Doig, A. J., Errington, N., Iqbalsyah, T. M.: Stability and Design of alpha-Helices. Edited by Buchner, J., Kiefhaber, T. Weinheim, WILEY-VCH, 2005, p. pp. 247-313

[281] Vila, J. A., Ripoll, D. R., Scheraga, H. A.: Physical reasons for the unusual alpha-helix stabilization afforded by charged or neutral polar residues in alanine-rich peptides, Proceedings of the National Academy of Sciences of the United States of America 2000, 97:13075-9

[282] Luo, P., Baldwin, R. L.: Interaction between water and polar groups of the helix backbone: an important determinant of helix propensities, Proceedings of the National Academy of Sciences of the United States of America 1999, 96:4930-5

[283] Hwang, P. M., Pan, J. S., Sykes, B. D.: Targeted expression, purification, and cleavage of fusion proteins from inclusion bodies in Escherichia coli, FEBS Lett 2014, 588:247-52

[284] Schwarzer, L.: Bachelor thesis, Universität Innsbruck 2014,