

# TECHNISCHE UNIVERSITÄT MÜNCHEN

Lehrstuhl für Biophysikalische Chemie

## Effects of Osmolytes and other Co-Solutes on the Dynamics of Loop Formation in Proteins and Peptides

Christian Nyffenegger

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. A. Itzen  
Prüfer der Dissertation:  
1. Univ.-Prof. Dr. T. Kiefhaber  
2. Univ.-Prof. Dr. M. Groll

Die Dissertation wurde am 23.08.2012 bei der Technischen Universität München  
eingereicht und durch die Fakultät für Chemie  
am 24.09.2012 angenommen.

---

# Contents

<b>1. Introduction</b>	<b>1</b>
1.1. Proteins . . . . .	1
1.2. The Native State of Proteins . . . . .	1
1.3. The Unfolded State of Proteins . . . . .	3
1.4. Protein Folding Kinetics . . . . .	6
1.4.1. Mechanisms of Protein Folding . . . . .	6
1.4.2. Protein Folding Barriers . . . . .	7
1.4.3. Effect of Friction on Chemical Reactions . . . . .	9
1.5. Protein Stability . . . . .	10
1.6. The Role of Osmolytes and other Co-Solutes in Protein Stability . . . . .	11
1.6.1. The Tanford Transfer Model . . . . .	12
1.6.2. Schellman's Weak Binding Model . . . . .	14
1.7. Loop Formation in Polypeptide Chains . . . . .	15
1.7.1. Triplet-Triplet Energy Transfer to Study Loop Formation . . . . .	16
1.7.2. Dynamics of Intramolecular Loop Formation in Polypeptide Chains . . . . .	19
1.8. Carp $\beta$ -Parvalbumin, as a Model to study Loop Formation by TTET in Different States of a Protein . . . . .	24
1.8.1. Loop Formation in Protein Fragments of Carp $\beta$ -Parvalbumin . . . . .	25
1.8.2. Introducing Non-Natural Amino Acids into Proteins by Native Chemical Ligation . . . . .	26
<b>2. Aim of Research</b>	<b>29</b>
<b>3. Material and Methods</b>	<b>31</b>
3.1. Used Materials . . . . .	31
3.2. Solid-Phase Peptide Synthesis . . . . .	31
3.3. Peptide Modification . . . . .	32

3.3.1.	Synthesis and Introduction of Chromophores for Triplet-Triplet Energy Transfer . . . . .	32
3.3.2.	Acetylation of the Peptide Amino-Terminus . . . . .	33
3.3.3.	Peptide Cleavage . . . . .	33
3.4.	Peptide Purification . . . . .	33
3.5.	Molecular Biology Methods . . . . .	34
3.5.1.	Site-Directed Mutagenesis . . . . .	34
3.5.2.	Preparation of Chemically Competent Cells . . . . .	35
3.5.3.	Transformation of Chemically Competent Cells . . . . .	35
3.6.	Protein Expression, Purification and Modification . . . . .	36
3.6.1.	Protein Expression in <i>E. coli</i> . . . . .	36
3.6.2.	PV <sup>1-77</sup> Purification and Intein Mediated Thioester Formation . . . . .	36
3.6.3.	PV C18S/H26F Purification . . . . .	38
3.6.4.	Native Chemical Ligation (NCL) . . . . .	38
3.6.5.	Sodium Dodecylsulfate Polyacrylamide Gel Electrophoresis . . . . .	39
3.6.6.	Protein Modification . . . . .	40
3.6.6.1.	Desulfurization of Cysteine Residues . . . . .	40
3.6.6.2.	Thiol Reactive Labeling . . . . .	40
3.7.	Sample Solutions and Viscosity . . . . .	41
3.8.	Calculation of Transfer Free Energies . . . . .	42
3.9.	Laserflash Photolysis . . . . .	42
3.10.	Circular Dichroism Spectroscopy . . . . .	43
3.11.	Equilibrium Transitions . . . . .	43
<b>4.</b>	<b>Results and Discussion</b>	<b>45</b>
4.1.	Chain Dynamics and Barriers for Loop Formation in Unfolded Polypeptide Chains . . . . .	45
4.1.1.	Loop Formation in Model Peptides . . . . .	45
4.1.2.	Effect of Solvent Viscosity on Loop Formation in Model Peptides . . . . .	47
4.1.3.	Barriers for Loop Formation in Model Peptides . . . . .	48
4.1.4.	Chain Dynamics and Barriers for Loop Formation in Natural Sequences . . . . .	51
4.1.5.	Effect of Solvent Viscosity on Loop Formation in Natural Sequences . . . . .	54
4.1.6.	Effect of Temperature on Loop Formation in Natural Sequences . . . . .	55
4.2.	Effect of Osmolytes and other Co-Solutes on the Dynamics of Loop Formation . . . . .	58

---

4.2.1.	Effect of Destabilizing Osmolytes and Co-Solutes on the Dynamics of Loop Formation . . . . .	59
4.2.2.	Comparison of the Effect of Destabilizing Osmolytes and Co-Solutes on the Dynamics of Loop Formation . . . . .	65
4.2.3.	Effect of Stabilizing Osmolytes and Co-Solutes on the Dynamics of Loop Formation . . . . .	69
4.2.4.	Comparison of the Effect of Stabilizing Osmolytes and Co-Solutes on the Dynamics of Loop Formation . . . . .	81
4.2.5.	Effect of Neutral Osmolytes on the Dynamics of Loop Formation . . . . .	82
4.2.6.	Comparison of the Effect of Neutral Osmolytes on the Dynamics of Loop Formation . . . . .	88
4.2.7.	Summary of the Effect of Osmolytes and Co-Solutes on the Dynamics of Loop Formation . . . . .	90
4.2.7.1.	Schellman's Weak Binding Model . . . . .	90
4.2.7.2.	Using Tanford's Transfer Model to Describe the Effect of Osmolytes and other Co-Solutes on Loop Formation . . . . .	94
4.2.7.3.	Summary of the Effect of Osmolytes and Co-Solutes on the Barriers for Loop Formation . . . . .	96
4.2.7.4.	Effect of Solvent Quality on the Dynamics of Loop Formation	100
4.3.	Contributions of Internal Friction and Solvent Friction to the Dynamics of Loop Formation in Unfolded Polypeptide Chains . . . . .	101
4.3.1.	Internal Friction in the Dynamics for Loop Formation . . . . .	109
4.4.	Dynamics of Loop Formation in the EF-Hand and Full Length Carp $\beta$ -Parvalbumin in the Unfolded and an Intermediate State . . . . .	112
4.4.1.	Loop Formation in Different Fragments from the EF-Hand Motif of Carp $\beta$ -Parvalbumin . . . . .	112
4.4.2.	Loop Formation in the Unfolded and in an Intermediate State of Full Length Carp $\beta$ -Parvalbumin . . . . .	115
4.4.2.1.	Semisynthesis of Carp $\beta$ -Parvalbumin for TTET Measurements . . . . .	116
4.4.2.2.	Loop Formation in the Unfolded and an Intermediate State of PV . . . . .	123
4.4.3.	Comparison of Loop Formation Dynamics in PV Fragments and Full Length PV . . . . .	126

## *Contents*

---

<b>5. Summary</b>	<b>129</b>
<b>A. Appendix</b>	<b>133</b>
<b>List of Figures</b>	<b>137</b>
<b>List of Tables</b>	<b>141</b>
<b>List of Abbreviations</b>	<b>143</b>
<b>Bibliography</b>	<b>147</b>
<b>Acknowledgements</b>	<b>165</b>
<b>Lebenslauf</b>	<b>167</b>

# 1. Introduction

## 1.1. Proteins

Proteins represent the major class of biomolecules in living organisms. They are responsible for a multitude of biological processes<sup>1</sup> such as catalysis, signal transduction, regulation, and transport. Proteins are composed of 20 canonical  $\alpha$ -L-amino acids, which are linked together by peptide bonds. Although the number of the amino acid building blocks is limited, proteins can fulfill the most various functions. To acquire biological function, the polypeptide molecule has to adopt a well-defined three dimensional structure, also referred to as the native state (N), in the protein folding process. The native state structure is fully encoded by the primary sequence, *i.e.* folding to N occurs spontaneously<sup>2</sup> but may depend on the binding of ligands<sup>3-5</sup>. Non-covalent interactions such as electrostatics, hydrogen bonds, hydrophobic interactions and van-der-Waals interactions are formed intramolecularly and with the surrounding solvent, thereby determining the native state structure of proteins and its stability.

Although folding occurs spontaneously, it may be catalyzed by disulfide isomerases<sup>6</sup> and peptidyl-prolyl isomerases<sup>7</sup>, which ensure efficient folding *in vivo*. Protein misfolding might lead to aggregation into nonfunctional structures, and is associated with several diseases such as Alzheimer and Creutzfeldt-Jacob disease<sup>8</sup>.

The goal of studying protein folding is to elucidate, which mechanisms lead to the native state of proteins, and how folding is coupled to functionality.

## 1.2. The Native State of Proteins

The native state (N) of proteins describes the thermodynamic state populated by proteins with a defined conformation under folding conditions. N is most often associated with the biological function in proteins. An exception are disordered proteins which are intrinsically unstructured under ambient conditions<sup>9</sup>. Structures of the native state of a vast number of proteins could be resolved with atomic resolution using X-ray diffraction crystallography<sup>10</sup>

## 1. Introduction

---

and nuclear magnetic resonance (NMR) for proteins in solution<sup>11-13</sup>. Structural elements of different hierarchical levels can be found in structures of native proteins. The primary structure is formed by the unbranched polypeptide chain. Interactions between groups close in primary structure (short-range interactions), mostly hydrogen bonds between backbone amide groups, are responsible for the formation and stabilization of the most common secondary structural elements like  $\alpha$ -helices and  $\beta$ -sheets<sup>14,15</sup>. The third level of structural hierarchy (tertiary structure) describes the spatial arrangement of secondary structure elements, stabilized by long-range interactions. They are packed together in order to avoid an unfavorable loss in solvent entropy, which would arise from the ordering of water molecules around solvent exposed hydrophobic side chains<sup>16,17</sup>. This driving force, known as the hydrophobic effect, leads to the burial of hydrophobic groups and the formation of a hydrophobic core in proteins. A tertiary structured unit, which can fold independently is called a domain. Domains remain folded in isolation and have a size rarely exceeding 200 amino acid residues<sup>18</sup>. The largest known proteins consist of several hundred domains forming proteins with mass weights up to 3 MDa<sup>19,20</sup>.

Native state structures in protein crystals and in solution are often similar with only minor differences<sup>21</sup>. This is a consequence from the location of the native state in a narrow free energy minimum, which results in a narrow conformational distribution of structures. However, folded proteins show flexibility over different magnitudes of motion and timescales. They are ranging from thermal fluctuations and group movements within femtoseconds and picoseconds up to large scale rearrangements or induced fit mechanisms in enzymes within microseconds and milliseconds, respectively<sup>22</sup>. Native states of proteins are thus not as rigid as one would expect from the static X-ray pictures. The native state rather comprises a collection of structurally similar conformational sub-states, where dynamics play an essential role in its function and movement on the free energy surface<sup>23</sup>.

Although a vast number of native state structures has been solved, the driving forces for attaining a specific three dimensional structure are largely unknown. Atoms or groups with particular properties show only limited tendencies to be found at specific positions in the native state<sup>24</sup>. Knowledge about the structure of the native state is thus not sufficient for understanding, how the native state is formed from the ensemble of molecules in the unfolded state.



## 1.3. The Unfolded State of Proteins

The unfolded state (U) differs from the well defined native state. It is hard to characterize U under conditions which favor N, since only an extremely small number of unfolded protein molecules exist. Therefore the unfolded state appears to be transient. However, U can be populated under conditions which favor unfolding. Denaturing conditions like acidic or alkaline pH, high temperature and pressure, or denaturants may lead to the population of U, constituting a defined state over a wide range of denaturing conditions<sup>25,26</sup>. It is not clear, whether the unfolded state under denaturing conditions is comparable to the unfolded state under conditions which favor the native state. (The most stable unfolded form under native conditions was speculated to be a collapsed but dynamic state (molten globule), rather than being comparable to the unfolded state under denaturing conditions<sup>24</sup>.) Properties of the unfolded state might be investigated under native conditions, using protein fragments<sup>27-32</sup>, short and flexible model peptides<sup>33-37</sup>, as well as proteins destabilized by reduced disulfide bonds or mutations<sup>38</sup>.

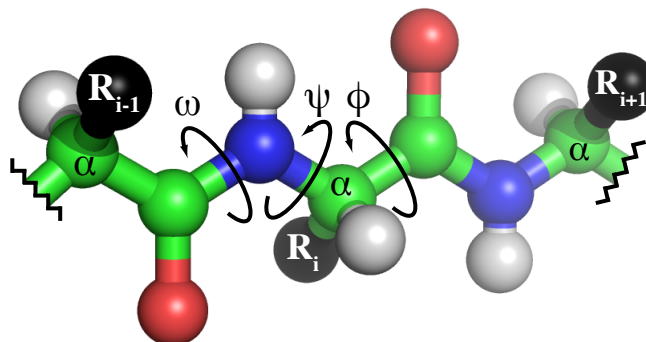
In strong denaturing conditions Tanford and others found the polypeptide chain to be highly expanded and solvent exposed. They concluded, that the unfolded chain in denaturing conditions could be described as a structureless random coil<sup>39-41</sup>. A random coil or freely jointed chain is constituted by a number of  $n$  bonds, each having a fixed length  $l$ <sup>42</sup>. The chain segments are thought to perform a random walk in three dimensional space (random flight). Angles at bond junctions may adopt all values with identical probability, non-correlated with one another. Given these requirements, a vast number of conformations with identical energies can be sampled by the chain. The root mean square end-to-end distance  $\langle \mathbf{r}^2 \rangle$  for all conformations can be written as

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

with the end-to-end vector  $\mathbf{r}$ . Brackets denote the statistical average over all conformations. The distribution of end-to-end distances  $W(\mathbf{r})$  is shown to be Gaussian.

$$W(\mathbf{r}) = \left( \frac{3}{2\pi\langle \mathbf{r}^2 \rangle} \right)^{3/2} \cdot \exp \left( -\frac{3\mathbf{r}^2}{2\langle \mathbf{r}^2 \rangle} \right) \quad (1.2)$$

This simple chain model fails to describe all features of unstructured polypeptide chains. In polypeptides, the number of possible angles at bond junctions constituting the peptide backbone is restricted. The peptide dihedral angles  $\omega$ ,  $\psi$  and  $\phi$  (Figure 1.1) specify the geometry for each amino acid in the chain and define the overall conformation. Ramachandran *et al.*



**Figure 1.1.:** Structure of the peptide backbone with side chains represented by  $R_i$ . Dihedral angles  $\omega$ ,  $\psi$  and  $\phi$  are indicated. Carbon atoms colored in green.  $C_\alpha$  atoms labeled with  $\alpha$ . Nitrogen atoms colored in blue, oxygen in red and hydrogen in white.

constructed maps with allowed pairs for  $\phi$  and  $\psi$ , displaying the dependence of each other<sup>43</sup>. Flory assumed, that each  $\phi, \psi$  pair is independent from pairs of other chain segments<sup>42</sup>. However, calculations showed that this assumption does not hold and that the accessible conformational space is restricted<sup>44</sup>. If hydrogen bonding constraints are taken into account, the region for allowed pairs of angles is further restricted<sup>45</sup>, thereby lowering the conformational entropy of  $U$  in favor of the native state. These restrictions result in stiffer chains than would be expected for a random coil. As a consequence, the end-to-end distance for a polypeptide is larger than calculated from Equation 1.1. Accounting for this phenomenon, Flory introduced the characteristic ratio  $C_n$ ,

$$\langle \mathbf{r}^2 \rangle = C_n \cdot nl^2 \quad (1.3)$$

a measure for the ideality of a polymer. Freely jointed long chains have a characteristic ratio of one, yielding Equation 1.1. In non-ideal chains,  $C_n$  is larger than unity and depends on  $n$ . With increasing length,  $C_n$  approaches a limiting value, yielding the constant  $C_\infty$  for very long chains ( $n \rightarrow \infty$ ). Equation 1.3 can be rewritten.

$$\langle \mathbf{r}^2 \rangle = C_\infty \cdot nl^2 \quad (1.4)$$

In homo-polypeptides of the canonical  $\alpha$ -L-amino acids, the most flexible poly-glycine chain exhibits the lowest value for  $C_\infty$  while the stiff poly-proline chain exhibits the highest value for  $C_\infty$ .

In addition to interactions between groups close in the polymer chain, interactions between chain segments distant along the chain (long-range effect) need to be considered. Since real

polymers like polypeptide chains have a finite volume, two segments cannot occupy the same spatial volume. This repulsive interaction called the excluded volume effect leads to a swelling of the polymer<sup>46</sup>. Edwards developed a statistical chain model, which accounts for excluded volume effects<sup>47</sup> (but not attractive forces). His model describes the probability to find the  $n^{\text{th}}$  chain segment along the polymer at a distance  $r$  from the origin. Considering excluded volume effects, Haas *et al.*<sup>48</sup> used a skewed Gaussian function (Equation 1.5) as an approximation to describe the end-to-end distance distribution,

$$W(r) \propto 4\pi r^2 \cdot \exp\left(-\frac{r - r_0}{\sigma}\right)^2 \quad (1.5)$$

where  $r$  is the end-to-end distance.  $\sigma$  is related to the half-width of the distribution and  $r_0$  is proportional to the excluded volume.

Many attempts have been made to describe the dynamics of polymers<sup>49-51</sup>. Rouse proposed to model the dynamics of a polymer in solution by representing the polymer by a set of beads along the chain connected by springs<sup>49</sup>. Brownian motion of these beads define the chain's dynamic behavior. An extension for this model was worked out by Zimm<sup>50</sup>, where hydrodynamic interactions between different parts of the chain and the solvent are taken into account. A pronounced event in polymer dynamics is the intrachain loop formation. The probability for loop formation ( $p_{\text{loop}}$ ) was calculated by Jacobson and Stockmayer<sup>52</sup>. For an ideal chain,  $p_{\text{loop}}$  can be shown to be a function of the number of bonds (see also Equation 1.1)

$$p_{\text{loop}} \propto n^{-3/2} \quad (1.6)$$

In a first passage time approach, Szabo, Schulten and Schulten calculated the time dependent probability distribution for end-to-end distances<sup>51</sup>. Assuming encounters between chain termini being diffusion controlled, the kinetics for loop closure reactions can be described by a single exponential decay

$$\Sigma(t) \approx \Sigma_{\text{approx}}(t) = e^{-t/\tau} \quad (1.7)$$

where the average reaction time  $\tau$  is related to the probability  $\Sigma(t)$  that loop formation not yet happened at time  $t$ . Loop formation thus yields a single rate constant given that the interconversion between chain conformations is fast compared to the loop formation and that the fraction which forms loops at equilibrium conditions is small.

## 1.4. Protein Folding Kinetics

The number of possible conformations for a polypeptide is vast. Assuming each residue could only be found in two configurations, a protein with 100 amino acids would have  $2^{100}$  ( $\sim 10^{30}$ ) possible conformations. Levinthal pointed out in the 60's, that the time needed for a polypeptide chain to sample all the possible conformations in a random search for the native state would be astronomic<sup>53</sup>. Yet proteins fold in milliseconds to seconds<sup>54-56</sup>. He thus concluded that a random search is not an effective way to find the native state and as a consequence, pathways and folding intermediates must exist, which speed up the folding process<sup>53,57</sup>. By introducing a small energy bias in favor for the native state, calculations showed that the folding time reduces to a few seconds<sup>58</sup>.

### 1.4.1. Mechanisms of Protein Folding

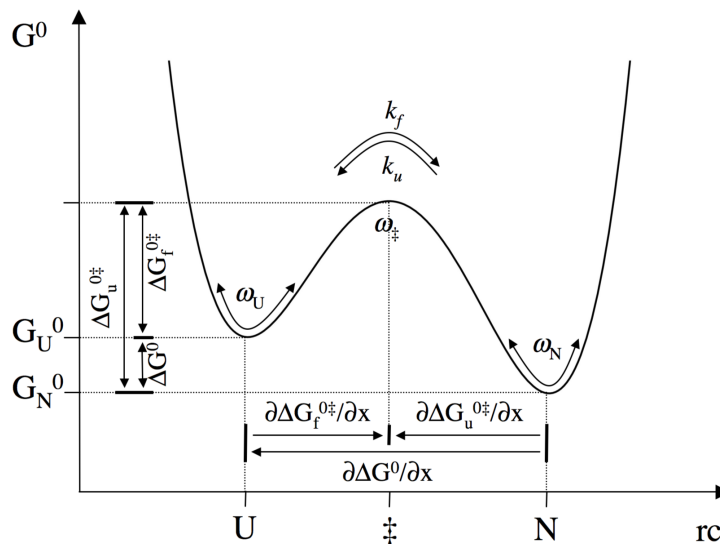
Several hierarchical models exist to describe the protein folding reaction, which uncouple the formation of secondary and tertiary structure. In the nucleation-growth model<sup>18</sup>, secondary structure elements are expected to be formed by neighboring residues in sequence. These elements are thought to act as a nuclei for the folding reaction, from where the structure propagates in a stepwise manner. In the framework model or diffusion collision model<sup>59,60</sup>, local secondary structure elements are speculated to be able to form independently of tertiary structure. The secondary structure elements diffuse together, collide and combine. The formation of a secondary structured folding nucleus can be seen as the rate limiting step in the nucleation-condensation model<sup>61,62</sup>, where the protein is assumed to condense around the nucleus in order to form the native state. A hydrophobic collapse<sup>63</sup> is expected to exist for several proteins, where the polypeptide chain collapses around hydrophobic side chains, followed by a rearrangement of the so-formed intermediate into the native state.

Folding from the unfolded state ensemble can often be described apparently by a two state model,



where  $k_f$  and  $k_u$  are the microscopic rate constants for folding and unfolding, respectively. A single exponential function can be used to describe folding and unfolding kinetics for many small single domain proteins, yielding the single observable apparent rate constant  $\lambda$ , which is the sum of  $k_f$  and  $k_u$ . In apparent two state folders, high energy intermediates may exist.

They are not detectable directly by spectroscopic methods due to their low stability compared to the folded and the unfolded state<sup>64,65</sup>. However, their existence can be verified and they were shown to speed up folding<sup>66</sup>. Deviations from apparent two state folding can be found for proteins with kinetic intermediates, which are populated during the folding reaction. These proteins show multi-exponential kinetics, giving rise to more than a single apparent rate constant. Existing intermediates can be on-pathway or off-pathway, or the folding pathway might be branched, which leads to parallel pathways. Intermediates have been found both for refolding<sup>67</sup> and unfolding<sup>68</sup> in kinetic experiments. They are speculated to be necessary for guiding the protein on the folding pathway, thereby speeding up the folding reaction<sup>69,70</sup>. This leads to a complex free energy surface with defined saddle points, which guide the protein folding reaction.



**Figure 1.2.:** Hypothetical free energy surface for two state folding. U and N are separated along the reaction coordinate by a free energy barrier, with the transition state located on the barrier top. Motions in the wells and at of the transition state are subject to the motional frequencies  $\omega$ .

### 1.4.2. Protein Folding Barriers

The different states of a protein are separated by a major free energy barrier. The characterization of this barrier along with the transition state on the barrier top and eventually existing intermediate states along with their position on the reaction coordinate is necessary to describe the free energy surface for a protein folding reaction. Figure 1.2 shows a hypothetical free energy surface for two state folding. Transition state theory can be applied<sup>71</sup>, where the rate

## 1. Introduction

---

constant for a reaction is shown to depend on the free energy barrier  $\Delta G^{0\dagger}$ .

$$k = k_0 \cdot \exp\left(-\frac{\Delta G^{0\dagger}}{RT}\right) \quad (1.9)$$

The pre-exponential factor  $k_0$  represents the maximum rate constant for a reaction in the absence of free energy barriers. For bimolecular reactions in solution, the upper limit defined by  $k_0$  is set by diffusion of the reaction partners. In reactions, where covalent bonds are formed or broken,  $k_0$  corresponds to the time associated with a single bond vibration for small molecules in the gas phase, with  $k_0 \approx 6 \cdot 10^{12} \text{ s}^{-1}$  at room temperature<sup>71</sup>. For protein folding reactions, the pre-exponential factor depends on the respective protein and is expected to be in the range of  $10^7 - 10^8 \text{ s}^{-1}$ <sup>34</sup>.

To characterize the free energy barrier and the transition state for the protein folding reaction, the relationship between the folding or unfolding rate constant and the respective equilibrium free energy can be consulted. Leffler found a linear relationship between changes in activation free energy and the equilibrium free energy (Rate equilibrium free energy relationship or REFER)<sup>72</sup>. He defined a proportionality constant  $\alpha_x$ , which is a measure for the sensitivity of the transition state relative to the ground states, caused by a perturbation  $\partial x$ .

$$\alpha_x = \frac{\partial \Delta G^{0\dagger} / \partial x}{\partial \Delta G^0 / \partial x} \quad (1.10)$$

An  $\alpha_x = 1$  represents the case, where the transition state has the same property as the product, where in the case of  $\alpha_x = 0$ , the transition state resembles the educt. By rewriting the Gibb's Fundamental Equation (Equation 1.17) it can be applied to the activation free energy of folding ( $\Delta G_f^{0\dagger}$ ) and unfolding ( $\Delta G_u^{0\dagger}$ ), assuming a transition state between U and N.

$$\Delta G_{f,u}^{0\dagger} = \Delta V_{f,u}^{0\dagger} dp - \Delta S_{f,u}^{0\dagger} dT + m_{f,u} \cdot [D] \quad (1.11)$$

From Equation 1.11 can be seen, that  $\Delta G_f^{0\dagger}$  and  $\Delta G_u^{0\dagger}$  are subject to different perturbations. By varying the pressure ( $p$ ) at constant temperature, an  $\alpha_p$  can be obtained which gives information about the volume of the transition state ( $\Delta V_f^{0\dagger}$ ). Informations about the entropy ( $\Delta S_f^{0\dagger}$ ) or heat capacity ( $\Delta C_{p,f}^{0\dagger}$ ) of the transition state can be obtained accordingly by varying the temperature at constant pressure ( $\alpha_T$ ). In addition to the above mentioned REFERs which are medium- or solvent-induced, structural information of the transition state can be obtained by amino acid exchanges ( $\phi_f$ -values). In several cases, deviation from linear REFERs can

be observed<sup>64,73–75</sup>. In that case, transition state movement along the reaction coordinate or changes in the unfolded state as well as the existence of sequential or parallel transition states might be responsible for the non-linear behavior. Characterizing the transition state can be further achieved for ligand binding proteins as a function of ligand concentration. If the ligand is bound in the transition state, folding proceeds faster in the presence of ligands than in the absence.

### 1.4.3. Effect of Friction on Chemical Reactions

Considering the influence of the solvent on chemical reactions, Kramers developed a transition state theory, where the crossing of a hypothetical particle over a barrier of height  $E$  is related to Brownian motion<sup>76</sup>. The probability for a transition of the particle between the two states (native and unfolded) depends, aside from  $E$  also on temperature  $T$  and inversely on the friction  $\gamma$ , a factor accounting for the solvent damping of the rate constant. The theory resembles models, where chemical reactions are treated as a diffusive problem. Therefore, it appears to be a well suited formalism to describe reactions in solution.

$$k = \frac{\omega_0 \omega_{\ddagger}}{2\pi \cdot \gamma} \cdot \exp\left(-\frac{E}{RT}\right) \quad (1.12)$$

In this formalism,  $\omega_0$  and  $\omega_{\ddagger}$  describe the frequency of motion of the system in the starting well and on the barrier top, respectively. In the case where  $\gamma$  would be solely represented by the solvent viscosity  $\eta_{\text{solv}}$ ,  $k$  will be proportional to  $1/\eta_{\text{solv}}$ . Deviations from the inverse viscosity dependence have been encountered where a fractional dependence on the solvent viscosity can be observed.

$$k = k_0 \cdot \left(\frac{\eta}{\eta_0}\right)^{\beta} \quad (1.13)$$

In this case, the exponent  $\beta$  will deviate from -1 ( $\beta > -1$ ). In addition to a fractional viscosity dependence of  $k$ , Equation 4.11 predicts, that  $k$  would become infinite with  $\eta_{\text{solv}} \rightarrow 0$ . It was argued, that such a behavior is implausible, since dissipative interactions within the polypeptide chain would eventually take control of the dynamics<sup>77</sup>. Intrachain interactions or internal friction effects<sup>78</sup> eventually set an upper limit for  $k$ , leading to a deviation from the simple behavior  $k \propto \eta_{\text{solv}}^{-1}$  at low solvent viscosities.  $\tau$  ( $= 1/k$ ) was instead suggested to be expressed by the sum of the solvent-controlled relaxation time  $\tau_{\text{solv}}$  (which depends on solvent viscosity

## 1. Introduction

---

$\eta_{\text{solv}}$ ) and a solvent-independent timescale  $\tau_{\text{int}}$ .

$$\tau = \tau_{\text{int}} + \tau_{\text{solv}} \cdot \left( \frac{\eta}{\eta_0} \right)^{-\beta} \quad (1.14)$$

By measuring  $\tau$  at different solvent viscosities, it is possible to determine contributions from internal friction ( $\tau_{\text{int}}$ ) through extrapolation of  $\eta_{\text{solv}} \rightarrow 0$ . Hagen *et al.* used this strategy to determine  $\tau_{\text{int}}$  by setting  $\beta$  to -1. For Tryptophan Cage folding and unfolding, they found  $\tau_{\text{int}}$  to be  $0.7 \mu\text{s}$  and  $3 \mu\text{s}$ <sup>79</sup>. For the folding of cytochrome c  $\tau_{\text{int}} = 8 \mu\text{s}$  at  $20^\circ\text{C}$ <sup>80</sup>. In both cases, increasing temperature led to a decrease in  $\tau_{\text{int}}$  and  $\tau_{\text{solv}}$ . However in several reported cases,  $\tau_{\text{int}}$  was found to be zero when determined with a  $\beta$  of -1<sup>81-84</sup>. Kinetics in these cases were in the order of milliseconds to seconds, while the internal friction was found to be in the microsecond or nanosecond range<sup>77,79,80,85,86</sup>. The determination of  $\tau_{\text{int}}$  from kinetics with time constants of milliseconds thus comes with a large error. To determine the internal friction time constant more accurate, fast reactions need to be investigated with time constants close to the internal friction time constant. Therefore, techniques are necessary, which are able to monitor kinetics with a high time resolution.

To date remains unclear, whether a  $\beta$ -value of -1 should be used when determining  $\tau_{\text{int}}$  according to Equation 4.12 or if  $\beta$  might deviate from -1 resulting in a power law for determining  $\tau_{\text{int}}$ .

## 1.5. Protein Stability

Most small single-domain proteins exhibit a thermodynamically cooperative transition between U and N, which can be described by two state behavior with only the fully folded or unfolded state populated.



This does not mean that intermediate states must be absent. They are rather energetically unstable compared to U or N and only transiently populated. The unfolded state comprises many thermodynamic sub-states (see Chapter 1.3). Since these different structural forms of U are in rapid equilibrium and their interchange is much faster than the folding/unfolding reaction, it is valid to group these sub-states into a single thermodynamic state. The same holds true for N with the difference, that the ensemble of different structural sub-states is much narrower. The thermodynamic stability of the native state relative to the unfolded state



is given by

$$\Delta G^0 = -RT \ln K_f, \quad K_f = \frac{[N]}{[U]} \quad (1.16a,b)$$

with the equilibrium constant  $K_f$  and is speculated to be the state where the Gibbs free energy of the system is minimal<sup>2</sup>.  $\Delta G^0$  values for protein stability are usually small due to compensating enthalpic and entropic contributions of the protein-solvent system with typical values of  $-15$  to  $-60 \frac{kJ}{mol}$ <sup>55,87</sup>. As a result of the low stability of the native state, a small fraction of unfolded molecules exists even under optimal native conditions and protein molecules unfold and refold many times during their lifetime, even at equilibrium.

The two state model fails to explain the folding of large, multi-domain or ligand-binding proteins, since they often fold through intermediates or depend on additional factors such as peptidyl-prolyl isomerases<sup>7</sup>, chaperones<sup>88</sup>, disulfide isomerases<sup>6</sup> or ligands. Calcium ( $Ca^{2+}$ ) binding proteins like  $\alpha$ -lactalbumin or  $\beta$ -parvalbumin form partially folded intermediate structures in the absence of  $Ca^{2+}$ <sup>3,89</sup>. Whether two state behavior is observed can be tested by differential scanning calorimetry<sup>87</sup>. The calculated van't Hoff enthalpy change based on the two state assumption will only be equivalent to the calorimetrically determined enthalpy change if the two state criterion is fulfilled.

From the Gibbs fundamental equation it can be seen that protein stability is susceptible to perturbations which change the difference in volume  $\Delta V^0$ , entropy  $\Delta S^0$  or chemical potential  $\Delta \mu_i^0$ .

$$d\Delta G^0 = \Delta V^0 dp - \Delta S^0 dT + \sum_i \Delta \mu_i^0 dn_i \quad (1.17)$$

This can be achieved through modulation of pressure, temperature or solvent composition. Modulating the latter is possible through the addition of co-solutes, which alter the chemical potential.

## 1.6. The Role of Osmolytes and other Co-Solutes in Protein Stability

The question how different co-solutes act on protein stability has a long history. In the 19th century, Hofmeister studied the amount of various salts needed to precipitate proteins from whole egg white<sup>90</sup>. The series of anions and cations bearing his name could later be associated with changes in the protein stability. However, the molecular mechanism by which Hofmeister

## 1. Introduction

---

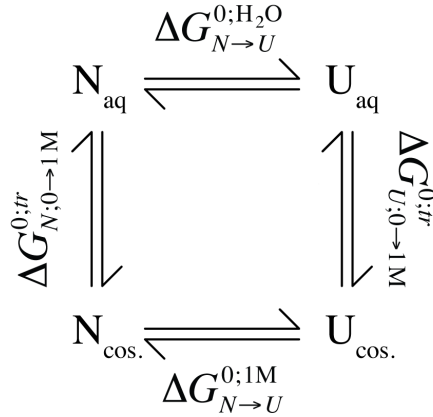
ions stabilize or destabilize the native state of proteins to date remains controversial<sup>91</sup>. In addition to ions, small organic co-solutes were found in various organisms, with the ability to regulate protein stabilization and destabilization in order to protect the proteins in the cell from environmental stress conditions<sup>92</sup>. These substances called osmolytes, together with other co-solutes influence the stability of proteins by modulating the solvent quality, a concept already used in polymer science<sup>42</sup>. They are capable of turning the solvent either into a good or a poor solvent for the polypeptide chain. In a good solvent, the polypeptide chain mainly interacts with solvent molecules, while in a poor solvent it mainly interacts with itself<sup>93,94</sup>. Therefore it is possible to tune the strength of intramolecular interactions by modulating the solvent quality. Denaturing co-solutes like urea and guanidinium chloride (GdmCl) render a good solvent for proteins and lead to unfolding due to interaction of protein surface with the solvent<sup>39,40</sup> or direct binding of denaturant molecules to the polypeptide<sup>95,96</sup>. Protecting osmolytes like sarcosine and trimethylamine *N*-oxide (TMAO) render a poor solvent for the polypeptide chain, thereby stabilizing the native state relative to the unfolded state. Different models exist to describe the effect of co-solutes and osmolytes on proteins. In the following, the two models believed to describe the effect with high accuracy are described in detail.

### 1.6.1. The Tanford Transfer Model

The Transfer Model developed by Tanford<sup>40,97</sup> describes the effect of osmolytes and other co-solutes on protein stability. It allows for the calculation of the free energy associated with the transfer either of a protein in the folded or unfolded state from water to a solvent containing a certain type of osmolyte or co-solute and gives insight into the interaction of protein groups with solute molecules. Transfer energies can be determined by summing the free energy contributions from hydrophilic and hydrophobic groups exposed to the solvent in the native and denatured state of a protein, assuming that the energetic contributions of the groups are additive. The transfer free energy for a protein from 0 to 1 M co-solute can be calculated by Equation 1.18a,b for the denatured state (a) or the native state (b) for all groups  $i$ ,

$$\Delta G_{U;0 \rightarrow 1M}^{0;tr} = \sum_i \alpha_{i;U} \Delta g_{i;0 \rightarrow 1M}^{0;tr} \quad , \quad \Delta G_{N;0 \rightarrow 1M}^{0;tr} = \sum_i \alpha_{i;N} \Delta g_{i;0 \rightarrow 1M}^{0;tr} \quad , \quad (1.18a,b)$$

where  $\alpha_i$  is the solvent accessibility of each group of type  $i$  and  $\Delta g_{i;0 \rightarrow 1M}^{0;tr}$  is the group transfer free energy (GTFE) for each individual group. Tanford combined these calculated free energies to a thermodynamic cycle depicted in Figure 1.3, where the following relationship



**Figure 1.3.:** Thermodynamic cycle of the Transfer Model. Horizontal equilibria represent unfolding in water (top) or 1 M of co-solute (bottom). Vertical equilibria represent transfer of either the native (left) or unfolded state (right) from water (aq) to 1 M of co-solute (Cos).

holds:

$$\Delta G_{N \rightarrow U}^{0;1M} - \Delta G_{N \rightarrow U}^{0;H_2O} = \Delta G_{U;0 \rightarrow 1M}^{0;tr} - \Delta G_{N;0 \rightarrow 1M}^{0;tr} \quad (1.19)$$

The difference in stability between the native state in 1 M co-solute ( $\Delta G_{N \rightarrow U}^{0;1M}$ ) and water ( $\Delta G_{N \rightarrow U}^{0;H_2O}$ ) is equal to the difference between the transfer energy of the denatured state ( $\Delta G_{U;0 \rightarrow 1M}^{0;tr}$ ) and the native state ( $\Delta G_{N;0 \rightarrow 1M}^{0;tr}$ ) from water to 1 M osmolyte, respectively. In theory, the difference between  $\Delta G_{N \rightarrow U}^{0;1M}$  and  $\Delta G_{N \rightarrow U}^{0;H_2O}$  or between  $\Delta G_{U;0 \rightarrow 1M}^{0;tr}$  and  $\Delta G_{N;0 \rightarrow 1M}^{0;tr}$  is equal to the protein folding  $m_{eq}$ -value, which can be determined experimentally by the linear extrapolation method<sup>98–100</sup>,

$$m_{eq} = \frac{\partial \Delta G_{N \rightarrow U}^{0;[Cos]}}{\partial [Cos]} \quad (1.20)$$

where [Cos] is the co-solute concentration. The comparison of the experimentally determined  $m_{eq}$ -value and the difference between  $\Delta G_{U;0 \rightarrow 1M}^{0;tr}$  and  $\Delta G_{N;0 \rightarrow 1M}^{0;tr}$  provides a way to test the validity of the thermodynamic cycle and the additivity of the GTFE's.

Bolen *et al.*<sup>101</sup> determined the difference between  $\Delta G_{U;0 \rightarrow 1M}^{0;tr}$  and  $\Delta G_{N;0 \rightarrow 1M}^{0;tr}$  according to Equation 1.18a,b. They calculated  $\alpha_N$  by summation of the surface exposed area of backbone units and amino acid side chains in the native state.  $\alpha_U$  was calculated by a similar summation, except that the exposed surface area was based on denatured state models. Two extreme models were used to represent the denatured state ensemble<sup>102</sup>. In the first model, the denatured state is represented by a fully extended chain in a good solvent. In the second model, the denatured state is represented by a collapsed chain in a poor solvent. The mean denatured

state accessibility of these two models was used for the determination of  $\alpha_U$ .

Predicted  $m_{eq}$ -values were found to be identical with experimental  $m_{eq}$ -values for induced folding or unfolding<sup>101,103</sup>, thereby confirming the validity of the thermodynamic cycle and the additivity of GTFE's. It is thus possible to predict solvent dependent cooperative folding and unfolding free-energy changes on the basis of GTFE's from water to various osmolytes.

### 1.6.2. Schellman's Weak Binding Model

In the weak binding model derived by Schellman, co-solute molecules are thought to bind to sites along the polypeptide chain<sup>95,104-106</sup>. The number of binding sites in U is larger than in N due to a higher solvent accessible surface area (SASA) of the former. The difference in the number of binding sites between U and N was found empirically to be proportional to the difference in the SASA<sup>99</sup>. According to the classical binding theory<sup>107</sup>, the free energy for binding of a co-solute molecule to equivalent binding sites is given by

$$\Delta G^0 = \Delta G^0(H_2O) - \Delta n \cdot RT \ln(1 + K_b[Cos]) \quad (1.21)$$

with  $\Delta n$  being the difference in the number of binding sites between U and N, and  $K_b$  the binding constant. For simplicity, all binding sites are assumed to be identical and binding to the different sites is independent. Equation 1.21 works well for strong binding ligands but cannot be used in the case of denaturants or osmolytes, where high co-solute concentrations (several molar) are necessary in order to exert their action<sup>39,40,94,108</sup>. A non-zero free energy associated with binding would be found in the case of weakly interacting co-solute molecules with proteins, even if the occupancy of co-solute molecules on the protein surface is identical to that in bulk solvent. This thermodynamic inconsistency was clearly pointed out by Schellman, which resolved this paradox by developing a "site exchange" formalism to describe the weak interaction of co-solutes with proteins<sup>95,109-111</sup>.

$$\Delta G^0 = \Delta G^0(H_2O) + \Delta n \cdot RT \ln(1 + (K_{ex} - 1)X_{Cos}) \quad (1.22)$$

Equation 1.22 is usually written as a function of the co-solute mole fraction ( $X_{Cos}$ ) rather than the molar concentration.  $K_{ex}$  is a dimensionless equilibrium constant for the exchange of a water molecule with a co-solute molecule at a binding site. When water and co-solute interact with an equal free energy with the protein,  $K_{ex}$  becomes unity and the intrinsic free energy associated with the exchange of a water molecule for a co-solvent molecule vanishes. In the

case where  $K_{\text{ex}} \gg 1$ , then  $K_{\text{ex}} - 1 \approx K_{\text{ex}}$  and Equation 1.22 becomes similar to the ordinary binding polynomial (Equation 1.21).

The findings of the Tanford Transfer Model and the weak binding model from Schellman do not contradict each other. Both models are able to describe the thermodynamics of folding under varying co-solute compositions.

## 1.7. Loop Formation in Polypeptide Chains

In order to understand the dynamic behavior in proteins, it is essential to measure intrachain diffusion in peptides, protein fragments or full-length proteins and compare obtained absolute rate constants with scaling laws from polymer theory. A variety of experimental systems addressed intrachain diffusion by measuring loop formation through contact formation between specific sites in the polypeptide chain. Haas *et al.* used time-resolved Förster resonance energy transfer (FRET) between two fluorophores, coupled to the polypeptide chain to measure dynamics in the unfolded state<sup>112</sup>. Difficulties arose for recovering both distance distributions and segmental diffusion due to strong correlation between the fitting parameters. However, using two different donor-acceptor pairs, distance distributions and diffusion constants could be obtained from which rate constants for loop formation were calculated<sup>36</sup>.

Intrachain loop formation was further studied by contact formation between specific points along the polypeptide chain. Hagen *et al.* estimated the intrachain diffusion rate by measuring the rate for an intramolecular bond formation in denatured cytochrome *c*<sup>85</sup>. A methionine reacts with a heme group, separated by approximately 50 amino acids, yielding time constants for contact formation around 35-40  $\mu\text{s}$ . However, the reaction of the methionine with the heme group is not diffusion controlled, which results in slower apparent rate constants compared to rate constants measured with diffusion-controlled systems. Loop formation was studied further by intramolecular triplet-triplet energy transfer from either thioxanthone or xanthone to naphthalene<sup>33</sup>. The results from these measurements are described in Chapter 1.7.1. Lapidus *et al.*<sup>113</sup> measured loop formation by triplet quenching of the tryptophan triplet state upon contact formation with a cysteine. The triplet state was populated selectively by a laserflash and the triplet decay was monitored by absorption. Since the excitation of the tryptophan into the triplet state is slow<sup>114</sup> and quenching does not occur upon every encounter<sup>115</sup>, additional information is required to extract absolute rate constants. Intrachain loop formation was measured in unfolded cytochrome *c* by electron transfer from the excited Zn-porphyrine group in

## 1. Introduction

---

the triplet state to a Ru-complex attached to a specific histidine<sup>116</sup>. The electron transfer rate constant was shown to be close to the diffusion limit allowing the determination of absolute rate constants for chain diffusion. A time constant of 250 ns was obtained for the formation of a 15-residue loop<sup>116</sup>.

In addition to triplet transfer and triplet quenching studies, fluorescence quenching has been used to measure loop formation. Intrachain loop formation was monitored by fluorescence contact quenching of 2,3-diacabicyclo[2.2.2]oct-2-ene (DBO) by tryptophan<sup>117,118</sup>. The fluorescence quenching rate was found to be close to the diffusion limit. However, due to the limited donor lifetime of  $\leq 1 \mu\text{s}$ <sup>117</sup> this method is restricted to relatively short peptides.

Intrachain loop formation was measured on the single molecule level, where tryptophan quenched the fluorescent state of an oxazine derivative (MR121)<sup>119</sup>. This led to fluorescence fluctuations, which could be followed by confocal fluorescence microscopy. However, rate constants obtained from this method have been shown to be significantly lower compared to those determined by more direct methods<sup>120</sup>, indicating that fluorescence quenching is not solely diffusion controlled.

Although some of the methods presented appear not to be diffusion controlled, the aforementioned studies reveal the importance of methods, which are able to measure intrachain diffusion. In order to obtain absolute rate constants for loop formation, model free methods are necessary.

### 1.7.1. Triplet-Triplet Energy Transfer to Study Loop Formation

A method perfectly suited to study intramolecular loop formation is triplet-triplet energy transfer (TTET)<sup>33,34,120</sup>. The triplet state of a donor (D) group is transferred to an acceptor (A) group upon contact formation with D ending up in the ground state. This process can be described by



where  $^*(T_1)$  describes the first electronically excited triplet state and  $(S_0)$  the singlet ground state. Dexter reported a quantum mechanical treatment for TTET where he described the energy transfer as a simultaneous exchange of two electrons<sup>121</sup>. He proposed the rate constant for energy transfer ( $k_{ET}$ ) to be strongly distance dependent, where the orbitals of D and A

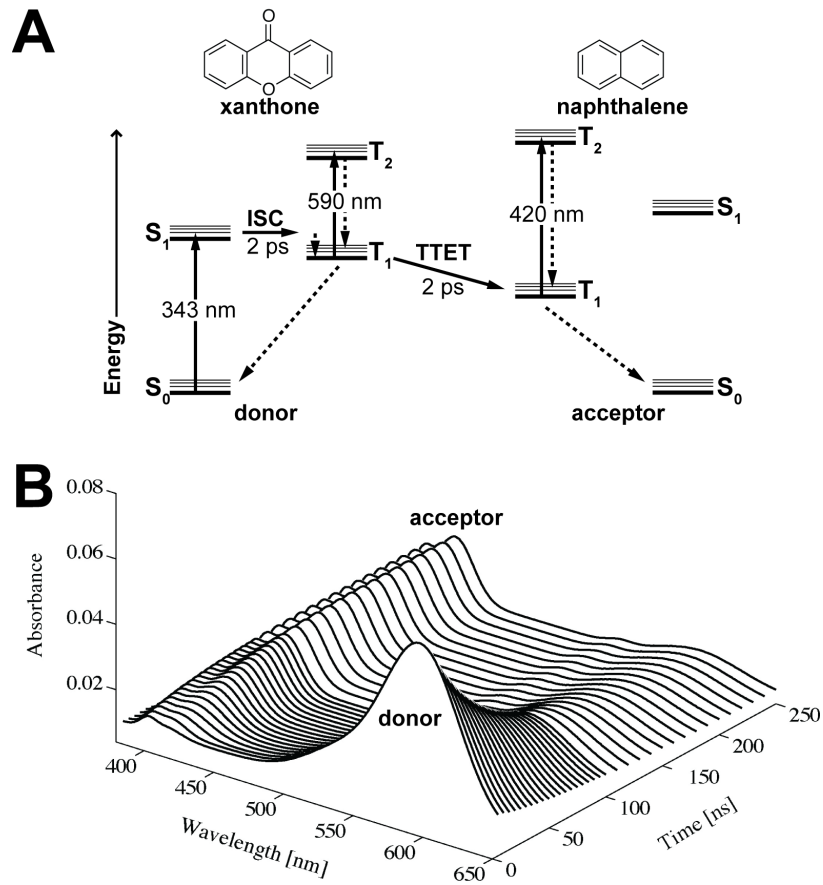
need to overlap in order to allow electron exchange.

$$k_{\text{ET}} \propto \exp\left(-\frac{2R_{\text{DA}}}{L}\right) \quad (1.24)$$

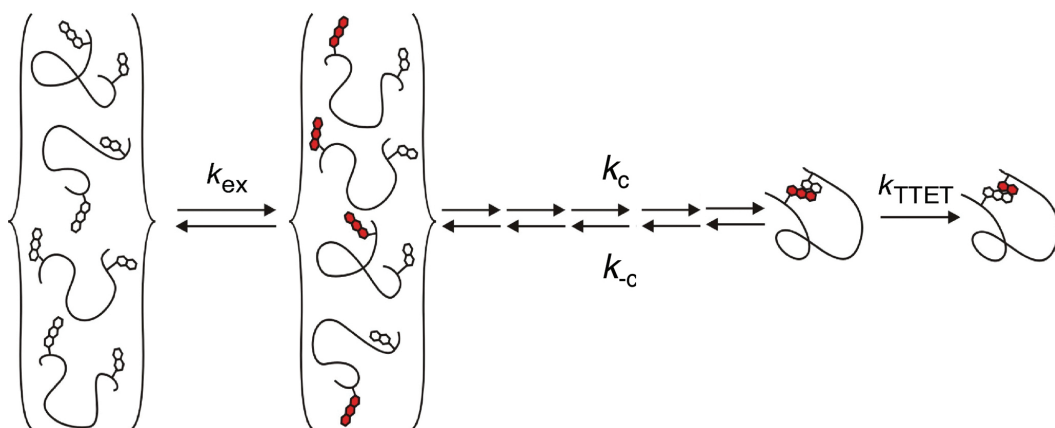
The critical electron transfer distance  $R$  between  $D$  and  $A$  thus becomes essentially equal to the sum of their van-der-Waals radii,  $L$ <sup>121</sup>. Energy transfer from  $D$  to  $A$  is not possible through a coulombic (Förster-type) mechanism due to an extremely small oscillator strength for singlet-triplet absorption by  $A$ . TTET thus differs from FRET where energy transfer occurs via dipole-dipole interactions and displays a weak distance dependence with  $k_{\text{ET}} \propto 1/r^6$ .

Suitable pairs of chromophores for TTET measurements need to fulfill several requirements. Intersystem crossing from the excited singlet state to the excited triplet state in the donor needs to be fast and virtually irreversible with a high quantum yield. Irreversibility in ISC is normally given due to an increase in entropy caused by the higher density of states in the lower-energy electronic state<sup>122</sup>. The triplet state of the donor should be long-lived, giving rise to a large experimental time window. Relaxation from the triplet state to the ground-state is quantum mechanically forbidden (unless magnetic interactions are included) as it is associated with a spin flip which changes the multiplicity of the system. Triplet state relaxation therefore is usually slow. Further, the triplet energy of the donor should exceed the triplet energy of the acceptor in order to allow efficient and irreversible energy transfer. It is also necessary to be able to exclusively excite the donor molecule. This can be achieved by selecting a pair of molecules which absorb at different wavelengths.

Xanthone and naphthalene have been shown to be a suitable donor-acceptor pair in TTET measurements fulfilling the requirements listed above. Triplet states of xanthone are formed within  $\sim 2$  ps with a high quantum yield of  $\sim 99\%$ <sup>123,124</sup>. Energy transfer was shown to be diffusion controlled with a time constant for triplet transfer of  $\sim 2$  ps with the chromophores in close contact<sup>34,123</sup>. In well-degassed solutions, the intrinsic donor lifetime exceeds  $10 \mu\text{s}$ <sup>125</sup>. The energy transfer can be directly followed by time-resolved absorbance measurements at 590 nm and 420 nm, where xanthone and naphthalene have strong triplet absorption bands, respectively (Figure 1.4 (A) and (B)). When xanthone and naphthalene are attached to a polypeptide chain, TTET gives information about chain dynamics (Figure 1.5). Loop formation brings the chromophores into proximity with energy transfer happening when they form van-der-Waals contact. This mechanism called the through-space mechanism is the single possibility for energy transfer between  $D$  and  $A$ , where a through-bond mechanism was found to be inexistent in bichromophoric systems with  $\geq 5$  C-C bonds between the labels<sup>126</sup>. Since  $k_{\text{ET}} \gg k_{\text{c}}, k_{-\text{c}}$



**Figure 1.4.:** Triplet-triplet energy transfer between the donor xanthone and the acceptor naphthalene. (A) Jablonski diagram. (B) Time-based change in the absorption spectra upon excitation by a short laser pulse. Figure B taken from<sup>34</sup>.



**Figure 1.5.:** Scheme for intramolecular triplet-triplet energy transfer in an ensemble of unstructured peptides to monitor site-specific interactions. Figure adapted from<sup>34</sup>.



(Figure 1.5), absolute rate constants for intrachain loop formation can be directly obtained from 10 ps to several  $\mu\text{s}$ .

Polypeptide sequences from naturally occurring proteins commonly consist of a balanced set of amino acids with a wide variety in size and chemical properties amongst their side chains. In order to measure TTET in natural sequences, the influence of the canonical amino acids on the triplet state of xanthone was tested. Several amino acids interact with the triplet state of xanthone by triplet quenching or TTET. Methionine and Histidine showed to efficiently quench the xanthone triplet state with a bimolecular rate constant close to the diffusion limit. For a complete list of amino acids affecting the xanthone triplet state, see Table 1.1. To perform

**Table 1.1.:** List of amino acids interacting with the triplet state of xanthone. Table adapted from<sup>32</sup>.

Amino acid	$k_q$ ( $\text{M}^{-1}\cdot\text{s}^{-1}$ ) <sup>a</sup>	Conditions	Interaction type
Naphthyl acetic acid	$(4.0 \pm 0.1)\cdot 10^9$	Water, pH 7	TTET
Trp	$(3.0 \pm 0.1)\cdot 10^9$	Water, pH 7	TTET <sup>b</sup>
NAla	$(2.8 \pm 0.1)\cdot 10^9$	Water, pH 7	TTET
Tyr	$(2.5 \pm 0.1)\cdot 10^9$	Water, pH 7	TTET <sup>b</sup>
Met	$(2.0 \pm 0.1)\cdot 10^9$	Water, pH 7	Triplet quenching
His	$(1.8 \pm 0.1)\cdot 10^9$	0.1 M KP, pH 8	Triplet quenching <sup>b</sup>
His <sup>⊕</sup>	$(2.8 \pm 0.1)\cdot 10^7$	0.1 M NaOAc, pH 4	Triplet quenching
Cys	$(5.1 \pm 0.2)\cdot 10^7$	Water, pH 7	Triplet quenching
N-terminal NH <sub>3</sub> <sup>⊕</sup>	$(2.0 \pm 0.5)\cdot 10^6$	Water, pH 7	Triplet quenching

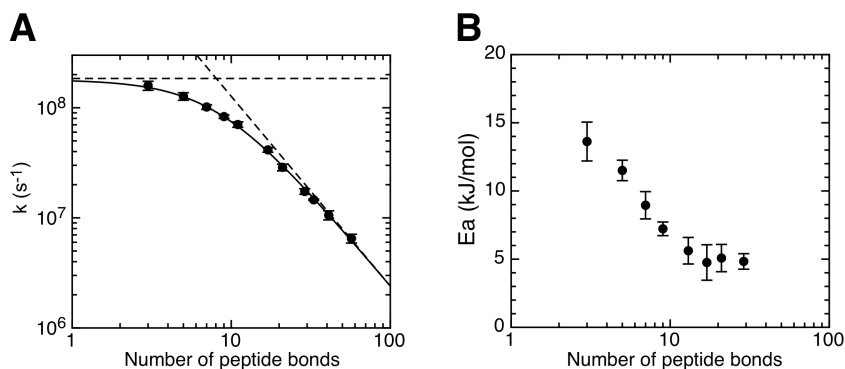
<sup>a</sup>Bimolecular quenching constants were measured under pseudo-first-order conditions.

<sup>b</sup>Radical formation as side reaction.

TTET in proteins, it is thus important to find suitable models, where amino acids interfering with the energy transfer reaction from xanthone to naphthalene are absent.

### 1.7.2. Dynamics of Intramolecular Loop Formation in Polypeptide Chains

Measuring intramolecular loop formation kinetics in peptides is important to get information about the flexibility and the dynamics in polypeptide chains. The kinetics for the chain to form a loop were measured in Glycine-Serine (Gly-Ser) peptides of different length with 1-28 repeats. The length dependent loop formation in the poly(Gly-Ser) peptides could be described



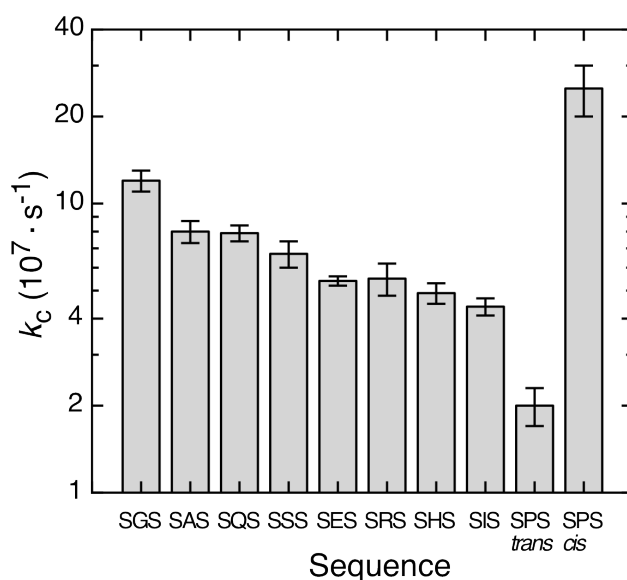
**Figure 1.6.:** Length dependence of the rate constant (A) and the activation energy (B) for loop formation in poly(Gly-Ser) peptides. Data from Krieger *et al.*<sup>34,127</sup>.

by Equation 1.25,

$$k_c = \frac{1}{1/k_0 + 1/(k_i \cdot N^\gamma)} \quad (1.25)$$

where  $N$  is the number of interchromophore peptide bonds. For loops with  $N > 20$ , the loop formation rate constant scales with  $N^{-1.7 \pm 0.1}$ , where  $k_i = (6.7 \pm 1.6) \cdot 10^9 \text{ s}^{-1}$ <sup>34</sup>. This exponent is in good agreement with the model for end-to-end diffusion, when excluded volume effects are considered<sup>120,127</sup>. For  $N < 20$ , the linear scaling law does not hold. In very short peptides, a limiting value for loop formation of  $k_0 = (1.8 \pm 0.2) \cdot 10^8 \text{ s}^{-1}$  is reached, corresponding to a time constant of 5.6 ns. This behavior can be explained by different processes, limiting the peptide chain motions over long and short segments. Chain stiffness is expected to dominate the dynamics in short chains. This view is supported by differences in enthalpic barrier height for loop formation in Gly-Ser chains of different length<sup>34</sup>. While long loops with  $N > 17$  encounter very low activation barriers ( $\sim 4 \text{ kJ/mol}$ ), short loops experience significant barriers ( $\sim 14 \text{ kJ/mol}$  for  $N = 3$ , see Figure 1.6 B). Polyserine chains are stiffer due to their lack of flexible glycylic residues. Loop formation is almost independent of loop size in short loops ( $N < 5$ ), where  $k_0 = (8.7 \pm 0.8) \cdot 10^7 \text{ s}^{-1}$  and two to three times slower than in poly(Gly-Ser) peptides. The regime where the rate constant for loop formation scales linearly with  $N$  in poly-serine peptides is only reached for longer chains compared to poly(Gly-Ser), which is a sign for decreased flexibility of the chain. This is in good agreement with the characteristic ratio, which is larger for polyserine than for poly(Gly-Ser) chains<sup>120</sup>. A slightly stronger effect of increasing loop size on the loop formation rate constant was found for poly-serine chains compared to poly(Gly-Ser), with  $\gamma = 2.1 \pm 0.3$  and  $k_i = (1.0 \pm 0.8) \cdot 10^{10} \text{ s}^{-1}$ <sup>34</sup>. Denaturants have been shown to affect loop formation kinetics in poly(Gly-Ser) peptides. In the presence of

8 M GdmCl,  $k_c$  scales with the peptide length already for shorter chains, with  $\gamma = -1.8 \pm 0.1$ . Polymer theory states, that chain dynamics are subject to the amino acid composition of the polypeptide. The characteristic ratio can be used to estimate the flexibility of the different amino acids. Glycine with the lowest  $C_\infty$  amongst all canonical amino acids thus appears to be the most flexible residue, whereas proline with the highest  $C_\infty$  appears stiff<sup>128,129</sup>. To verify the effect of different amino acids on chain dynamics, a series of short host-guest peptides was investigated by TTET with the uniform structure Xan-Ser-Xaa-Ser-Nal-Ser-Gly, where Xaa denotes the variable guest amino acid (Xaa = Gly, Ala, Gln, Ser, Glu, Arg, His, Ile, Pro)<sup>34</sup> (Figure 1.7). The fastest loop formation was measured for the peptide with the flexible glycyl



**Figure 1.7.:** Effect of individual amino acids on the loop formation rate constant in host-guest peptides. Data taken from references<sup>34,130</sup>.

residue. The kinetics for the host peptide with proline yielded two rate constants, one for *cis*-proline and one for *trans*-proline, which showed the slowest loop formation rate constant. Amino acids with bulky side chains slightly slow down the dynamics, whereas charges have no influence. Although the amino acid composition influences chain dynamics, the effect of the different side chains is small<sup>34</sup>.

During protein folding, not only loops between terminal residues need to be formed but also between residues within the polypeptide chain. The influence of additional tails on loop formation was studied in peptides with extensions either on one or on both termini<sup>131</sup>. As predicted by polymer theory<sup>132</sup>, a limiting value for the effect of very long additional tails on loop for-

## 1. Introduction

---

mation kinetics was found. In this limit, loop formation is 2.5 times slower for chains with a single long extension. Formation of loops with two extensions are further slowed down by a factor of 1.7 compared to chains with a single extension. These findings suggest, that dynamics are faster at the chain ends due to higher flexibility compared to segments within the chain. Chain dynamics are strongly dependent on solvent viscosity<sup>35,133</sup>. A linear relation of  $\log k_c$  as a function of  $\log \eta$  was found (Equation 4.15). Long Gly-Ser chains ( $N \geq 15$ ) experience the full viscous effect with  $\beta = -1$  in the presence of viscosogens with approximately the size of the persistence length. Short peptides show  $\beta$ -values deviating from -1 and are not solely diffusion controlled.

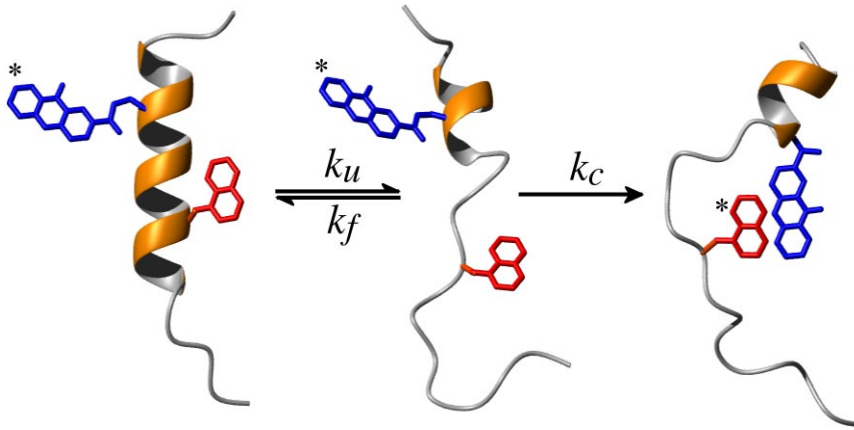
The effect of different sized co-solutes on chain dynamics was investigated for several model peptides<sup>130</sup>. The viscosity dependence can be described by Equation 4.15. An increase of the solvent viscosity slows down loop formation kinetics for all investigated co-solutes, but the magnitude of the effect depends on their size.  $|\beta|$  was found to decrease with increasing co-solute size even for long polypeptide chains. These findings show, that polypeptides do not experience the full macroscopic viscous effect in the presence of large viscosogens, which is in good agreement with large crowding agents having only a small effect on bimolecular diffusion processes<sup>134</sup>.

Denaturants have been shown to influence intrachain diffusion in unstructured peptides<sup>34,35</sup>. A linear change of  $\ln k_c$  with denaturant concentration  $[D]$  was observed,

$$\ln k_c = \ln k_c^{\text{H}_2\text{O}} - m_c \cdot \frac{[D]}{RT} \quad (1.26)$$

yielding an  $m_c$ -value ( $\partial \ln k_c / \partial [D]$ ) two times higher for GdmCl than for urea. In the absence of denaturant, the rate constant for loop formation is equal to  $\ln k_c^{\text{H}_2\text{O}}$ , the rate constant for loop formation in water. When correcting for the viscous effect,  $\ln k_c$  depends no longer linearly on  $[D]$ . The interaction of denaturant molecules with the peptide chain can then be described by the weak binding model from Schellman.

In addition to unstructured polypeptides, TTET through intramolecular loop formation was used to monitor dynamics in folded and partially folded structures, when linked to a folding/unfolding equilibrium<sup>125,135,136</sup>. The TTET labels were attached to the polypeptide chain, unable to form contact in the folded structure. Loop formation can thus only happen by local or global unfolding. Figure 1.8 shows the folding/unfolding equilibrium and the irreversible probing step in a folded structure. Two apparent rate constants were observed, which allowed to determine the microscopic rate constants for folding ( $k_f$ ), unfolding ( $k_u$ ) and loop forma-



**Figure 1.8.:** Probing local dynamics and stability in  $\alpha$ -helical peptides. Helix unfolding ( $k_u$ ) can be followed by refolding ( $k_f$ ) or by loop formation ( $k_c$ ). Figure taken from Fierz *et al.*<sup>135</sup>.

tion in the unfolded state ( $k_c$ ), using a three state model. To determine the microscopic rate constants more reliably, experiments were performed at different urea concentrations. A linear effect of urea was assumed on  $\ln k_f$  and  $\ln k_u$ .

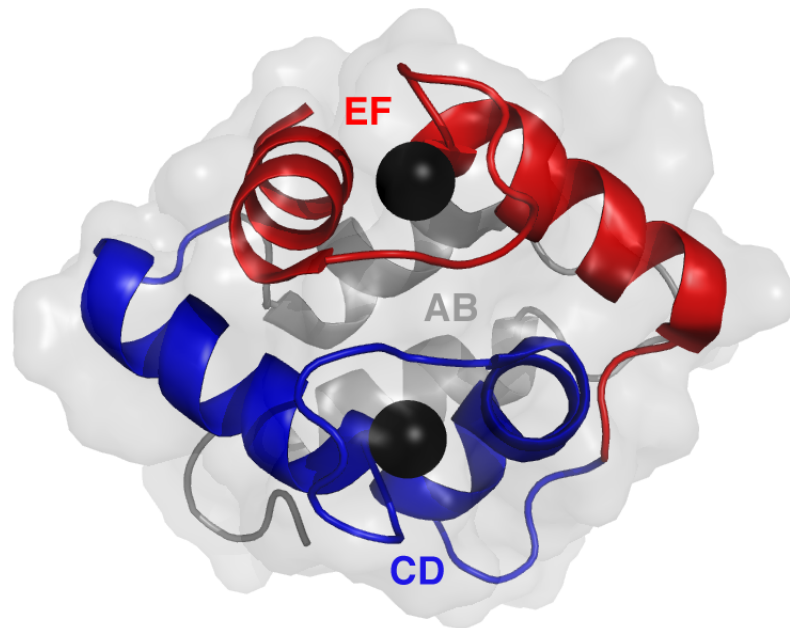
In helical peptides, loop formation occurs on the time scale of  $\sim 100$  ns, similar to loop formation in unstructured peptides. Local helix formation kinetics are position independent, revealing a time constant of  $\sim 400$  ns. Helix unfolding however is fast at the termini but slow in the center of the helix and varies between  $\sim 250$  ns and  $\sim 1.4 \mu\text{s}$ <sup>135</sup>.

Conformational fluctuations were investigated in the 35 amino acid long villin headpiece subdomain (HP35)<sup>125</sup>. Two native states with a slightly different conformation in the C-terminal region could be found. They exhibit different conformational dynamics and are connected by an unlocking/relocking equilibrium. Unlocking occurs with a time constant of  $\sim 1 \mu\text{s}$  at  $5^\circ\text{C}$ . The unlocked state exhibits a fast local conformational fluctuation with a time constant of 170 ns, which can be addressed to C-terminal helix unfolding and undocking from the core. Global unfolding occurs from the unlocked state on a much slower time scale. Global unfolding is only possible after an initial unlocking of the native state, leading to a more flexible native-like conformation.

Up to date, no full-length protein has been investigated by TTET. It has proven challenging to find a suitable model, where amino acids interfering with the triplet state of xanthone are absent. In addition, site specific labeling in full length protein is not straightforward.

## 1.8. Carp $\beta$ -Parvalbumin, as a Model to study Loop Formation by TTET in Different States of a Protein

Parvalbumins constitute a class of highly soluble all  $\alpha$ -type calcium binding proteins with low molecular weight. They exist in fish<sup>137</sup>, amphibians<sup>138</sup> and mammals<sup>139</sup> and mediate  $\text{Ca}^{2+}$ -exchange between the sarcoplasmic reticulum and myofibrils<sup>140</sup>. Mouse mutants deficient in parvalbumins exhibit tetanic contractures after muscular stimulation<sup>141</sup>.  $\beta$ -Parvalbumin (PV) from *cyprinus carpio* (common carp) can be found in white tissue of carp muscles and fast firing neurons, with intracellular concentrations in the milli-molar range<sup>142</sup>. PV has an acidic isoelectric point (pI) of 4.25 and an unbalanced composition of amino acids. Its primary sequence of 108 amino acids is rich in alanine, aspartic acid, glutamic acid and phenylalanine with a single histidine and cysteine. Methionine, proline, tyrosine and tryptophan are lacking completely<sup>143</sup>. Kretsinger and Nockolds resolved the crystal structure in 1973 to a resolution of 1.9 Å by the multiple isomorphous method<sup>144</sup>. This structure was refined by restrained least-squares analysis to 1.5 Å resolution<sup>145</sup>.



**Figure 1.9.:** Structure of carp  $\beta$ -parvalbumin with the helix-loop-helix motifs AB, CD and EF. Calcium ions are shown as black spheres. The figure was prepared using MacPyMOL and the protein databank (PDB) file 4CPV<sup>145,146</sup>.

The overall structure of PV consists of six helices, A through F. They are associated in three

helix-loop-helix motifs, AB, CD, and EF, giving rise to a spherical shape. Two calcium ions can be bound per protein molecule, one by the CD and EF-motif with dissociation constants of  $10^{-7}$ - $10^{-9}$  M for both  $\text{Ca}^{2+}$ -binding loops<sup>147</sup> (Figure 1.9). It was this EF-motif which served as eponym for the well-known EF-hand structural motif. Although the AB region does not bind calcium, it has a structure similar to the CD and EF regions, which show an approximate intramolecular two-fold axis and may have resulted from gene triplication with loss of calcium binding ability of the AB-loop<sup>144</sup>. Hydrophobic amino acids like phenylalanine, isoleucine, leucine and valine point to the center of the structure and form a well-defined hydrophobic core. Polar side chains, except those associated in  $\text{Ca}^{2+}$ -coordination reside on the surface and may interact with surrounding solvent.

The two  $\text{Ca}^{2+}$ -ions can be removed independently from the  $\text{Ca}^{2+}$ -binding loops of PV through chelation with ethylene glycol tetraacetic acid (EGTA), displaying the non-cooperativity of  $\text{Ca}^{2+}$ -binding. Removal of the first  $\text{Ca}^{2+}$  has no significant effect on the helical content while the removal of the second ion leads to reduced helical content<sup>89</sup>.

Folding of PV was expected to yield simple kinetics due to the absence of proline residues<sup>148</sup>. However, a transient folding intermediate accumulates in the dead-time of stopped-flow re-folding<sup>149</sup>. This fast pre-equilibrium between the unfolded and intermediate state has been shown to be a  $\text{Ca}^{2+}$ -dependent mechanism where the first ion is bound during the transition from U to I with the EF-hand  $\text{Ca}^{2+}$ -binding domain being formed first and rapidly<sup>150</sup>.

### 1.8.1. Loop Formation in Protein Fragments of Carp

#### $\beta$ -Parvalbumin

TTET was used to monitor dynamics in different fragments of PV<sup>32,127,130,151</sup>. Common to all studied PV fragments are two flanking phenylalanines, which form contact in the hydrophobic core of the native PV. The phenylalanines were replaced by xanthone and naphthylalanine during solid-phase peptide synthesis (SPPS).

Dynamics in the 18-amino-acid long, calcium binding EF helix-loop-helix motif (PV residues 85-102) were studied. Single exponential kinetics were observed in TTET-measurements, with the time constant for loop formation of 50 ns. Chain dynamics in the EF-loop are comparable to the dynamics in polyserine peptides with the same number of peptide bonds, although the former contains amino acids with larger side chains. These were shown to slow down intra-chain dynamics<sup>32</sup>. The high glycine content of the EF-loop might on the other hand increase chain dynamics, in summary resulting in chain dynamics comparable to those of polyser-

ine peptides of the same length. Replacing the charged amino acids aspartate and glutamate showed no effect on loop formation dynamics, neither did the addition even of high concentrations of calcium.

The 16 amino acid long DE-loop (PV residues 70-85) does not bind calcium but also connects two helices, D and E. Single exponential kinetics were observed with the time constant of 70 ns. Dynamics of the DE-loop are slower compared to those of polyserine peptides of identical length due to the presence of larger amino acid side chains. Where this effect was compensated in the EF-loop by its high glycine content, this is not the case in the DE-Loop, which consist only a single glycine residue.

A short fragment of 5 amino acids (PV residues 66-70) represents the N-terminal end of the D-helix. The position of the phenylalanines along the helix is  $i/i+4$ , with the side chains contributing to the hydrophobic core. Surprisingly, identical intrachain dynamics compared to the polyserine of the same length were found, although the PV fragment contains only amino acids with larger side chains and completely lacks Gly.

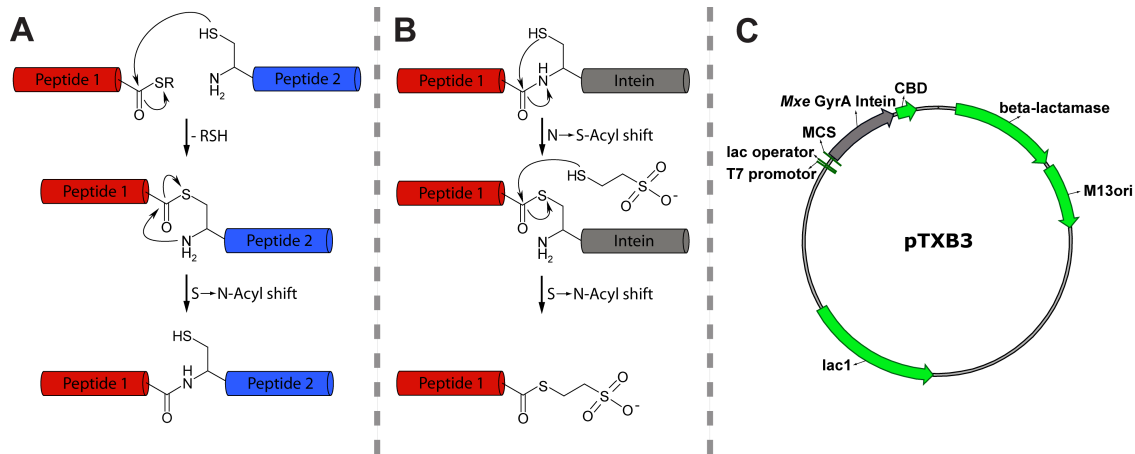
### 1.8.2. Introducing Non-Natural Amino Acids into Proteins by Native Chemical Ligation

To apply TTET to a full length protein, xanthone and naphthalene have to be attached at specific positions in the polypeptide chain. The chromophores are usually introduced site-specifically into peptides in a fully protected manner during SPPS. Despite the relatively small size of PV (108 amino acids), it exceeds the size of polypeptides accessible by SPPS. Recombinant protein expression is usually used for the generation of long polypeptides and proteins but fails incorporating non-canonical amino acids with a xanthone- and a naphthalene moiety directly. To circumvent this limitation, a condensation reaction known as native chemical ligation can be used to generate proteins bearing non-canonical amino acids or biophysical probes at defined positions. The condensation strategy known as native chemical ligation (NCL) was developed by Dawson *et al.*<sup>152</sup>, to circumvent the limitations of SPPS and recombinant protein expression. In NCL, two unprotected polypeptides, either synthesized or expressed, can be ligated together in aqueous condition at physiological pH, yielding a native peptide bond at the ligation site. The principle of NCL is shown in Figure 1.10 A.

The first step in NCL is the nucleophilic attack of the side chain of an N-terminal cysteine of peptide 2 on a C-terminal thioester of peptide 1. This reversible transthioesterification step is followed by an S→N-acyl transfer of the so-formed thioester. This results in the ligated



## 1.8. Loop Formation in Different States of Carp $\beta$ -Parvalbumin



**Figure 1.10.:** Native chemical ligation. (A) Principle, (B) intein mediated thioester generation, (C) commercially available expression vector for peptide-thioester formation.

product with a native peptide bond at the ligation site.

Despite the ease of NCL, generating the  $\alpha$ -carboxyl thioester has shown to be the bottleneck in NCL<sup>153</sup>. Different attempts have been made to obtain peptide thioesters using solid phase tert-butyloxycarbonyl (Boc) or 9-fluorenylmethoxycarbonyl (Fmoc) peptide synthesis<sup>152,154</sup>. NCL can be applied to protein semi-synthesis due to its compatibility with all naturally occurring amino acids. However, a strategy is needed for the preparation of recombinant polypeptide  $\alpha$ -carboxyl thioesters in case a synthetic peptide needs to be ligated to the C-terminus. The discovery of inteins provides a solution for the need of  $\alpha$ -carboxyl thioesters. Inteins can be seen as the protein counterpart of introns, self-excision modules, which ligate their flanking portions together in a process called protein splicing<sup>155,156</sup>. Modified inteins with a mutation of the crucial residue lose their ability to catalyze their *ipso* excision<sup>157,158</sup>. These inteins can only promote the first step of protein splicing, getting stalled after the formation of a C-terminal thioester at an N-terminally located protein (Figure 1.10 B). The intein-thioester can be substituted by a small, nucleophilic thiol, forming a thioester better suited for subsequent NCL. Modified inteins from e.g. *Mycobacterium xenopi*<sup>159,160</sup> (*Mxe GyrA*) cloned into special expression vectors are commercially available (see figure 1.10 C).

A successful application of NCL would for the first time allow to prepare full length proteins bearing the chromophores xanthone and naphthalene. Getting insights into dynamics in different states of proteins using TTET would become possible and eventually reveal how these dynamics are coupled to protein folding.



## 2. Aim of Research

The aim of this thesis is to compare loop formation dynamics in model peptides, in protein fragments and in a full-length protein, in order to determine the effect of amino acid side chains on the dynamics of polypeptide chains. Further, I wanted to study, how the nature of the solvent and solvent viscosity affects the dynamics of loop formation in model peptides and natural sequences.

**Effect of Osmolytes and other Co-solutes on Loop Formation:** The dimensions and dynamics of a polypeptide chain depend on the properties of the solvent. For polypeptides, a solvent can be good or poor, depending on the strength of peptide intramolecular interactions in comparison with peptide-solvent interactions. In good solvents, peptide-solvent interactions are more favorable than intramolecular interactions. In poor solvents, however, intramolecular interactions are more favorable than peptide-solvent interactions. Small organic compounds exist, which are able to modulate solvent quality. We planned to investigate the influence of stabilizing, destabilizing and neutral co-solutes on the dynamics and barriers of loop formation in the unfolded state of different polypeptide chains. By comparing the effect on poly(Gly-Ser) homopolymers and on natural sequences, we wanted to determine the contributions from side chains and intramolecular interactions on loop formation dynamics. Further, we wanted to test, whether the effect of the co-solutes on the polypeptide chain could be described by Tanford's Transfer Model or by the weak binding formalism proposed by Schellman.

**Internal Friction for Loop Formation:** Diffusive motions in polypeptide chains lead to the exploration of conformational space during protein folding. It was suggested, that intramolecular peptide motions are subject to friction evoked by the solvent and to internal friction within the polypeptide chain. The internal friction effect, expectedly caused by intramolecular interactions and chain stiffness, should be independent of solvent viscosity and would set an upper limit to the dynamics of loop formation. We wanted to investigate, whether internal friction exists for loop formation in unstructured polypeptide chains and which theoretical model is able to describe the contributions from internal friction. We planned to deter-

## 2. Aim of Research

---

mine the effect of amino acid composition, solvent quality and temperature on solvent friction and internal friction, by investigating poly(Gly-Ser) homopolymers and fragments from naturally occurring proteins in the presence of stabilizing, destabilizing and neutral co-solutes.

**Dynamics in Different States of Carp  $\beta$ -Parvalbumin:** Understanding the dynamics in different states of a protein is important to understand its free energy landscape. Little is known about the dynamics in unfolded and in intermediate states of full-length proteins. The introduction of triplet labels into a natural occurring protein would enable us to characterize the dynamics in different states. We chose to study carp  $\beta$ -parvalbumin (PV) by TTET, since it is folded in the presence of calcium and forms an equilibrium intermediate (I) in the absence of calcium. Amino acids interfering with the energy transfer reaction are absent in PV. We wanted to attach the chromophores selectively at the termini of the E-helix in the context of full-length PV, using native chemical ligation to couple a synthesized C-terminal fragment bearing the triplet donor to an expressed N-terminal fragment. The triplet acceptor was planned to be attached using a thiol reactive naphthyl derivative. With this double-labeled full-length PV, we wanted to measure dynamics in U and I. We further wanted to investigate, how long range interactions affect dynamics in a protein in different states by comparing the result with those obtained from identically labeled isolated protein fragments.

## 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), VWR International, LLC (West Chester, PA, USA) or Honeywell Riedel de Haën (Seelze, Germany) if not stated otherwise. Guanidinium chloride (GdmCl) AA grade was purchased from Nigu Chemie GmbH (Waldkraiburg, Germany), urea ultrapure from Gerbu Biotechnik GmbH (Heidelberg, Germany). Fmoc-protected amino acids were purchased from Novabiochem (Hohenbrunn, Germany) or Iris Biotech (Marktredwitz, Germany).

### 3.2. Solid-Phase Peptide Synthesis

All peptides were synthesized using 9-fluorenylmethoxycarbonyl (Fmoc) solid-phase peptide synthesis (SPPS) on a ABI 433A peptide synthesizer from Applied Biosystems (Foster City, CA, USA). A slightly modified version with prolonged coupling times of the *FastMoc*<sup>TM</sup>-chemistry protocol on the 0.1 mmol scale in a 10 ml reaction vessel was applied. 1 mmol or 0.5 mmol of dry  $\alpha$ -amino Fmoc-protected amino acid was activated with *O*-(Benzotriazol-1-yl)-*N,N,N',N'*-tetramethyluroniumhexafluorophosphate (HBTU) or *N*-hydroxybenzotriazol (HOBT) / HBTU and *N,N*-diisopropylethylamine (DIPEA) in *N,N*-Dimethylformamide (DMF) forming an active ester. The N-terminus of the growing polypeptide was deprotected with 20% piperidine in *N*-methyl-pyrrolidone (NMP). Fmoc-deprotection efficiencies were controlled via UV feedback monitoring at 301 nm. Deprotection steps were repeated, until the absorbance was less than 5% of the initial value.

Key amino acids were coupled manually. 0.3 mmol of Fmoc protected amino acids were activated with benzotriazol-1-yl-oxytrypyrrolidinophosphonium hexafluorophosphate (PyBOP) in DMF and 130  $\mu$ l of 44% *N*-methylmorpholine (NMM) in DMF. The side chains remained protected during the whole process in automated and manual peptide synthesis.

### 3. Material and Methods

Solid supports (resins) were purchased from Rapp Polymere GmbH (Tübingen, Germany) for all syntheses according to Table 3.1.

**Table 3.1.:** Solid supports, double building blocks and activation reagents used for SPPS.

Peptide	Resin ( <i>TentaGel</i> )	double building blocks	Activation
poly(Gly-Ser)	Trt-Gly Fmoc S PHB-Gly Fmoc	Fmoc-Gly-Ser( $\psi^{\text{Me,Me}}$ )-OH	HBTU
PV <sup>70-85</sup>	R RAM	Fmoc-Asp(OtBu)-(Dmb)Gly-OH	HBTU
PV <sup>75-108</sup>	S RAM	Fmoc-Asp(OtBu)-(Dmb)Gly-OH Fmoc-Phe-Thr( $\psi^{\text{Me,Me}}$ )-OH Fmoc-Lys-Thr( $\psi^{\text{Me,Me}}$ )-OH	HBTU
PV <sup>78-108</sup>	S RAM R RAM R PHB-Ala Fmoc	Fmoc-Asp(OtBu)-(Dmb)Gly-OH Fmoc-Phe-Thr( $\psi^{\text{Me,Me}}$ )-OH Fmoc-Lys-Thr( $\psi^{\text{Me,Me}}$ )-OH	HBTU HATU
PV <sup>78-91</sup>	R RAM	Fmoc-Asp(OtBu)-(Dmb)Gly-OH Fmoc-Lys-Thr( $\psi^{\text{Me,Me}}$ )-OH	HBTU
PV <sup>79-91</sup>	S RAM	Fmoc-Lys-Thr( $\psi^{\text{Me,Me}}$ )-OH	HBTU
PV <sup>85-102</sup>	R RAM	Fmoc-Asp(OtBu)-(Dmb)Gly-OH	HBTU
GB1 Hairpin	R RAM	-	HBTU

Special procedures were applied for difficult sequences. Aspartimide formation for sequences containing an aspartate-glycine motif was prevented by the use of a 2,4-dimethoxybenzyl protected aspartate-glycine building block (Fmoc-Asp(OtBu)-(Dmb)Gly-OH) from Novabiochem (Hohenbrunn, Germany). Aggregation of the peptides during synthesis was prevented by using oxazolidine dipeptide building blocks from Novabiochem. The more efficient activating reagent 2-(7-Aza-1H-benzotriazole-1-yl)-1,1,3,3-tetramethyluronium hexafluorophosphate (HATU) was used for the coupling of difficult amino acids.

## 3.3. Peptide Modification

### 3.3.1. Synthesis and Introduction of Chromophores for Triplet-Triplet Energy Transfer

9-oxoxanthene-2-carboxylic acid (xanthonic acid) was synthesized as described earlier<sup>161</sup> and either coupled to the N-terminus or the selectively deprotected side chain of diaminopropionic acid (Dpr) while the rest of the side chains remained protected. The *N*-methyltrityl

(Mtt) protecting group was removed from Dpr by shaking the resin four times six minutes in 1% trifluoroacetic acid (TFA), 5% triethylsilane (TES) in dichloromethane (DCM), followed by thorough washing with DMF. A three fold excess (relative to the resin functionality) of xanthonic acid dissolved 1:1 with PyBOP in DMF was used for labeling. 5 equivalents N-Methylmorpholine (NMM) were added to the mixture right before transferring it to the peptide. After 30 minutes of reaction, the resin was washed with DMF and DCM. For incorporating the triplet acceptor naphthalene into polypeptides, the nonnatural Fmoc protected amino acid 1-(L)-naphthylalanine (BACHEM, Switzerland) was introduced during peptide synthesis. Alternatively, 2-bromo-*N*-(1-naphthyl)acetamide (BNAA) provided by A. Reiner<sup>162</sup> was used to attach a naphthyl moiety to the side chain of cysteine residues in an S<sub>N</sub>2 reaction. Peptides were labeled with BNAA as described in Chapter 3.6.6.2, except that labeling was performed in 100 mM Tris/HCl pH 8.48 in DMF (1:1). Labeling reactions were carried out for 40 minutes at room temperature. The success of the labeling reaction was tested with high-performance liquid chromatography (HPLC) and matrix-assisted laser desorption/ionization-time of flight mass spectrometry (MALDI-TOF).

#### 3.3.2. Acetylation of the Peptide Amino-Terminus

10% acetic anhydride, 10% DIPEA in DMF were added to the resin bound peptide and shaken for 10 minutes, followed by washing of the resin with DMF. The procedure was repeated three times and in a final step the peptide was washed with DCM.

#### 3.3.3. Peptide Cleavage

The peptide was simultaneously deprotected and released from the resin by a cleavage cocktail of 95% (v/v) TFA, 2.5% (v/v) TES and 2.5% (v/v) H<sub>2</sub>O for 2 hours while stirring under argon atmosphere. No additional precautions were necessary to avoid oxidation of cysteine containing peptides.

### 3.4. Peptide Purification

Peptides were precipitated from the cleavage cocktail in ~ 40 ml ice-cold methyl *tert*-butyl ether. The precipitate was pelleted by centrifugation and lyophilized. 20 mg of dried raw product was dissolved in 1 ml TFA, diluted with 5 ml of a 20% (v/v) acetonitrile (ACN)/water

### 3. Material and Methods

---

mixture and loaded onto a HPLC column. Small or hydrophobic peptides, which did not precipitate in ether, were directly loaded onto the column without a previous precipitation step. An ACN/water gradient in the presence of 0.1% TFA was applied to elute the peptides from the column. Peak fractions were collected and lyophilized. The purity of the fractions was approved by analytical HPLC. Fractions containing the desired product were identified by MALDI-TOF or electrospray ionization mass spectrometry (ESI-MS).

## 3.5. Molecular Biology Methods

### 3.5.1. Site-Directed Mutagenesis

The carp  $\beta$ -parvalbumin (PV) double mutant C18S/H26F was obtained by site-directed mutagenesis of the triple PV mutant C18S/H26F/F102W provided from Langheld<sup>150</sup>. Mutagenesis was performed using the QuickChange<sup>®</sup> Site-Directed Mutagenesis Kit from Stratagene (La Jolla, CA, USA). The mutagenic primer and wild-type sequences are given in Table 3.2 and were obtained from Microsynth AG (Balgach, Switzerland).

**Table 3.2.:** Mutagenic primer sequences used for QuickChange<sup>®</sup> site-directed mutagenesis.

PV variant	Direction	Primer sequence
PV C18S/H26F	5' → 3'	CGG TGT TGA CGA ATT TAC CGC TCT GGT TAA AGC ATG AGG
	3' → 5'	CCT CAT GCT TTA ACC AGA GCG GTA AAT TCG TCA ACA CCG

Polymerase chain reaction (PCR)<sup>163</sup> was performed on a Advanced Primus 25 PCR thermocycler from PEQLAB (Erlangen, Germany) with the temperature program given in Table 3.3.

Parental DNA strands were digested with *Dpn* I for 60 minutes at 37 °C. In the kit included supercompetent XL1-Blue *E. coli* cells (genotype: endA1 gyrA96(nal<sup>R</sup>) thi-1 recA1 relA1 lac glnV44 F' [::Tn10 proAB<sup>+</sup> lacI<sup>q</sup> Δ(lacZ)M15] hsdR17(r<sub>K</sub><sup>-</sup> m<sub>K</sub><sup>+</sup>)) were transformed (see Chapter 3.5.3) with the *Dpn* I treated PCR product and incubated on 0.1 mg/ml ampicillin (Amp) containing lysogeny broth (LB) agar plates overnight at 37 °C. Overnight cultures, inoculated from freshly grown colonies, were prepared in LB and 0.1 mg/ml Amp to amplify the



**Table 3.3.:** PCR Temperature program

Temperature	Time	Cycles
95 °C	1 min	1x
95 °C	1 min	
58 °C	1 min	18x
72 °C	6 min	
4 °C	24 h	1x

mutated plasmid. A QIAprep Spin MiniPrep Kit from QIAGEN (Hilden, Germany) was used to extract and purify the plasmids. Plasmids were sequenced by GATC Biotech (Konstanz, Germany), in order to confirm the success of the site-directed mutagenesis.

### 3.5.2. Preparation of Chemically Competent Cells

Chemically competent cells were prepared after a modified protocol from Cohen *et al.*<sup>164</sup>. 2 ml of a BL21 (DE3) *E. coli* (genotype: F<sup>-</sup> ompT gal dcm lon hsdS<sub>B</sub>(r<sub>B</sub><sup>-</sup> m<sub>B</sub><sup>-</sup>) λ(DE3 [lacI lacUV5-T7 gene 1 ind1 sam7 nin5])) overnight culture were diluted into 100 ml dYT-medium and grown at 30 °C to an OD<sub>600</sub> of 1. Cells were incubated for 15 minutes on ice and pelleted by centrifugation at 4 °C. After resuspending and incubating the cells in 40 ml ice-cold 100 mM CaCl<sub>2</sub>-solution for 30 minutes, they were pelleted again and resuspended in 5 ml ice-cold 80 mM CaCl<sub>2</sub>, 20% glycerol (v/v). Aliquots of 200 μl were frozen in liquid nitrogen and stored at -80 °C.

### 3.5.3. Transformation of Chemically Competent Cells

200 μl of chemically competent BL12 (DE3) or XL1 blue *E. coli* cells were thawed on ice and mixed with an appropriate amount of vector, followed by an incubation on ice for 10 minutes. After a 2 minutes heat shock at 42 °C, ~200 μl dYT- or LB-medium were added immediately and the cells were incubated for 30 minutes at 37 °C. The cell suspension was transferred to 2-3 agar plates containing 0.1 mg/ml Amp. The plates were incubated overnight at 37 °C.

## 3.6. Protein Expression, Purification and Modification

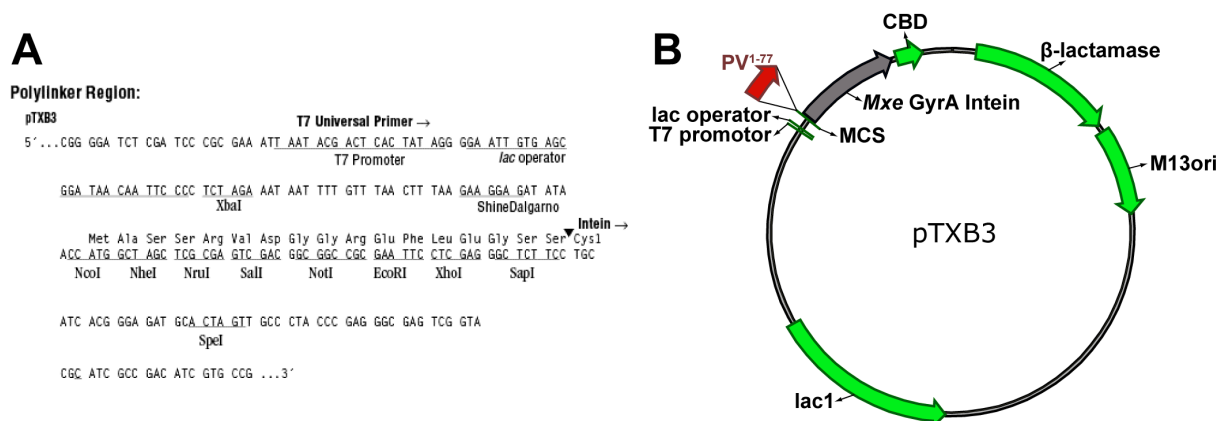
### 3.6.1. Protein Expression in *E. coli*

**Induced Expression:** The PV<sup>1-77</sup>-intein fusion protein was expressed as follows. Overnight cultures of transformed BL21 (DE3) *E. coli* cells were diluted into 500 ml LB medium (10 g/l tryptone, 5 g/l yeast extract, 5 g/l NaCl) to a starting OD<sub>600</sub> of 0.1. Protein expression was induced by the addition of 500  $\mu$ l 1 M  $\beta$ -D-1-thiogalactopyranose (IPTG) at an OD<sub>600</sub> of 0.6. The cells were harvested after 5 hours of expression and pelleted by centrifugation, frozen in liquid nitrogen and stored at -80 °C.

**Autoinduced Expression:** Expression of the PV mutant C18S/H26F was performed by autoinduction in a lactose medium. Pre-cultures of transformed BL21 (DE3) *E. coli* cells were diluted into 500 ml lactose medium (10 g/l tryptone, 5 g/l yeast extract, 25 mM Na<sub>2</sub>HPO<sub>4</sub>, 25 mM KH<sub>2</sub>PO<sub>4</sub>, 50 mM NH<sub>4</sub>Cl, 5 mM Na<sub>2</sub>SO<sub>4</sub>, 2 mM MgSO<sub>4</sub>, 0.18 mM CaCl<sub>2</sub>, 68 mM glycerol, 2.8 mM glucose, 5.8 mM lactose and 0.16 nM carbenicillin) to an OD<sub>600</sub> of 0.1 and grown overnight at 37 °C. The cells were harvested the next day, frozen in liquid nitrogen and stored at -80 °C.

### 3.6.2. PV<sup>1-77</sup> Purification and Intein Mediated Thioester Formation

For the preparation of the PV<sup>1-77</sup>-protein thioester, the pTXB3 expression vector of the IMPACT<sup>TM</sup> kit from New England Biolabs Inc. (Ipswich, MA, USA) was used. The PV<sup>1-77</sup> coding sequence was cloned in-frame into the polylinker (multiple cloning site, MCS) upstream of the *Mxe* GyrA-intein/chitin binding domain using the NcoI and SapI restriction sites (Figure 3.1 A). Induced expression of the fusion protein was performed as described in Chapter 3.6.1. Pelleted cells were thawed and resuspended in standard buffer (list of used buffers in Table 3.4). Lysis was performed with a SONOPLUS ultrasound device from Bandelin (Berlin, Germany) and a Basic Model Z microfluidizer from Constant Systems (Kennesaw, GA, USA). Inclusion bodies containing the fusion protein were pelleted by centrifugation for 15000 rpm, 30 min at 4 °C and the supernatant discarded. In order to remove the cell debris, the inclusion bodies were washed 2 times with washbuffer I followed by two 2 times with washbuffer II and two times with standard buffer. The inclusion bodies were resolved in unfolding buffer and stirred for 1 hour at room temperature. The urea concentration was lowered by dialysis for two hours against 3 M, 1.5 M and 0 M urea at room temperature, followed by a dilution



**Figure 3.1.:** Cloning of PV<sup>1-77</sup> in frame with the *Mxe GyrA*-intein/chitin binding domain into the pTXB3 vector using the NcoI and SapI restriction sites (A). The sequence in Figure A was taken from the New England Biolabs Inc. homepage.

with three volumes of standard buffer. Eventually accumulating precipitate was removed by centrifugation and filtration. The refolded protein from inclusion bodies was loaded onto an equilibrated chitin column with chitin buffer at a flow rate of 1 ml/min. After washing with 10 column volumes of chitin buffer, the column was flushed at a flow rate of 10 ml/min with 3 column volumes of cleavage buffer. Subsequently the flow was stopped and the on-column cleavage was performed overnight at room temperature. The PV<sup>1-77</sup>-MESNA thioester was eluted with chitin buffer the following day and analyzed by sodium dodecyl sulfate polyacrylamide gel electrophoresis (SDS-PAGE, Chapter 3.6.5) in the absence of thiols. The fractions were frozen in liquid nitrogen and stored at -20 °C.

**Table 3.4.:** List of buffers needed for protein thioester formation

Buffer name	Composition		
Standard buffer	20 mM HEPES <sup>a</sup> , pH 8	0.5 M NaCl	1 mM EDTA
Wash buffer I	1 mg/ml deoxycholic acid	-	-
Wash buffer II	PBS <sup>b</sup> , 1% triton X-100, 25% sucrose	-	5 mM EDTA
Unfolding buffer	20 mM HEPES <sup>a</sup> , pH 8, 6 M urea	0.5 M NaCl	1 mM EDTA
Chitin buffer	20 mM HEPES <sup>a</sup> , pH 8	0.5 M NaCl	-
Cleavage buffer	20 mM HEPES <sup>a</sup> , pH 8, 0.2 M MESNA <sup>c</sup>	0.5 M NaCl	-

<sup>a</sup>4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid, pH adjusted with HCl or NaOH

<sup>b</sup>phosphate buffered saline

<sup>c</sup>sodium 2-sulfanylethanesulfonate

#### 3.6.3. PV C18S/H26F Purification

The purification of the PV pseudo wild-type mutant C18S/H26F was performed by a modified protocol of Moncrieffe *et al.*<sup>142</sup>. Cell pellets were thawed on ice and resuspended in 40 mM 2-Amino-2-hydroxymethyl-propane-1,3-diol (Tris) pH 8, 2 mM DTT, 5 mM CaCl<sub>2</sub> (buffer A) and the protease inhibitors phenylmethanesulfonyl fluoride (PMSF, 1 mg per 10 ml of lysate), aprotinin (20 μg per 10 ml of lysate), pepstatin A (10 μg per 10 ml of lysate) and a spatula tip of benzamidine. After cell lysis by sonication with a SONOPLUS ultrasound device from Bandelin (Berlin, Germany), the sonicate was spun for 30 minutes at 18000 rpm. Ammonium sulfate was added slowly to the supernatant up to a concentration of 70% (w/v) and stirred overnight at 4 °C. Precipitate was spun down the next day and the supernatant dialyzed in a 6000-8000 mass weight cutoff (MWCO) Spectra/Por dialysis bag (Spectrum Laboratories, Inc, Rancho Dominguez, CA, USA) against 5 l of H<sub>2</sub>O for 1 hour, followed by two times dialysis for 4 hours against 5 l of 20 mM Tris, 1 mM DTT, 5 mM CaCl<sub>2</sub> and dialysis once for 4 hours against 10 l of buffer A. The solution was then loaded to a HiTrap diethylaminoethanol-sepharose fast flow column from GE Healthcare Life Sciences (Chalfont St Giles, UK) and the flow through containing the protein collected. Residual impurities were removed by HPLC using a Pack C8 column (10 mm x 250 mm), with 5 μm particle size and 20 nm pore size from YMC Europe GmbH (Dinslaken, Germany). Peak fractions were frozen and lyophilized. The purity was checked by analytical HPLC using a Pack C8 column (4.6 mm x 250 mm), with 5 μm particle size and 20 nm pore size from YMC Europe GmbH (Dinslaken, Germany). The peaks with the desired product were identified with MALDI-TOF and ESI-MS.

#### 3.6.4. Native Chemical Ligation (NCL)

Full length PV (residues 1-108) bearing the triplet donor xanthone was prepared as follows. Eluted fractions from the chitin column containing the PV<sup>1-77</sup>-MESNA thioester were thawed and pooled together. The thioester was concentrated and transferred to ligation buffer (50 mM Tris pH 8, 100 mM NaCl, 150 mM MESNA) using VIVASPIN 20 centrifugal concentrators with a mass weight cutoff of 3,000 Da from Sartorius (Göttingen, Germany). The final concentration of the PV<sup>1-77</sup>-MESNA thioester in the ligation buffer was determined with the method after Bradford<sup>165</sup> and the same concentration of the synthesized PV<sup>78-108</sup>Cys<sup>1</sup>Xan<sup>11</sup> fragment was added. Urea was then added to a final concentration of 1.5 M and the mixture shaken overnight at room temperature. The success of the ligation reaction was checked with

SDS-PAGE (Chapter 3.6.5), analytical HPLC with a Pack C8 column (4.6 mm x 250 mm), with 5  $\mu$ m particle size and 20 nm pore size from YMC Europe GmbH (Dinslaken, Germany) and MALDI-TOF. Eventually oxidized cysteines were reduced by incubation of the ligation product in 50 mM Tris pH 8, 100 mM NaCl, 2 mM dithiothreitol (DTT) overnight at room temperature. Reduced full length PV<sup>1-108</sup>Cys<sup>78</sup>Xan<sup>88</sup> was purified by HPLC using a Pack C8 column (10 mm x 250 mm), with 5  $\mu$ m particle size and 20 nm pore size from YMC Europe GmbH (Dinslaken, Germany).

### 3.6.5. Sodium Dodecylsulfate Polyacrylamide Gel Electrophoresis

Discontinuous sodium dodecylsulfate polyacrylamide gel electrophoresis (SDS-Page) after Laemmli<sup>166</sup> was used to analyze the purity of protein fractions or protein samples according to their size. The polyacrylamide gel containing a separating and a stacking gel were prepared using a Serva multiple gel caster from GE Healthcare Life Sciences (Chalfont St Giles, UK) according to Table 3.5. Protein samples were mixed 1:1 (v/v) with Laemmli Sample Buffer

**Table 3.5.:** SDS-PAGE casting of separating and stacking gel

Chemical	Separating Gel	Stacking gel
40% Acrylamide/Bisacrylamide 29:1	1.85 ml	288 $\mu$ l
Water	1.13 ml	1.95 ml
Gel buffer <sup>a</sup>	1.50 ml	750 $\mu$ l
TEMED <sup>b</sup>	4.5 $\mu$ l	7 $\mu$ l
10% (w/v) APS <sup>c</sup>	45 $\mu$ l	70 $\mu$ l

<sup>a</sup>Gel buffer: 3 M Tris, 0.3% SDS, pH 8.45

<sup>b</sup>Tetramethylethylenediamine

<sup>c</sup>Ammonium persulfate

from BioRad (Hercules, CA, USA) and kept at 95 °C for 5 minutes before loaded on the gel. When analyzing the purity of PV<sup>1-77</sup>-intein or PV<sup>1-77</sup>-MESNA, no  $\beta$ -mercaptoethanol was added. The molecular weight size marker Roti-Mark 10-150 from Carl Roth (Karlsruhe, Germany) was used as a reference. Electrophoresis was carried out at ~ 100 V, ~ 100 mA. Selective staining of the protein bands was achieved by incubating the gel for 1 hour in 0.1% (w/v) Coomassie Brilliant Blue G-250, 80% methanol, 20% acetic acid, followed by destaining overnight in 10% acetic acid, 5% methanol.

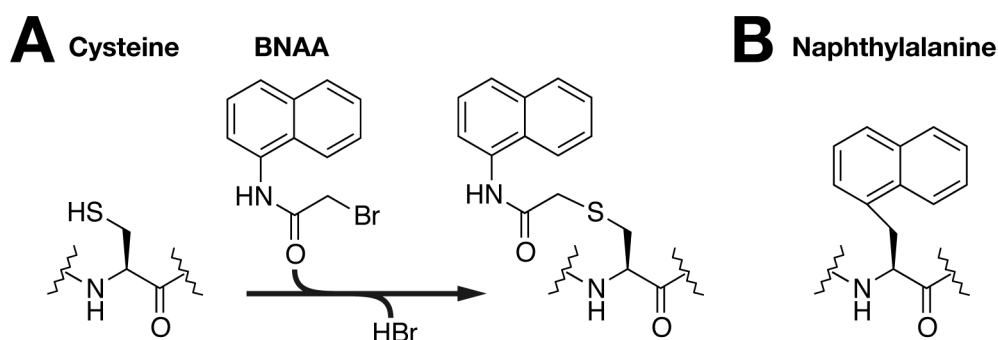
## 3.6.6. Protein Modification

### 3.6.6.1. Desulfurization of Cysteine Residues

Removal of the cysteine-sulfur was performed according to the selective free radical desulfurization method of Wan & Danishefsky<sup>167</sup>, which yields an alanyl-side chain. The cysteinyl-protein was dissolved in 200 mM sodium phosphate, pH 7, 4 M urea and an equal volume of 0.5 M tris(2-carboxyethyl)phosphine (TCEP) bond-breaker<sup>®</sup> solution from Thermo Fisher Scientific Inc. (Rockford, IL USA). The reaction was started by the addition of 5% (v/v) ethanethiol and 2.5% (v/v) 0.1 M 2,2'-Azobis[2-(2-Imidazolin-2-yl)propane]dihydrochloride (VA-044, radical initiator) from Wako Pure Chemical Industries, Ltd. (Osaka, Japan) dissolved in water. After shaking the mixture overnight at room temperature, the conversion of the cysteine residue into alanine was nearly quantitative as determined by analytical or semipreparative HPLC, using the Pack C8 column from YMC Europe GmbH (Dinslaken, Germany).

### 3.6.6.2. Thiol Reactive Labeling

Water soluble 2-bromo-*N*-(1-naphthalen-1-yl)acetamide (BNAA) was used to attach a naphthyl moiety to the side chain of cysteine residues in an S<sub>N</sub>2 reaction (Figure 3.2).



**Figure 3.2.:** Introduction of a naphthalene moiety into proteins and peptides for TTET. (A) Protein labeling of cysteine with 2-bromo-*N*-(naphthalen-1-yl)acetamide (BNAA). (B) Peptide labeling with naphthylalanine (Nal) during SPPS.

Labeling of purified, reduced PV<sup>1-108</sup>Cys<sup>78</sup>Xan<sup>88</sup> was performed in 100 mM Tris/HCl pH 8.48 in the presence of 1.8 equivalents of BNAA, added from a 11.7 mM stock solution in DMF. BNAA concentrations were determined by UV-VIS spectroscopy with an extinction coefficient of 5500 cm<sup>-1</sup>M<sup>-1</sup> at 296 nm. The reaction was incubated for 40 minutes at room temperature

while shaking and the success controlled by analytical HPLC and MALDI-TOF. Excess label was removed by preparative HPLC and fractions with pure PV<sup>1-108</sup>NAA<sup>78</sup>Xan<sup>88</sup> were frozen, lyophilized and stored at -20 °C.

### 3.7. Sample Solutions and Viscosity

Solvent viscosities were measured with a HAAKE falling-ball viscosimeter Type C from Thermo Scientific. The temperature of the solution was adjusted with a F20,HC/7 water bath from Julabo Labortechnik GmbH (Seelbach, Germany). The sample viscosity was calculated from the time  $t$  (s) needed by a ball to traverse a defined distance by

$$\eta = t * (\rho_1 - \rho_2) * K \quad (3.1)$$

where  $\rho_1$  and  $\rho_2$  are the density (g/cm<sup>3</sup>) of the ball and the solvent, and where  $K$  (mPa· $\frac{\text{cm}^3}{\text{g}}$ ) is a ball specific constant.

Peptide and Protein concentrations were determined spectrophotometrically with an Agilent 8453 UV-visible Spectroscopy System (Santa Clara, CA, USA) using the Lambert-Beer relationship. An extinction coefficient of 3900 cm<sup>-1</sup> M<sup>-1</sup> at 343 nm was used for polypeptides labeled with Xan. The concentration of the PV mutant C18S/H26F was determined with an extinction coefficient of 2200 cm<sup>-1</sup> M<sup>-1</sup> at 258 nm<sup>168</sup>.

Urea and GdmCl concentrations were calculated from the refractive indices in the presence ( $n$ ) and absence ( $n_0$ ) of denaturant after Pace *et al.*<sup>104</sup>.

$$[\text{urea}] = 117.66 \cdot (n - n_0) + 29.753 \cdot (n - n_0)^2 + 185.56 \cdot (n - n_0)^3 \quad (3.2)$$

$$[\text{GdmCl}] = 57.147 \cdot (n - n_0) + 38.68 \cdot (n - n_0)^2 - 91.60 \cdot (n - n_0)^3 \quad (3.3)$$

Measured solvent viscosities of the denaturant solutions were compared to calculated viscosities using an empirical equation from Perl *et al.*<sup>169</sup>. Densities and concentrations of TMAO, sarcosine and proline were calculated from the refractive indices using standard curves, which were kindly handed over from Wayne D. Bolen, University of Texas. A standard curve for determining the concentration by refractive index  $n$  of 10 mM potassium phosphate buffered arginine solutions was recorded by dilution of a concentrated arginine stock solution and could

### 3. Material and Methods

---

be described by the following equation

$$[\text{arginine}] = -33.3799 + 25.1866 \cdot n - 0.1140 \cdot n^2. \quad (3.4)$$

Co-solute mole fractions of the different solutions  $\chi_{\text{Cos}}$  were calculated from the co-solute concentrations [Cos] by

$$\chi_{\text{Cos}} = \frac{[\text{Cos}]}{[\text{Cos}] + (\rho - M_{\text{Cos}} \cdot [\text{Cos}]) / M_{\text{H}_2\text{O}}} \quad (3.5)$$

where  $\rho$  is the density of the solution,  $M_{\text{Cos}}$  the molar mass of the co-solute and  $M_{\text{H}_2\text{O}}$  the molar mass of water.

## 3.8. Calculation of Transfer Free Energies

Transfer free energies for the unfolded state were calculated according to Auton and Bolen<sup>101</sup>, using Equation 1.18a,b. The fractional solvent accessibility for the backbone ( $\alpha_i^{\text{bb}}$ ) or the side chains ( $\alpha_i^{\text{sc}}$ ) in the unfolded state were calculated by

$$\alpha_i^{\text{bb/sc}} = \frac{\sum_i n_i ASA_{i,U}^{\text{bb/sc}}}{n_i ASA_{i,G-X-G}^{\text{bb/sc}}} \quad (3.6)$$

where  $n$  is the number of the group of type  $i$  in the polypeptide sequence. The side chain and backbone solvent accessible surface area  $ASA_i$  were obtained from Creamer *et al.*<sup>102</sup> by interpolation between the upper and lower limit for the  $ASA$ .  $ASA_{i,G-X-G}^{\text{bb/sc}}$  describes the standard side chain or backbone solvent accessibility determined from Gly-Xaa-Gly tripeptides by Lee and Richards<sup>170</sup>, where Xaa is the amino acid residue of type  $i$ . Tables with solvent accessibility of the unfolded state and tripeptides as well as group transfer free energies ( $\Delta g_{i;0 \rightarrow 1M}^{0;tr}$ ) can be found in the Appendix (Tables A.1 and A.2).

## 3.9. Laserflash Photolysis

A commercial laser flash photolysis reaction analyzer model LKL.60 from Applied Photophysics Ltd (Surrey, UK) was used for TTET-measurements. 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). Relaxation from the triplet excited state of the xanthone was followed by the change in absorption at 590 nm, measured with a pulsed Xe short arc flash lamp from Osram (München, Germany). The concomitant formation of the triplet state of the acceptor was followed at 420 nm. At least four traces were recorded at each wavelength and averaged for a better signal to noise ratio. Peptide and protein concentrations were in the 15-100  $\mu\text{M}$  range. Samples were degassed prior to TTET measurements in order to minimize triplet quenching by dissolved oxygen.

### 3.10. Circular Dichroism Spectroscopy

All circular dichroism (CD) measurements were carried out on an AVIV DS62 or an AVIV 410 circular dichroism spectropolarimeter (Lakewood, NJ, USA). The fractional helix content  $f_H$  of the PV mutant C18S/H26F was calculated according to Luo & Baldwin<sup>171</sup> at 222 nm with

$$f_H = \frac{[\Theta]_{\text{obs}} - [\Theta]_{0\%}}{[\Theta]_{100\%} - [\Theta]_{0\%}}, \quad (3.7)$$

where 100% and 0%<sup>172</sup> helix content are given by:

$$[\Theta]_{100\%} = (-44000 + 250 \cdot T) \cdot (1 - 3/N) \text{ deg} \cdot \text{cm}^{-2} \cdot \text{dmol}^{-1} \quad (3.8)$$

$$[\Theta]_{0\%} = 13000 \text{ deg} \cdot \text{cm}^{-2} \cdot \text{dmol}^{-1}. \quad (3.9)$$

$[\Theta]_{\text{obs}}$  describes the observed mean residual ellipticity at 222 nm,  $T$  the temperature and  $N$  the number of residues.

### 3.11. Equilibrium Transitions

**Temperature Transitions** Temperature transitions were carried out on an AVIV 410 circular dichroism spectropolarimeter (Lakewood, NJ, USA) at a wavelength of 222 nm in a cuvette with a 10 mm path length while stirring the sample solution. The temperature was increased with a ratio of one degree celsius per minute using a peltier element. The signal was recorded in steps of one degree after temperature equilibration for five minutes.

### 3. Material and Methods

---

**Urea Transitions** Urea transitions of the PV mutant C18S/H26F and the TTET-variant PV<sup>1-108</sup>Cys(NAA)<sup>78</sup>Xan<sup>88</sup> were carried out on an AVIV 410 circular dichroism spectropolarimeter (Lakewood, NJ, USA) at a wavelength of 222 nm in a cuvette with a 1 mm path length at 25 °C and 1.5 nm bandwidth. The signal was recorded and averaged over 60 s and the buffer signal subtracted. Samples were prepared from a protein stock solution in 10 mM cacodylic acid, pH 7 and urea by dilution in buffers with different urea concentrations. Transitions were performed in the presence and absence CaCl<sub>2</sub>.

## 4. Results and Discussion

### 4.1. Chain Dynamics and Barriers for Loop Formation in Unfolded Polypeptide Chains

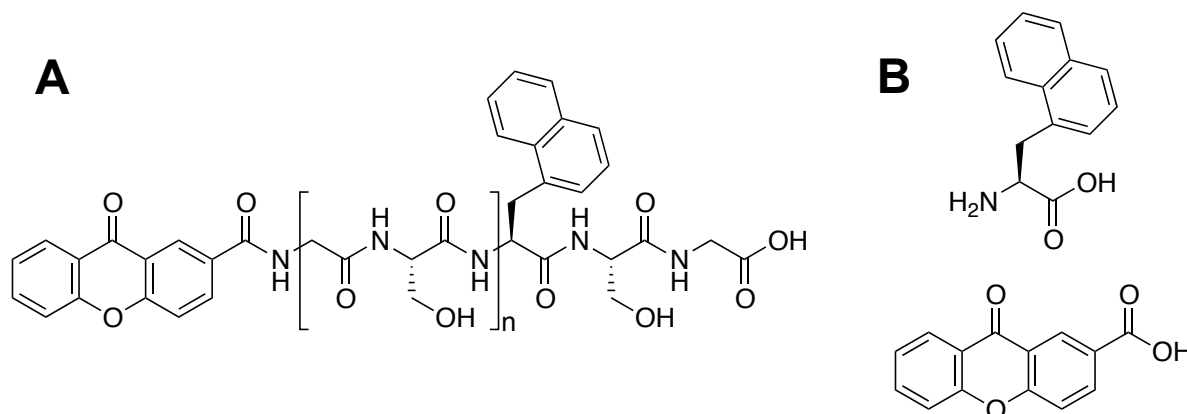
Chain dynamics in the unfolded state ensemble are the process, by which polypeptide chains explore conformational space on their way to acquire native-like interactions. Knowledge about the dynamics of the unfolded state ensemble and the barriers for forming specific interactions by loop formation are thus essential to understand elementary steps in protein folding.

#### 4.1.1. Loop Formation in Model Peptides

To study chain dynamics and barriers for loop formation in U, unstructured poly(Gly-Ser) peptides were used as a model for the unfolded state. These peptides serve as model for a simple chain, where amino acid side chains are absent and no other interaction than intramolecular or peptide-solvent hydrogen bonds can be formed. Loop formation was measured in Gly-Ser peptides of different length<sup>34</sup>. The peptides all share a uniform chemical structure, which is depicted in Figure 4.1, where  $n$  describes the number of Gly-Ser repeats. The repetitive architecture ensures, that only the influence of chain length but not amino acid composition on chain dynamics is observed. The triplet donor xanthone (Xan) was attached to the N-terminus and the triplet acceptor moiety naphthalene was introduced via the nonnatural amino acid L-1-naphthylalanine (Nal) in the vicinity of the C-terminus. Intramolecular TTET from Xan to Nal was measured by Krieger *et al.*<sup>34</sup> for peptides with up to 14 (GS)-repeats ( $N=29$ ). In short peptides with less than 20 peptide bonds, the rate constant for loop formation decreases only weakly with increasing peptide length. In peptides with  $N > 20$ , it decreases with  $N^{-1.7 \pm 0.1}$ . In short peptides, loop formation is controlled by contributions from chain stiffness, which limit local chain dynamics. In long peptides, contribution from chain stiffness are low and loop formation is governed by an entropical search in conformational space.

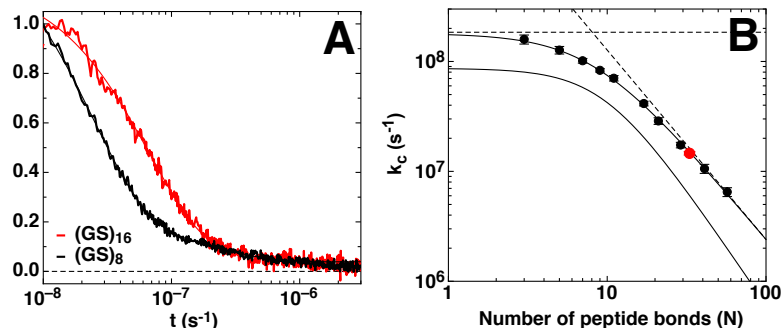
Chain dynamics were investigated in the intermediate long (GS)<sub>8</sub> ( $n = 8$ ) and the long (GS)<sub>16</sub>

#### 4. Results and Discussion



**Figure 4.1.:** (A) Chemical structure of poly(Gly-Ser) peptides. The repetitive sequence within the brackets is repeated  $n$  times. The triplet donor xanthone (Xan) was coupled to the N-terminus forming an amide bond while the triplet acceptor moiety naphthalene was introduced via the nonnatural amino acid L-1-Naphthylalanine (Nal) close to the C-terminus. (B) triplet chromophores Nal (top) and Xan (bottom)

( $n = 16$ ) peptide. In  $(GS)_{16}$ , the time based change in absorbance at 590 nm could be described by a double exponential function (Figure 4.2 A). A fast main phase with 94% amplitude and a



**Figure 4.2.:** Loop formation kinetics poly(Gly-Ser) peptides. (A) Time based change in absorbance at 590 nm. (B) Loop formation rate constant of  $(GS)_{16}$  (●) in comparison with (GS)-peptides of different length.

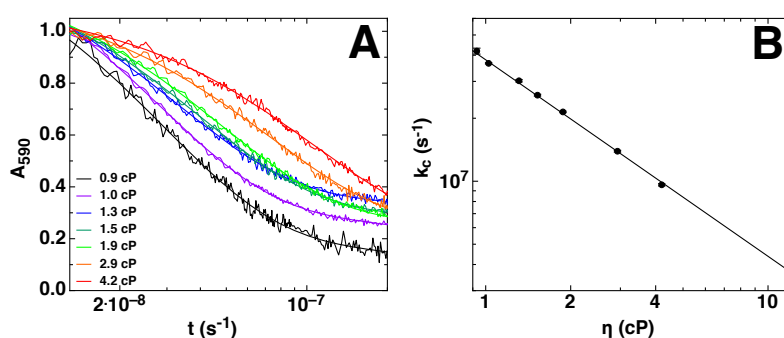
rate constant  $k_c = (1.46 \pm 0.05) \cdot 10^7 \text{ s}^{-1}$  corresponds to intramolecular TTET by loop formation. As expected from earlier findings,  $(GS)_{16}$  behaves like a random chain and lies in the limit, where  $k_c$  scales with  $N^{-1.7 \pm 0.1}$  (Figure 4.2 B). Additionally, a slow phase with low amplitude (6%) and a time constant of  $1.7 \pm 0.2 \mu\text{s}$  was found, which corresponds to the intrinsic donor lifetime and presumably arises from small peptide aggregates, in which the labels are unable to form contact.

### 4.1.2. Effect of Solvent Viscosity on Loop Formation in Model Peptides

The macroscopic solvent viscosity is expected to affect the polypeptide's exploration of the conformational space by chain diffusion. To test the effect of increased solvent viscosity on chain dynamics, TTET was measured in (GS)<sub>8</sub> at different solvent viscosities ( $\eta$ ), which were adjusted by the addition of the viscosifier glycerol. Figure 4.3 A shows the time based absorbance decay at 590 nm at the indicated solvent viscosity. A double exponential fit is necessary to describe the triplet decay of Xan. The main phase corresponds to intramolecular TTET and yields the rate constant for loop formation  $k_c$ . The minor phase corresponds to intrinsic Xan triplet decay in molecules unable to form contact with the acceptor. A linear relationship exists for  $\log k_c$  as a function of  $\log \eta$  (Figure 4.3 B). The data can be described by the empirical Equation 4.1

$$k_c = k_c^0 \cdot \left( \frac{\eta}{\eta_0} \right)^\beta \quad (4.1)$$

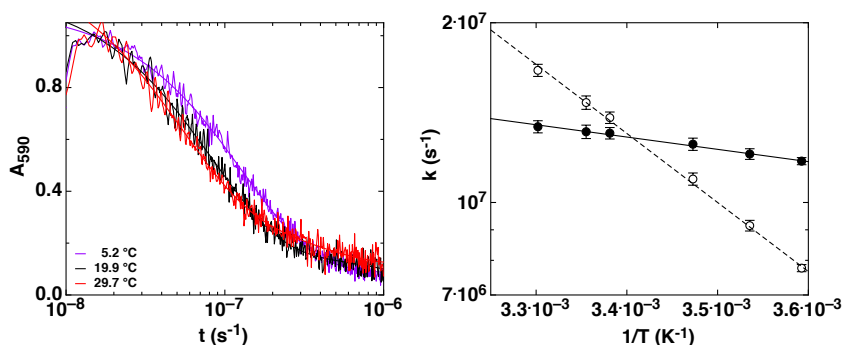
where  $k_c^0$  is the rate constant for loop formation at the reference solvent viscosity  $\eta_0$ . The exponent  $\beta$  reflects the sensitivity of loop formation to solvent viscosity. While a  $\beta$ -value of -1 indicates, that  $k_c$  is inversely proportional on solvent viscosity,  $k_c$  is independent of solvent viscosity for a  $\beta$ -value of 0. The fit yields a  $\beta$ -value of  $0.94 \pm 0.01$  and  $k_c^{\text{H}_2\text{O}} = (4.02 \pm 0.02) \cdot 10^7 \text{ s}^{-1}$ . The  $\beta$  value is close to -1, and reveals an almost inverse proportionality of chain dynamics with solvent viscosity.



**Figure 4.3.:** Effect of solvent viscosity on loop formation in (GS)<sub>8</sub>. (A) Time-based absorbance decay at 590 nm at the indicated solvent viscosity. (B) Rate constant for loop formation as a function of solvent viscosity. The solid line corresponds to a fit of Equation 4.1 to the data.

### 4.1.3. Barriers for Loop Formation in Model Peptides

In order to get insight into the barriers for loop formation dynamics, intramolecular TTET was measured in (GS)<sub>16</sub> between 5 °C and 30 °C. Figure 4.4 shows the time-based decay in absorbance at 590 nm. A significant acceleration for loop formation can be observed with



**Figure 4.4.:** Chain dynamics of (GS)<sub>16</sub> at different temperatures. (A) Time based change in absorbance at 590 nm at the indicated temperatures. (B) Loop formation rate constants as a function of 1/T. The solid lines correspond to a fit of Equation 4.2 to the data.

increasing temperature. A double exponential fit was necessary to describe the decay traces at 590 nm. The fast main phase corresponds to TTET from Xan to Nal and yields the rate constant for loop formation. The origin of the second minor phase is discussed above. When plotting  $\log k_c$  as a function of  $1/T$ , a linear relationship can be found, which can be described by the Arrhenius Equation (Equation 4.2),

$$k_c = A \cdot \exp\left(-\frac{E_A}{RT}\right) \quad (4.2)$$

where  $A$  is the pre-exponential factor and  $E_A$  the activation energy. The apparent activation energy and the pre-exponential factor for loop formation contain contributions from temperature based changes in solvent viscosity. In order to correct  $A$  and  $E_A$  for these contributions, we calculated the viscosity-corrected rate constant for loop formation  $k'_c$  from the observed rate constants  $k_c$  using Equation 4.3 with  $\beta = -1$ .

$$k'_c = k_c \cdot \left(\frac{\eta}{\eta_0}\right)^{-\beta} \quad (4.3)$$

In this Equation,  $\eta$  describes the macroscopic solvent viscosity at the respective temperature, measured by viscosimetry (see Appendix, Figure A.1). As for  $k_c$ , the change in the viscosity-

#### 4.1. Chain Dynamics and Barriers for Loop Formation in Unfolded Polypeptide Chains

corrected rate constant  $k'_c$  with  $1/T$  could also be described by Equation 4.2, yielding the corrected values for  $A$  and  $E_A$ . Within the measured temperature range, no deviation from linearity could be observed for  $\log k_c$  or  $\log k'_c$  as a function of  $1/T$ . This indicates that the change in the activation heat capacity  $C_p^\ddagger$  is negligible in the temperature range between 5 °C and 30 °C. The results from the fit of the Arrhenius Equation to the data are summarized in Table 4.1.

**Table 4.1.:** Parameters for the temperature dependence of loop formation in (GS)<sub>16</sub>

	$A$ (s <sup>-1</sup> )	$E_A$ (kJ/mol)	$\Delta H^{0\dagger}$ (kJ/mol)	$\Delta S^{0\dagger}$ <sup>a</sup> (J/mol/K)
$k_c$	$(1.1 \pm 0.3) \cdot 10^{11}$	$22.1 \pm 0.6$	$19.7 \pm 0.6$	$11.7 \pm 2.3$
$k'_c$	$(6.5 \pm 1.7) \cdot 10^7$	$3.9 \pm 0.6$	$1.5 \pm 0.6$	$-50.2 \pm 2.2$

<sup>a</sup>Calculated with  $k_0 = 10^{10} \text{ s}^{-1}$

When comparing the Arrhenius Equation to the general rate equation  $k = k_0 \cdot \exp(-\Delta G^{0\dagger}/RT)$  (Equation 1.9), the activation enthalpy  $\Delta H^{0\dagger}$  can be expressed by the activation energy  $E_A$  when differentiating  $\ln k$  with respect to the temperature. This results in the following relationship:

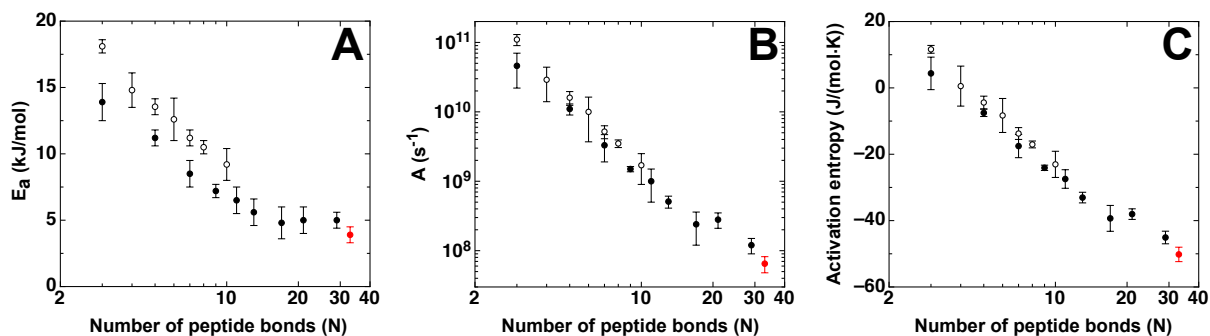
$$\Delta H^{0\dagger} = E_A - RT \quad (4.4)$$

For high activation energies,  $\Delta H^{0\dagger} \approx E_A$ . However, for low activation energies, the difference between  $E_A$  and  $\Delta H^{0\dagger}$  is significant. The viscosity-corrected loop formation rate constant for (GS)<sub>16</sub> reveals a  $\Delta H^{0\dagger}$ -value of 1.5 kJ/mol (Table 4.1). To get information about the entropy of activation associated with loop formation, Equation 4.4 can be used to express the Arrhenius pre-exponential factor  $A$  as a function of the pre-exponential factor  $k_0$  (rate constant in the absence of any barriers), the entropy of activation  $\Delta S^{0\dagger}$  and the universal gas constant  $R$ .

$$A = k_0 \cdot \exp\left(\frac{\Delta S^{0\dagger} + R}{R}\right) \quad (4.5)$$

It is not possible to directly calculate  $\Delta S^{0\dagger}$  from  $A$ , since the correct  $k_0$  value for loop formation is not known. Single bond rotations, which bring the chromophores into van der Waals contact are expected to set the upper limit for intrachain loop formation. We assume  $k_0$  for this process to be approximately  $10^{10} \text{ s}^{-1}$ . Figure 4.5 summarizes  $E_A$ ,  $A$ , and  $\Delta S^{0\dagger}$  for the different lengths of the poly(Gly-Ser) peptides. The pre-exponential factor  $A$  decreases with increasing chain length. Short (GS)-peptides with a narrow conformational distribution have the highest

#### 4. Results and Discussion



**Figure 4.5.:** Activation energy, Arrhenius pre-exponential factor and activation entropy in homo-polypeptides of different length. (A) Activation energy  $E_A$ , (B) Arrhenius pre-exponential factor  $A$  and (C) activation entropy  $\Delta S^{0\ddagger}$  in homo-polypeptides of different length. Data for (GS)<sub>16</sub> is shown as (•). (GS)<sub>n</sub> (•), (S)<sub>n</sub> (○). Figure adapted from Joder<sup>130</sup>. Black data points taken from References<sup>127,130</sup>.

Arrhenius pre-exponential factor. With increasing chain length, a higher number of conformations can be adopted by the polypeptide chain.  $A$  thus decreases from  $4.6 \pm 2.4 \cdot 10^{10} s^{-1}$  in (GS)<sub>1</sub>, to  $6.5 \pm 1.7 \cdot 10^7 s^{-1}$  in the long (GS)<sub>16</sub> peptide. The entropy of activation was found to depend on peptide length, too with negative values for all homo-polypeptides except the shortest homo-polypeptides. While the loss in entropy upon loop formation is smallest in short peptides, it increases with increasing chain length. The loss in conformational entropy is larger in long chains and loop formation is therefore a more unfavorable process. Although loop formation is less unfavorable in the short peptides compared to that in long peptides,  $\Delta S^{0\ddagger}$  is still expected to be negative for loop formation, even in the shortest peptides. This indicates that the value for  $k_0$  is in fact larger than  $10^{10} s^{-1}$ .

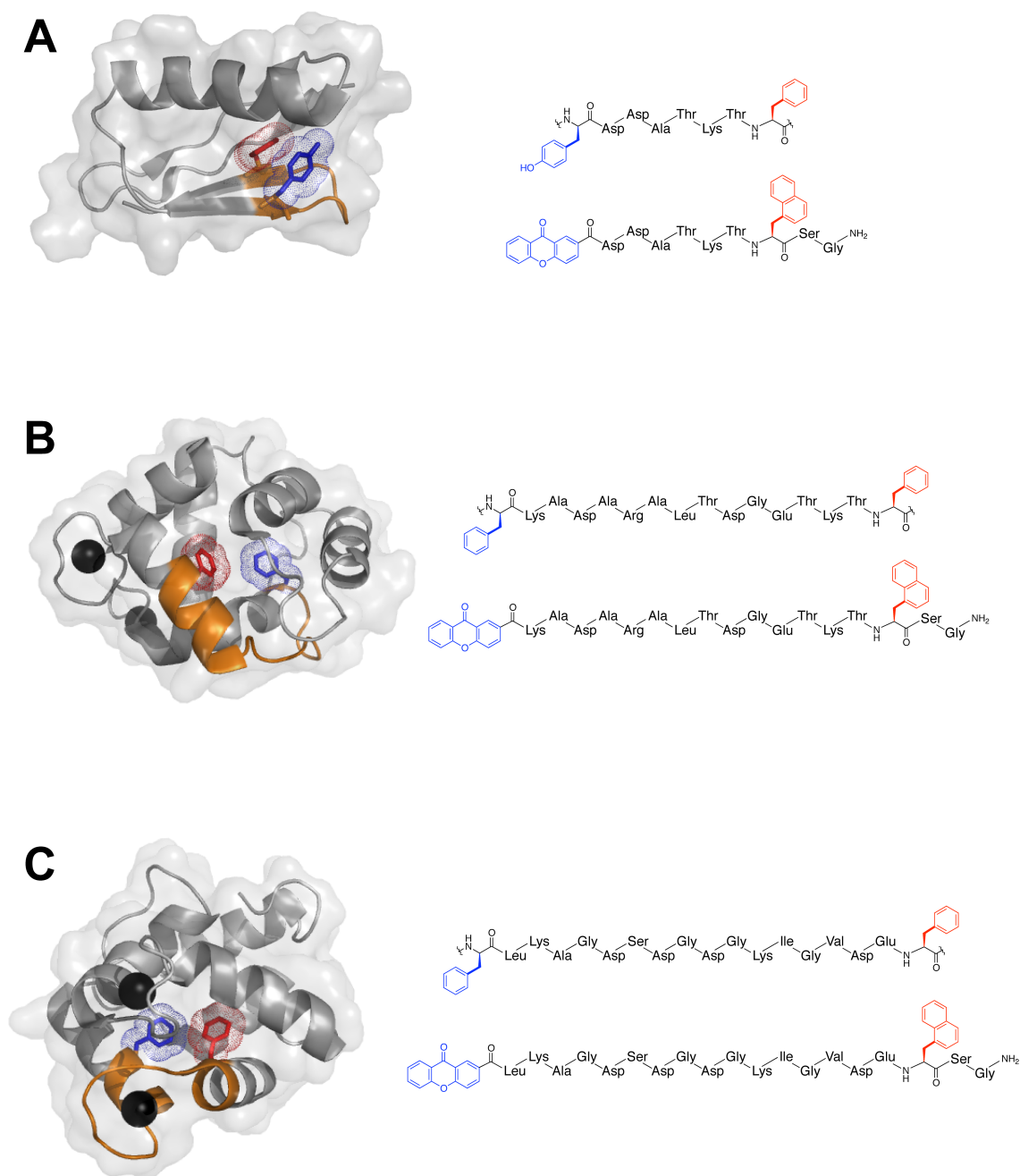
The activation energy  $E_A$  decreases in short (GS)<sub>n</sub> peptides with increasing chain length up to  $n = 6$ , while it is independent of chain length for peptides with  $\geq 8$  peptide bonds. Short and long peptides thus experience different barriers for loop formation. For the short (GS)<sub>1</sub> peptide, an  $E_A = 13.9$  kJ/mol was measured<sup>127</sup>, which corresponds to a  $\Delta H^{0\ddagger}$ -value of 11.4 kJ/mol. Chain stiffness in short peptides is responsible for the high  $E_A$ . In long (GS)<sub>n</sub> peptides ( $n \geq 8$ ), which exhibit statistical chain behavior, an  $E_A$  of 4-5 kJ/mol was found, independent of chain length. This corresponds to a  $\Delta H^{0\ddagger} \approx 2$  kJ/mol at room temperature. However,  $\Delta H^{0\ddagger}$  is assumed to be inexistent in diffusion controlled reactions, where dissipative forces would be absent. Intrachain loop formation in long and flexible (GS) peptides is thus not completely diffusion controlled. A  $\Delta H^{0\ddagger} \neq 0$  in long and flexible peptides could be explained by residual intramolecular interactions, which have to be broken before loop formation, thus impeding loop formation.



### 4.1.4. Chain Dynamics and Barriers for Loop Formation in Natural Sequences

The primary sequence of fragments from naturally occurring proteins normally consist of a variety of amino acids with side chains differing in size and chemical property. We wanted to test, what effect amino acid side chains have on loop formation by comparing chain dynamics in natural sequences (Figure 4.6) and model peptides. We chose three protein fragments, which are part of common local folds such as  $\beta$ -hairpins or the helix-loop-helix structural motif in the native state of the full length protein, but which are unstructured in isolation. They are free of the amino acids histidine (His), methionine (Met), tryptophane (Trp) and tyrosine (Tyr), which would interact with the TTET reaction. It is thus possible to obtain absolute rate constants for intramolecular loop formation. Two fragments from the 108 amino acid long calcium binding protein carp  $\beta$ -parvalbumin (PV) were chosen to be investigated. The loop region connecting helices D and E (residues 70-85) and the conserved calcium binding loop between the helices E and F (residues 85-102, Figure 1.9). 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 Phe was replaced by Xan while the C-terminal Phe was substituted by the nonnatural amino acid Nal, introduced in the course of solid phase peptide synthesis (SPPS). The PV-DE-loop contains a single Gly residue, while the PV-EF-loop contains three between the chromophores. A Gly residue can be found each at the two loop ends flanking the C-terminal end of helix E and the N-terminal end in helix F in the PV-EF-loop. The presence of Gly at the ends of helices is common, since it acts as a helix breaker<sup>175</sup>, due to the high conformational flexibility, which makes it entropically unfavorable for Gly to adopt  $\alpha$ -helical structures. The EF-loop contains further alternating parts of negative sidechains and small uncharged sidechains between residues 89 and 96, with aspartate (Asp) being a frequent neighbour of Gly. The aspartic acid residues at position 90, 92 and 94 are associated in calcium binding by complexing the  $\text{Ca}^{2+}$ -ion in the native state. While calcium was shown to be bound by PV fragments<sup>176,177</sup>, no influence was found on loop formation kinetics<sup>127</sup>. Additionally to the PV sequences, an eight amino acid long turn sequence (residues 45-52) of the 56 amino acid long and highly stable B1 domain of protein G was investigated. In the native state structure, the side chain of Tyr at position 45 is close to the side chain of Phe at position 52. Phe45 was replaced by the triplet donor Xan while Phe52 was substituted by Nal. Amino acid sequences of the synthesized peptide fragments from naturally occurring proteins and the  $(\text{GS})_n$  are summarized in Table 4.2.

## 4. Results and Discussion



**Figure 4.6.:** Peptide fragments from naturally occurring proteins used to study chain dynamics. The structures are shown in the context of the full-length protein on the left in orange with the respective amino acid sequence given on the right. (A) Hairpin residues 45-52 of the B1 domain from protein G, (B) DE-loop of parvalbumin and (C) calcium binding EF-loop of parvalbumin. Blue residues were replaced by Xan, red residues by Nal. The Figure was prepared using PyMOL<sup>173</sup> and the pdb<sup>146</sup> files 4CPV<sup>145</sup> for parvalbumin and 1GB1<sup>174</sup> for the B1 domain of protein G.

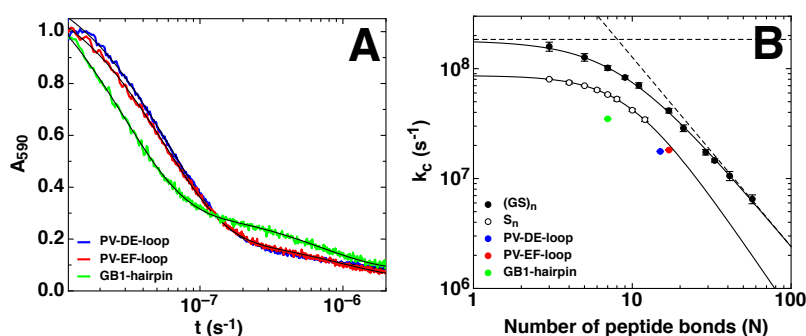
#### 4.1. Chain Dynamics and Barriers for Loop Formation in Unfolded Polypeptide Chains

**Table 4.2.:** Amino acid sequences of investigated peptides. Poly(Gly-Ser) peptides and natural fragments. 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 all fragments in the proximity of the resin in the course of SPPS by Ser-Gly at the C-terminus.

Name	Sequence	charges <sup>a</sup>		net charge <sup>a</sup>
		+	-	
(GS) <sub>8</sub>	Xan-(GS) <sub>8</sub> -Nal-SG-O <sup>⊖</sup>	0	1	-1
(GS) <sub>16</sub>	Xan-(GS) <sub>16</sub> -Nal-SG-O <sup>⊖</sup>	0	1	-1
PV <sup>DE</sup> -loop	Xan-KADARALTDGETKT-Nal-SG-NH <sub>3</sub> <sup>⊕</sup>	4	3	+1
PV <sup>EF</sup> -loop	Xan-LKAGDSDGDGKIGVDE-Nal-SG-NH <sub>3</sub> <sup>⊕</sup>	3	5	-2
GB1-hairpin	Xan-DDATKT-Nal-SG-NH <sub>3</sub> <sup>⊕</sup>	2	2	0

<sup>a</sup>at pH 7

Loop formation in the natural sequences was measured in 10 mM potassium phosphate at pH 7 and 22.5 °C (Figure 4.7). The time based change of the absorbance at 590 nm could be



**Figure 4.7.:** Loop formation kinetics in natural sequences. (A) Time based change in absorbance at 590 nm. (B) Loop formation rate constant in comparison to (GS)<sub>n</sub> and (S)<sub>n</sub> peptides.

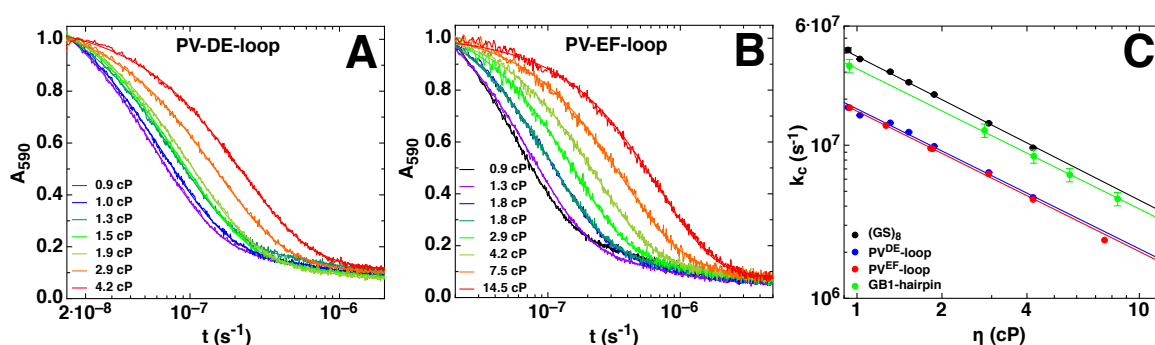
described by a double exponential fit. The main fast phase corresponds to loop formation. A second phase with low amplitude and a time constant around 1 μs was found in all peptides, which corresponds to the intrinsic donor lifetime from small peptide aggregates, in which the labels are unable to form contact. The rate constants for loop formation for all peptides are plotted in Figure 4.7 B in comparison to poly(GS) and poly(S) model peptides. They are in good accordance with the values from Reference<sup>127</sup>. A rate constant for loop formation of  $(1.82 \pm 0.01) \cdot 10^7 \text{ s}^{-1}$  was found for the PV-EF-loop. Its chain dynamics are comparable to the dynamics in polyserine peptides of identical length, although the former contains amino acids with larger side chains, which were shown to slow down intrachain dynamics in host-

## 4. Results and Discussion

guest studies<sup>34</sup>. However, the high glycine content of the EF-loop might on the other hand increase flexibility, in summary resulting in chain dynamics comparable to those observed in polyserine peptides (black open points in Figure 4.8) of the same length. Loop formation rate constants of  $(1.77 \pm 0.01) \cdot 10^7 \text{ s}^{-1}$  and  $(3.49 \pm 0.03) \cdot 10^7 \text{ s}^{-1}$  were found for the PV-DE-loop and the GB1 hairpin, respectively. They are significantly smaller than the rate constants observed in polyserine chains of identical length. This might be due to the presence of bulky amino acids and a low Gly content.

### 4.1.5. Effect of Solvent Viscosity on Loop Formation in Natural Sequences

To investigate the effect of increased solvent viscosity on loop formation in the natural sequences and to compare it with the effect in model peptides, TTET was measured in different glycerol/water mixtures at pH 7 and 22.5 °C (Figure 4.8). Panels A and B of Figure 4.8 show



**Figure 4.8.:** Intramolecular loop formation in natural sequences. (A, B) Time dependent absorbance decay at 590 nm at different solvent viscosities. (C) Rate constants for loop formation as a function of solvent viscosity. (D) Comparison of rate constants for loop formation of natural sequences with  $(GS)_n$  and  $S_n$  peptides. Data for  $(GS)_n$  and  $S_n$  taken from Reference<sup>34</sup>

the time based change in absorbance at the indicated solvent viscosities. A significant deceleration can be observed for the kinetics at high solvent viscosities. The traces could be described by a double exponential fit, with the main phase yielding the rate constant for loop formation, which is shown in Figure 4.8 C as a function of solvent viscosity. A linear decrease of  $\log k_c$  as a function of  $\log \eta$  could be observed and described by Equation 4.1. The obtained parameters from the fit are summarized for the natural sequences in Table 4.3, in comparison to  $(GS)_8$ . For all natural fragments, the  $\beta$ -values are close to -1, revealing an almost inverse viscosity

## 4.1. Chain Dynamics and Barriers for Loop Formation in Unfolded Polypeptide Chains

**Table 4.3.:** Parameters for the viscosity dependence of chain dynamics.

Peptide	$k_c^{\text{H}_2\text{O}}$ ( $\text{s}^{-1}$ )	$-\beta$
(GS) <sub>8</sub>	$(4.02 \pm 0.02) \cdot 10^7$	$0.94 \pm 0.01$
PV-DE-loop	$(1.82 \pm 0.01) \cdot 10^7$	$0.91 \pm 0.01$
PV-EF-loop	$(1.76 \pm 0.01) \cdot 10^7$	$0.92 \pm 0.01$
GB1-hairpin <sup>a</sup>	$(3.27 \pm 0.03) \cdot 10^7$	$0.90 \pm 0.05$

<sup>a</sup>Data taken from Reference<sup>127</sup>.

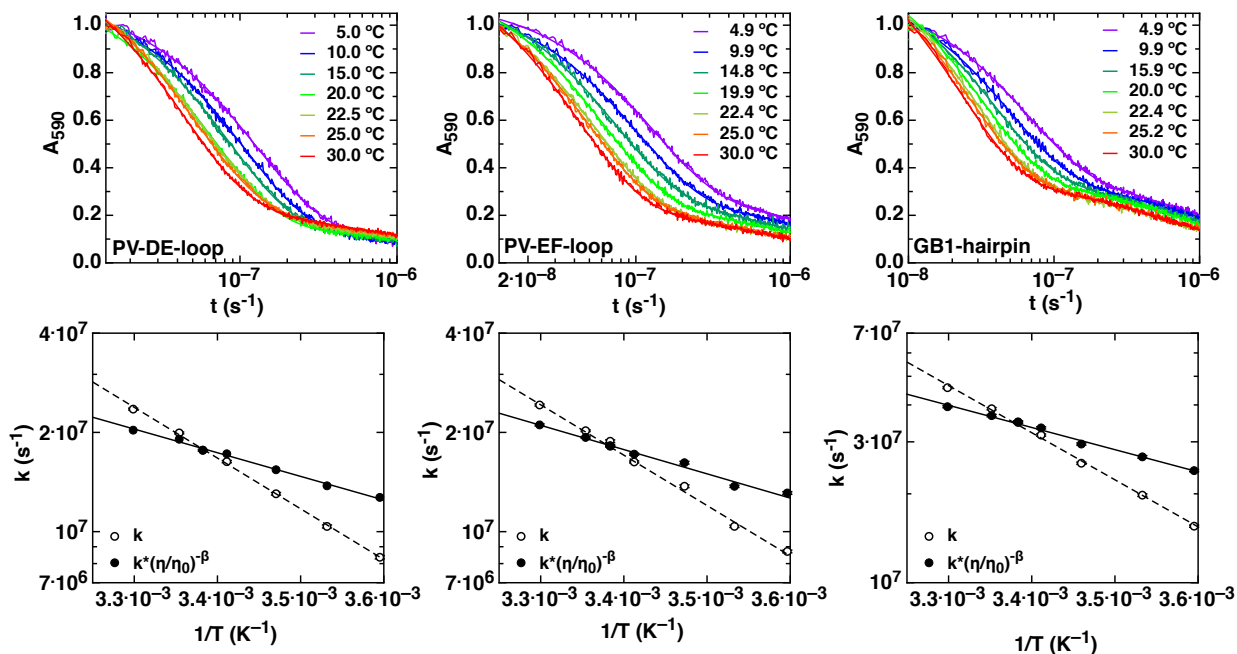
dependence. They are slightly smaller than that of (GS)<sub>8</sub>, due to the influence of the amino acid side chains.

### 4.1.6. Effect of Temperature on Loop Formation in Natural Sequences

To characterize the barriers for intrachain loop formation in natural sequences and to compare them with model peptides, TTET was measured in the PV-DE-loop, the PV-EF-loop and the GB1-hairpin (Figure 4.6) at different temperatures. The temperature dependence of loop formation kinetics are shown in Figure 4.9. As for the (GS)-peptides, loop formation is speeded up significantly with increasing temperature. Double exponential kinetics were necessary to describe the time-based decay at 590 nm at the different temperatures. The main phase corresponding to TTET yields the rate constants for loop formation ( $\circ$ ), which was corrected for changes in solvent viscosity with temperature ( $\bullet$ ) as described in Chapter 4.1.3. The broken and solid lines represent a fit of the Arrhenius Equation (Equation 4.2) to the uncorrected and viscosity-corrected loop formation rate constants, respectively. Table 4.12 summarizes the measured activation energies and pre-exponential factors for the natural sequences in comparison with the (GS)-model peptides.

The natural sequences exhibit both a higher activation energy and a larger pre-exponential factor when compared with (GS)-peptides of the same length. The highest values for  $E_A$  and  $A$  were found for the GB1-hairpin, since chain stiffness is high and large side chains affect loop formation. In the longer PV fragments,  $E_A$  and  $A$  are lower, indicating less stiff chains. The slightly lower  $A$  and  $E_A$  in the PV-EF-loop compared with the PV-DE-loop can be explained by the higher Gly content in the former sequence. The lowest values for  $A$  and  $E_A$  are found for (GS)<sub>8</sub> and (GS)<sub>16</sub>. Due to their long chains and the high Gly content, chain stiffness is low, allowing for efficient loop formation. The Gly content and chain length thus seem to have a

#### 4. Results and Discussion



**Figure 4.9.:** Loop formation dynamics as a function of temperature. Top: Time-based absorbance decay at the indicated temperature. Bottom: uncorrected ( $\circ$ ) and viscosity-corrected ( $\bullet$ ) rate constants for loop formation as a function of temperature. The solid and broken lines represent a fit of Equation 4.2 to the data.

**Table 4.4.:** Arrhenius parameters for the indicated peptide in water.

Peptide	$A$ ( $s^{-1}$ )	$E_A$ (kJ/mol)
$(GS)_1$ <sup>a</sup>	$(4.1 \pm 2.4) \cdot 10^{10}$	$13.6 \pm 1.4$
$(GS)_8$	$(2.4 \pm 1.2) \cdot 10^8$	$4.8 \pm 1.2$
$(GS)_{16}$	$(6.5 \pm 1.7) \cdot 10^7$	$3.9 \pm 0.6$
PV-DE-loop	$(4.6 \pm 1.4) \cdot 10^9$	$14.3 \pm 0.7$
PV-EF-loop	$(1.9 \pm 0.3) \cdot 10^9$	$12.5 \pm 0.7$
GB1-hairpin	$(1.8 \pm 0.7) \cdot 10^{10}$	$15.7 \pm 0.9$

<sup>a</sup> taken from Reference<sup>127</sup>

#### 4.1. Chain Dynamics and Barriers for Loop Formation in Unfolded Polypeptide Chains

large contribution to the chain flexibility and thereby strongly influence the barriers for loop formation.

## 4.2. Effect of Osmolytes and other Co-Solutes on the Dynamics of Loop Formation

Solvent quality describes the characteristics of a solvent for a particular polymer in relation to the characteristics of a reference solvent. In good solvents, peptide-solvent interactions are more favorable than intramolecular interactions. As a consequence, the conformations of a polymer will on average be more stretched out. In poor solvents, intramolecular interactions are more favorable than peptide-solvent interactions and the polypeptide chain will on average be more compact.

We want to modulate the solvent quality for the polypeptide chain and thereby tuning the strength of intramolecular peptide interactions or peptide-solvent interactions, in order to study their effect on loop formation dynamics and barriers. Modulation of the solvent quality can be achieved by adding a third component (co-solute) to the polypeptide-water mixture. Nature found a way to modulate solvent quality by using co-solutes with specific properties, which are called osmolytes. They are present in many organisms, which experience environmental stress conditions like extreme temperatures, desiccation, osmotic pressure or denaturing stress<sup>178</sup>. Osmolytes are small organic molecules and can be produced by the organisms itself as a means of stress response. The function of osmolytes includes counteracting osmotic pressure and adjusting protein stability by modulation of the solvent quality. Protein destabilizing osmolytes or co-solutes render a good solvent, increase solubility and stabilize the unfolded state relative to the native state. Examples of destabilizing compounds are the osmolyte urea and the co-solute GdmCl. Stabilizing osmolytes (stabilizers) were found to do the opposite, *i.e.* lowering solubility and stabilizing the native state of proteins relative to the unfolded state by rendering a poor solvent. Examples for stabilizing osmolytes are trimethylamine *N*-oxide (TMAO) and sarcosine. Additionally, osmolytes exist, which have only a marginal effect on protein stability. Examples for these neutral osmolytes are proline and arginine. Their function might be to increase the equilibrium solubility of both N and U, therefore suppressing aggregation of denatured protein<sup>179</sup> without stabilizing the native state<sup>180–182</sup>. All mentioned substances are soluble up to concentrations of several molar and therefore allow measurements at high co-solute concentrations. This is crucial, since the effect on the polypeptide chain might be small and only prominent at high co-solute concentrations.

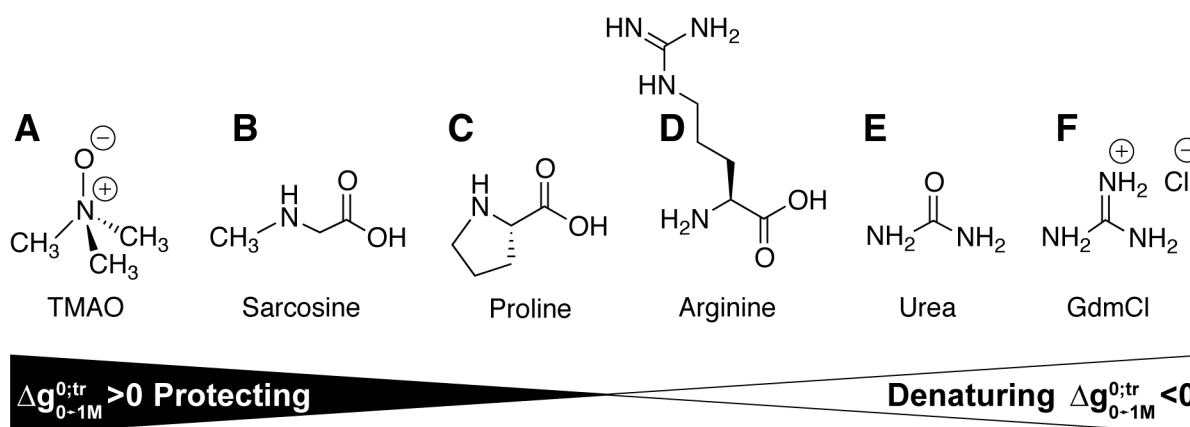
In proteins, osmolytes have been found to interact mainly with the backbone, less with side chains and hardly with hydrophobic groups. Since the main effect is exerted on the backbone,



## 4.2. Effect of Osmolytes and other Co-Solutes on the Dynamics of Loop Formation

it is expected that solvents of different quality lead to the formation and breaking of peptide hydrogen bonds or peptide-solvent hydrogen bonds<sup>93</sup>.

In contrast to proteins, which populate both the unfolded and the native state in equilibrium, we wanted to investigate, whether co-solutes affecting protein stability also affect the dynamics of loop formation in peptide chains, unable to form a folded state. By studying loop formation in unstructured peptides, it is possible to observe the effect of co-solutes they solely exert on the unfolded state.



**Figure 4.10.:** Chemical structure of co-solutes. Stabilizers: Trimethylamine *N*-oxide (TMAO, A), Sarcosine (B). Neutral osmolytes: L-proline (C), L-arginine (D). Denaturants: urea (E), guanidinium chloride (GdmCl, F). Compounds are ordered according to their measured  $\Delta g_{0 \rightarrow 1M}^{0;tr}$  values of a backbone unit from water to 1 M co-solute<sup>183</sup>. Value for L-arginine was provided by A. Tischer and M. Auton (personal communication). Compounds A-E belong to the class of osmolytes.

The influence of different co-solutes on chain dynamics was tested for model peptides (Figure 4.1) and natural sequences (Figure 4.6). By comparing the effect of osmolytes in the different proteins, we are able to determine the influence of osmolytes on the peptide backbone and the amino acid side chains.

### 4.2.1. Effect of Destabilizing Osmolytes and Co-Solutes on the Dynamics of Loop Formation

The destabilizing compounds GdmCl and urea are used to determine protein stability and to investigate kinetics and mechanisms for protein folding. Empirically, a linear relationship between the logarithm of the microscopic rate constant for folding ( $k_f$ ) or unfolding ( $k_u$ ) and the denaturant concentration has been found, yielding the kinetic  $m$ -values  $m_f$  and  $m_u$  for the

#### 4. Results and Discussion

---

folding or unfolding reaction, respectively. However, it is difficult to disentangle the effect of denaturants on protein stability and folding, since both the native and the unfolded state are affected. It is thus not possible to determine the effect of denaturants separately on U or N. However, using flexible chains unable to attain the native state, the effect of denaturants on unstructured polypeptides can be investigated and the contribution of chain dynamics to the  $m_f$ -value assessed.

Intramolecular TTET was measured by time resolved change in absorbance at 590 nm in the presence of different urea or GdmCl (Figure 4.11) concentrations in the peptides used before (Figure 4.6). A double exponential fit is necessary to describe the data, where the main phase with high amplitude (90-95%) corresponds to TTET from Xan to Nal, yielding the rate constant for loop formation  $k_c$ . In all peptides, the denaturants urea and GdmCl reduce the rate constant for loop formation. The value decreases in (GS)<sub>16</sub> from  $(1.39 \pm 0.03) \cdot 10^7 \text{ s}^{-1}$  in water by a factor of 4 to  $(3.45 \pm 0.02) \cdot 10^6 \text{ s}^{-1}$  in 10 M urea and by a factor of 7 to  $(2.00 \pm 0.03) \cdot 10^6 \text{ s}^{-1}$  in 8 M GdmCl. In the PV<sup>DE</sup>-loop  $k_c$  was found to decrease from  $(1.50 \pm 0.01) \cdot 10^7 \text{ s}^{-1}$  in water by a factor of 4 to  $(4.11 \pm 0.01) \cdot 10^6 \text{ s}^{-1}$  in 10 M urea and a factor of 6 to  $(2.50 \pm 0.02) \cdot 10^6 \text{ s}^{-1}$  in 8 M GdmCl. The logarithm of the rate constant for loop formation depends in all investigated peptides linearly on the molar denaturant concentrations of urea and GdmCl. Denaturants thus affect  $k_c$  in a comparable way as they influence the microscopic rate constant for protein folding<sup>34,35</sup>. The effect of denaturants on  $k_c$  can be described by Equation 4.6

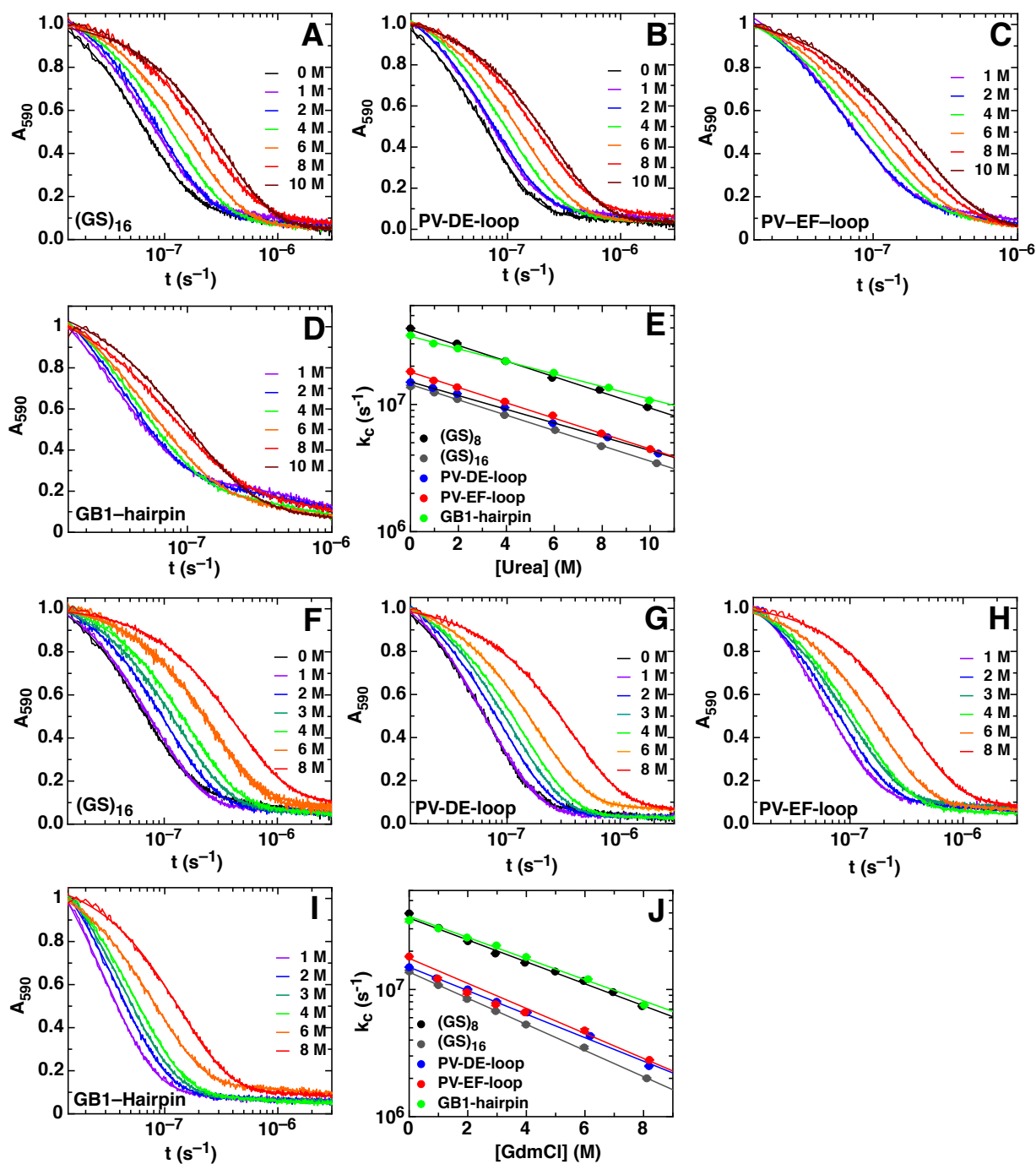
$$\ln k_c = \ln k_c^{\text{H}_2\text{O}} - m_c \cdot \frac{[\text{D}]}{RT} \quad (4.6)$$

where  $k_c^{\text{H}_2\text{O}}$  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. The fit parameters are summarized in Table 4.5. In all peptides, the effect of GdmCl on  $k_c$  is stronger than that of urea, which results in 1.4 - 1.7 times higher  $m_c$ -values in GdmCl compared to  $m_c$ -values in urea.

Möglich *et al.*<sup>35</sup> found a weak length dependence of the  $m$ -value for loop formation in GdmCl and Urea in poly(GS) peptides. The  $m_c$ -value for GdmCl and urea determined for (GS)<sub>16</sub> are slightly larger or identical when compared with the values of (GS)<sub>8</sub>. The three natural sequences also show a weak length dependence of  $m_c$  for both denaturants, where the highest  $m_c$  was found for the longest peptide, the PV-EF-loop in GdmCl.

The presence of denaturants in buffer solutions is known to drastically increase solvent viscosity<sup>169</sup>, which affects chain diffusion. Aside from a viscous effect, urea and GdmCl are

## 4.2. Effect of Osmolytes and other Co-Solutes on the Dynamics of Loop Formation



**Figure 4.11.:** Chain dynamics at various denaturant concentrations. (A-D): Effect of urea on loop formation, monitored by time-based absorbance change at 590 nm. (E) Rate constants for loop formation of the indicated peptide as a function of the urea concentration. (F-I) Effect of GdmCl on loop formation, monitored by time-based absorbance change at 590 nm. (J) Rate constants for loop formation of the indicated peptide as a function of the GdmCl concentration. The solid lines in E and J represents a fit of Equation 4.6 to the data. All measurements were performed in 10 mM potassium phosphate at pH7 and 22.5 °C.

#### 4. Results and Discussion

**Table 4.5.:** Denaturant dependent fitting parameters of Equation 4.6.

Peptide	GdmCl		urea		$\frac{m_c(\text{GdmCl})}{m_c(\text{urea})}$
	$k_c^0$ (s <sup>-1</sup> )	$m_c$ (J/mol/M)	$k_c^0$ (s <sup>-1</sup> )	$m_c$ (J/mol/M)	
(Gly-Ser) <sub>8</sub> <sup>a</sup>	$(3.65 \pm 0.01) \cdot 10^7$	$488 \pm 2$	$(3.86 \pm 0.02) \cdot 10^7$	$348 \pm 2$	1.4
(Gly-Ser) <sub>16</sub>	$(1.37 \pm 0.02) \cdot 10^7$	$582 \pm 13$	$(1.43 \pm 0.01) \cdot 10^7$	$341 \pm 1$	1.7
PV <sup>DE</sup> -loop	$(1.50 \pm 0.01) \cdot 10^7$	$511 \pm 7$	$(1.50 \pm 0.01) \cdot 10^7$	$299 \pm 2$	1.7
PV <sup>EF</sup> -loop	$(1.76 \pm 0.01) \cdot 10^7$	$557 \pm 3$	$(1.80 \pm 0.01) \cdot 10^7$	$343 \pm 1$	1.6
GB1-hairpin	$(3.64 \pm 0.03) \cdot 10^7$	$447 \pm 10$	$(3.43 \pm 0.02) \cdot 10^7$	$280 \pm 5$	1.6

<sup>a</sup>Data taken from Möglich *et al.*<sup>35</sup>

expected to exert an additional, denaturant specific effect. In order to determine this effect, the rate constant for loop formation needs to be corrected for contributions from increased solvent viscosity to chain dynamics. Equation 4.7 is applied to calculate the viscosity-corrected rate constant  $k'_c$  from the loop formation rate constant  $k_c$

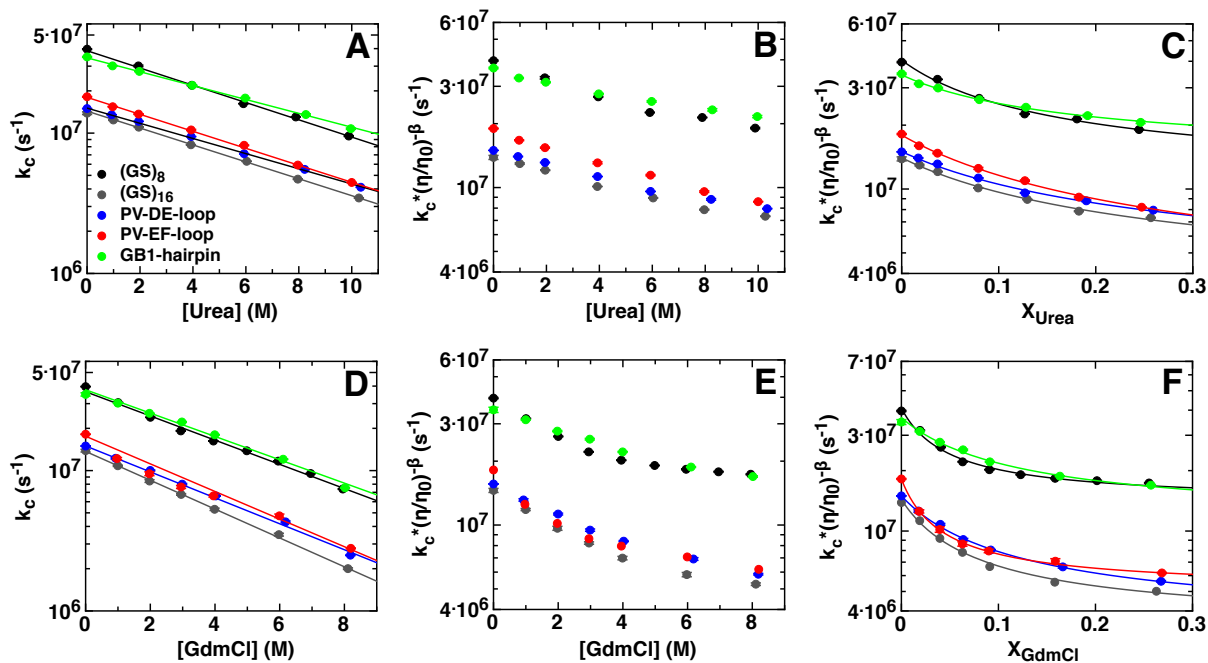
$$k'_c = k_c \cdot \left( \frac{\eta}{\eta_0} \right)^{-\beta} \quad (4.7)$$

$\eta$  is the measured solvent viscosity of the buffer containing urea or GdmCl (see Appendix, Figure A.1). For  $\beta$ , we use the values, which have been determined for each peptide according to Equation 4.1 in the presence of glycerol (Table 4.3).

Figure 4.12 shows  $k'_c$  as a function of the molar denaturant concentration and the denaturant mole fraction in comparison to the uncorrected data (A and D, see also Figure 4.11). In all peptides, the viscosity-corrected rate constants for loop formation  $k'_c$  become slower in the presence of denaturants. They asymptotically approach a limiting value at high urea and GdmCl concentrations, respectively. This indicates, that the specific effect of destabilizing co-solutes becomes saturated at high concentrations. The effect is similar for the (GS)<sub>n</sub> peptides and the natural sequences, pointing towards a unifying mechanism, independent of the amino acid sequence. The effect of urea and GdmCl on  $k_c$  can therefore be explained by the sum of the denaturant specific effect and an increase in solvent viscosity, while the latter effect dominates  $k_c$  at high co-solute concentrations, where the co-solute specific effect is already saturated.

In order to describe the specific effect of destabilizing co-solutes on loop formation, Schellman's weak binding model was applied. In this model, co-solutes and osmolytes are assumed

## 4.2. Effect of Osmolytes and other Co-Solutes on the Dynamics of Loop Formation



**Figure 4.12.:** Non-corrected (A, D) and viscosity-corrected rate constants for loop formation as a function of the molar denaturant concentration (B, E) and the denaturant mole fraction (C, F). The solid lines in C and F correspond to a fit of Equation 4.9 to the data.

to bind to independent sites along the polypeptide chain. Binding to the sites is weak and can only happen via an exchange mechanism where a water molecule needs to leave in order for a co-solute molecule to bind. An average fractional binding site occupancy  $\bar{\nu}$  can be expressed as a function of the co-solute mole fraction  $X_{\text{Cos}}$  by

$$\bar{\nu} = \frac{(K_{\text{ex}} - 1)X_{\text{Cos}}}{(K_{\text{Ex}} - 1)X_{\text{Cos}} + 1} \quad (4.8)$$

where  $K_{\text{Ex}}$  describes the equilibrium constant for the exchange of a water molecule with a denaturant molecule at a single binding site. The effect of co-solute binding on chain dynamics is assumed to be proportional to the fractional binding site occupancy  $\bar{\nu}$ , therefore treating all co-solute binding sites along the polypeptide chain as being identical and independent of each other. Above assumptions lead to the following Equation for the co-solute dependence of  $k_c'$ <sup>35</sup>

$$k_c' = k_c^0 - \gamma k_c^0 \cdot \bar{\nu} \quad (4.9)$$

where  $k_c^0$  is the rate constant for loop formation in the absence of co-solute and where  $\gamma$

#### 4. Results and Discussion

reflects a proportionality constant for the strength of the effect exerted on intrachain diffusion by weakly bound co-solute molecules. The solid lines in Figure 4.12 C and F represent a fit of the weak binding model (Equation 4.9) to the measured data. The fitting parameters  $\gamma$  and  $K_{\text{Ex}}$  are summarized in Table 4.11. Differences in  $\gamma$  for the various peptides are small while

**Table 4.6.:** Denaturant parameters for Schellman's weak binding model

Peptide	GdmCl		urea	
	$\gamma$	$K_{\text{Ex}}$	$\gamma$	$K_{\text{Ex}}$
(Gly-Ser) <sub>8</sub>	0.66 ± 0.01	33.8 ± 0.9	0.69 ± 0.01	13.3 ± 0.6
(Gly-Ser) <sub>16</sub>	0.77 ± 0.01	23.7 ± 2.0	0.78 ± 0.06	7.7 ± 1.3
PV <sup>DE</sup> -loop	0.78 ± 0.01	16.2 ± 0.7	0.78 ± 0.01	7.1 ± 0.2
PV <sup>EF</sup> -loop	0.72 ± 0.01	41.8 ± 2.8	0.83 ± 0.02	8.6 ± 0.6
GB1-hairpin	0.74 ± 0.02	10.4 ± 0.5	0.59 ± 0.03	9.4 ± 0.9

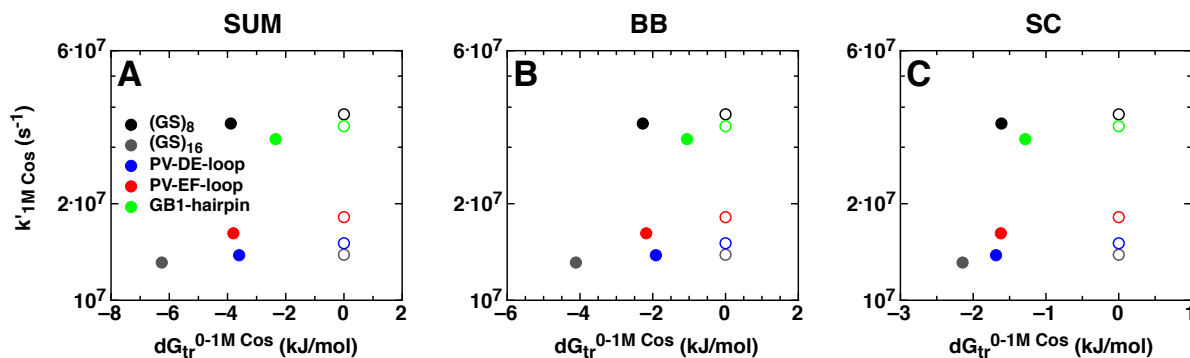
$K_{\text{Ex}}$  differs significantly. At high urea or GdmCl concentrations,  $k'_c$  is speeded up by a similar factor in all peptides but the concentration differs, at which the limiting regime is reached.

Further, Tanford's Transfer Model was used to correlate the thermodynamic effect of the denaturants with the loop formation kinetics. The transfer free energy is obtained by summing over the group transfer free energies (GTFE) of all groups within a polypeptide from water to the respective co-solute ( $\Delta g_{i;0 \rightarrow 1\text{M}}^{0;tr}$ ), multiplied with each groups solvent accessibility,  $\alpha_{i;U}$ .

$$\Delta G_{U;0 \rightarrow 1\text{M}}^{0;tr} = \sum_i \alpha_{i;U} \Delta g_{i;0 \rightarrow 1\text{M}}^{0;tr} \quad (4.10)$$

GTFEs are available for the side chains of the 20 canonical amino acids and the peptide backbone unit from water to 1 M for the co-solutes TMAO, sarcosine, proline, arginine and urea<sup>101,184</sup>. According to earlier findings<sup>101-103</sup>, we used the mean solvent accessibility of an extended and a collapsed unfolded chain for  $\alpha_{i;U}$ . The GTFEs of Trp was used to calculate the GTFE of Nal and multiplied with 1.5 for Xan. Calculated transfer free energies from the GTFE are usually negative (favorable) for the transfer of a polypeptide from water to 1 M denaturant while they are positive (unfavorable) from water to 1 M protectant. Figure 4.13 shows  $k'_c$  as a function of the free energy for the transfer from water to 1 M urea for each peptide sequence. For all peptides, the transfer free energy of both the side chains (SC) and the backbone (BB) from water to 1 M urea is favorable. Accordingly,  $k'_c$  is lower in all peptides at 1 M urea (filled points) compared to  $k'_c$  in water (open points).

## 4.2. Effect of Osmolytes and other Co-Solutes on the Dynamics of Loop Formation



**Figure 4.13.:** Viscosity-corrected rate constant for loop formation as a function of the transfer free energy from water to 1 M urea (closed points) in comparison with water (open points). BB: backbone, SC: sidechain.

In addition to the effect various concentrations of denaturants have on loop formation dynamics, we investigated the influence of destabilizing osmolytes and co-solutes on the barriers for loop formation. Loop formation dynamics were measured at different temperatures with a fixed concentration of urea or GdmCl, respectively. We chose high concentrations of denaturants, since the co-solute effects were expected to be small. Figure 4.14 shows  $k_c$  and  $k'_c$  as a function of temperature in the presence of 6 M GdmCl or 8 M urea.  $k'_c$  was calculated from  $k_c$  by Equation 4.3 with the solvent viscosity  $\eta$  of the buffer at the respective temperature.

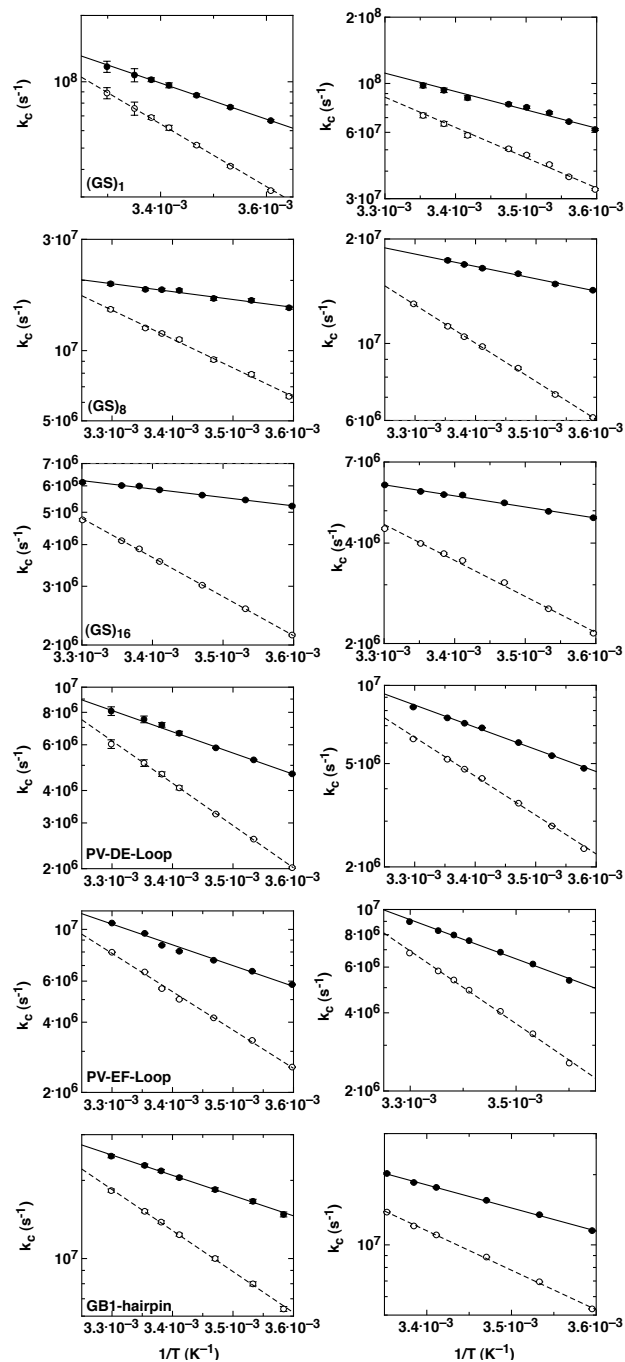
For all peptides,  $k_c$  and  $k'_c$  could be described by the Arrhenius Equation. The fit parameters for for the viscosity-corrected loop formation rate constant are given in Table 4.7 in comparison to the values obtained in water.

Interestingly, destabilizing co-solutes, which are expected to break intramolecular interactions, do not lead to lower activation energies for loop formation compared with those in water. In  $(GS)_{16}$ , a viscosity-corrected activation energy of  $6.4 \pm 0.2 \frac{\text{kJ}}{\text{mol}}$  and  $4.8 \pm 0.2 \frac{\text{kJ}}{\text{mol}}$  was found in the presence of 6 M GdmCl and 8.1 M urea, respectively. When comparing these values to the  $E_A = 3.9 \pm 0.6 \frac{\text{kJ}}{\text{mol}}$  in water, it can be seen that denaturants are not lowering the activation energy and the activation enthalpy but lead to a slight increase in  $E_A$  instead. A similar result was found for  $(GS)_1$ ,  $(GS)_8$  and the natural sequences.

### 4.2.2. Comparison of the Effect of Destabilizing Osmolytes and Co-Solutes on the Dynamics of Loop Formation

Our results show, that denaturants lead to a deceleration of the loop formation rate constant in all investigated peptides. The deceleration can be explained by the sum of two effects, namely

## 4. Results and Discussion



**Figure 4.14.:** Non-corrected and viscosity-corrected rate constants for loop formation as a function of  $1/T$  in the presence of 8 M urea (left) or 6 M GdmCl (right) for the indicated peptide.



**Table 4.7.:** Arrhenius parameters for the indicated peptide in the presence of the destabilizing co-solutes GdmCl and urea.

Peptide	Water		6 M GdmCl		8 M urea	
	$A$ ( $s^{-1}$ )	$E_A$ (kJ/mol)	$A$ ( $s^{-1}$ )	$E_A$ (kJ/mol)	$A$ ( $s^{-1}$ )	$E_A$ (kJ/mol)
(GS) <sub>1</sub> <sup>a</sup>	$(4.1 \pm 2.4) \cdot 10^{10}$	$13.6 \pm 1.4$	$(6.3 \pm 1.6) \cdot 10^{10}$	$16.0 \pm 0.6$	$(6.0 \pm 1.7) \cdot 10^{10}$	$15.7 \pm 0.6$
(GS) <sub>8</sub> <sup>a</sup>	$(2.4 \pm 1.2) \cdot 10^8$	$4.8 \pm 1.2$	$(3.1 \pm 0.3) \cdot 10^8$	$7.2 \pm 0.2$	$(2.5 \pm 0.5) \cdot 10^8$	$6.5 \pm 0.5$
(GS) <sub>16</sub>	$(6.5 \pm 1.7) \cdot 10^7$	$3.9 \pm 0.6$	$(7.5 \pm 0.6) \cdot 10^7$	$6.4 \pm 0.2$	$(4.2 \pm 0.4) \cdot 10^7$	$4.8 \pm 0.2$
PV-DE-loop	$(4.6 \pm 1.4) \cdot 10^9$	$14.3 \pm 0.7$	$(5.5 \pm 0.4) \cdot 10^9$	$16.3 \pm 0.2$	$(4.1 \pm 0.8) \cdot 10^9$	$15.7 \pm 0.4$
PV-EF-loop	$(1.9 \pm 0.3) \cdot 10^9$	$12.5 \pm 0.7$	$(2.8 \pm 0.2) \cdot 10^9$	$14.4 \pm 0.2$	$(2.2 \pm 0.2) \cdot 10^9$	$13.6 \pm 0.2$
GB1-hairpin	$(1.8 \pm 0.7) \cdot 10^{10}$	$15.7 \pm 0.9$	$(4.1 \pm 0.4) \cdot 10^{10}$	$18.9 \pm 0.3$	$(9.4 \pm 2.2) \cdot 10^9$	$14.9 \pm 0.6$

<sup>a</sup> Data for water taken from Reference<sup>127</sup>

#### 4. Results and Discussion

---

an increase in solvent viscosity and a denaturant specific effect. The denaturant specific effect becomes saturated at high concentrations and can be explained by Schellman's weak binding model. The fit yields higher  $K_{\text{Ex}}$  values for the more potent denaturant GdmCl compared to urea. However in the high concentration limit, the strength of the denaturant specific effect is identical for both denaturants.

The activation energies associated with loop formation are slightly increased in both denaturants in comparison to water, while the Arrhenius pre-exponential factors are similar to that in sole water. A possible explanation for the slightly increased  $E_A$  and the slower loop formation evoked by denaturants is, that denaturant molecules might bind to the backbone of the polypeptide, since the effect is independent of the amino acid sequence. This view is supported by the large favorable transfer free energy of the backbone in 1 M urea.

### 4.2.3. Effect of Stabilizing Osmolytes and Co-Solutes on the Dynamics of Loop Formation

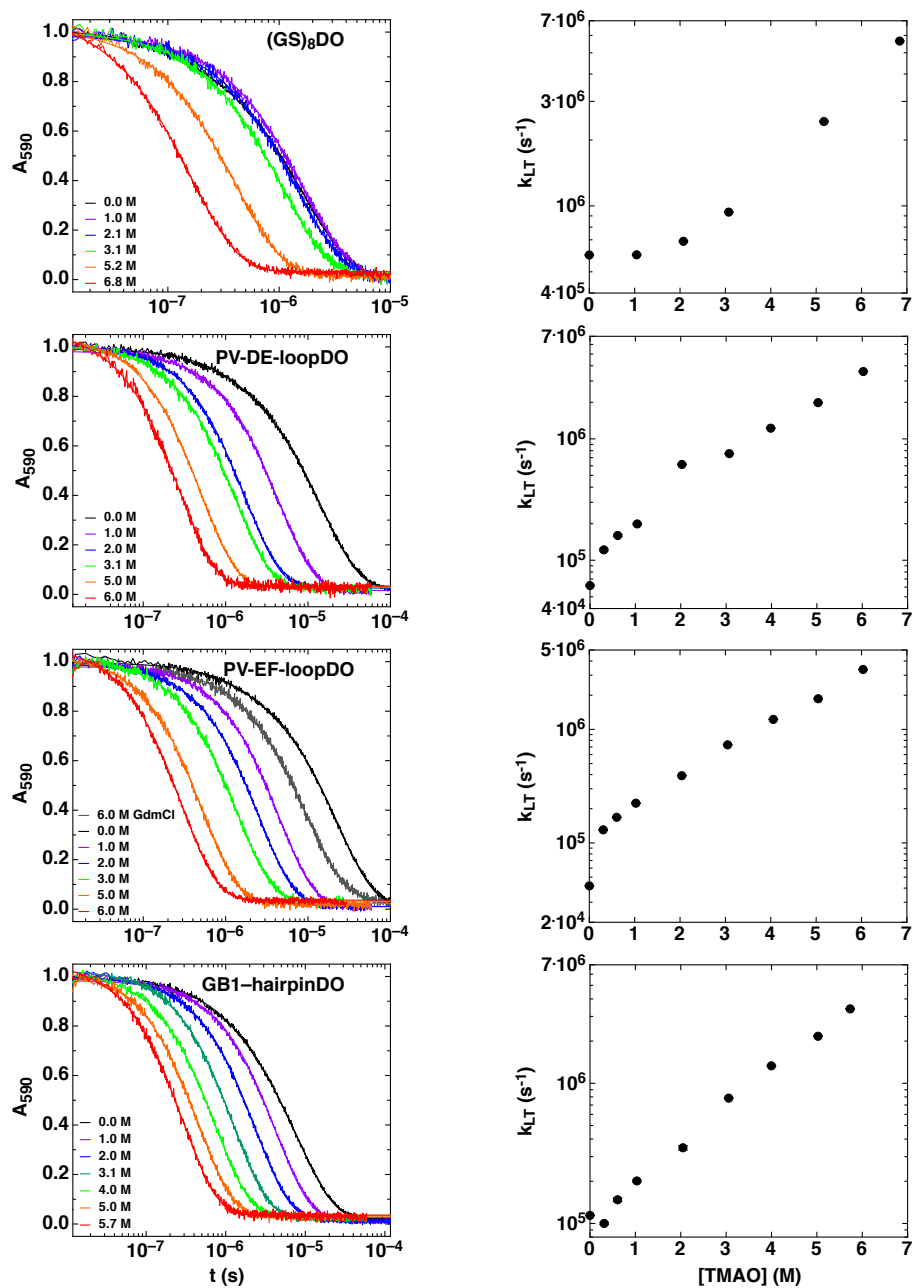
Loop formation kinetics were investigated in the presence of the stabilizing (protecting) osmolytes TMAO and sarcosine to assess their effect on the unfolded state dynamics. To test, whether these compounds affect the intrinsic donor lifetime of Xan, we investigated peptides bearing the donor but not the acceptor. Figure 4.15 and 4.16 show the effect of TMAO and sarcosine on these donor-only peptides, respectively.

The triplet state of xanthone of the donor-only peptides in the presence of co-solutes was followed by time based absorbance at 590 nm. A double exponential fit was necessary to describe the kinetic traces in sole buffer and low stabilizer concentration, as expected for a bimolecular triplet quenching process with approximate identical concentrations of the quenching compound and the donor-only peptide. In case of a quenching compound in excess of donor-only molecules, single exponential kinetics are expected. Accordingly, at co-solute concentrations higher than 1 M (TMAO) or 3 M (sarcosine), single exponential kinetics were observed due to an excess of triplet quenching compounds present in the buffer. While sarcosine and TMAO itself do not interfere with TTET, quenching of Xan is addressed to remaining impurities (*e.g.* residual heavy metal ions). These impurities were present even at the highest commercially available grade for both protectants and could not be removed efficiently by an additional purification step. It is thus not possible to directly obtain absolute rate constants for loop formation from the kinetics in the presence of TMAO or sarcosine. An apparent rate constant ( $k_{app}$ ) is obtained instead, which represents the sum of the rate constant for loop formation and the donor triplet quenching rate constant in the respective buffer. In order to determine the absolute rate constants for loop formation  $k_c$ , donor-acceptor and donor-only peptides were measured under identical conditions and the xanthone decay rate constant determined by the donor-only peptide subtracted from  $k_{app}$ . (GdmCl, urea,  $Na_2SO_4$  and proline did not influence the xanthone lifetime in donor-only peptides in the concentration range used. Therefore,  $k_c$  could be obtained directly from the kinetic traces at 590 nm in the presence of denaturants due to the long intrinsic donor lifetime).

Loop formation dynamics in the presence of TMAO (Figure 4.17) and sarcosine (Figure 4.18) were monitored by TTET in the peptides used above, bearing donor and acceptor.

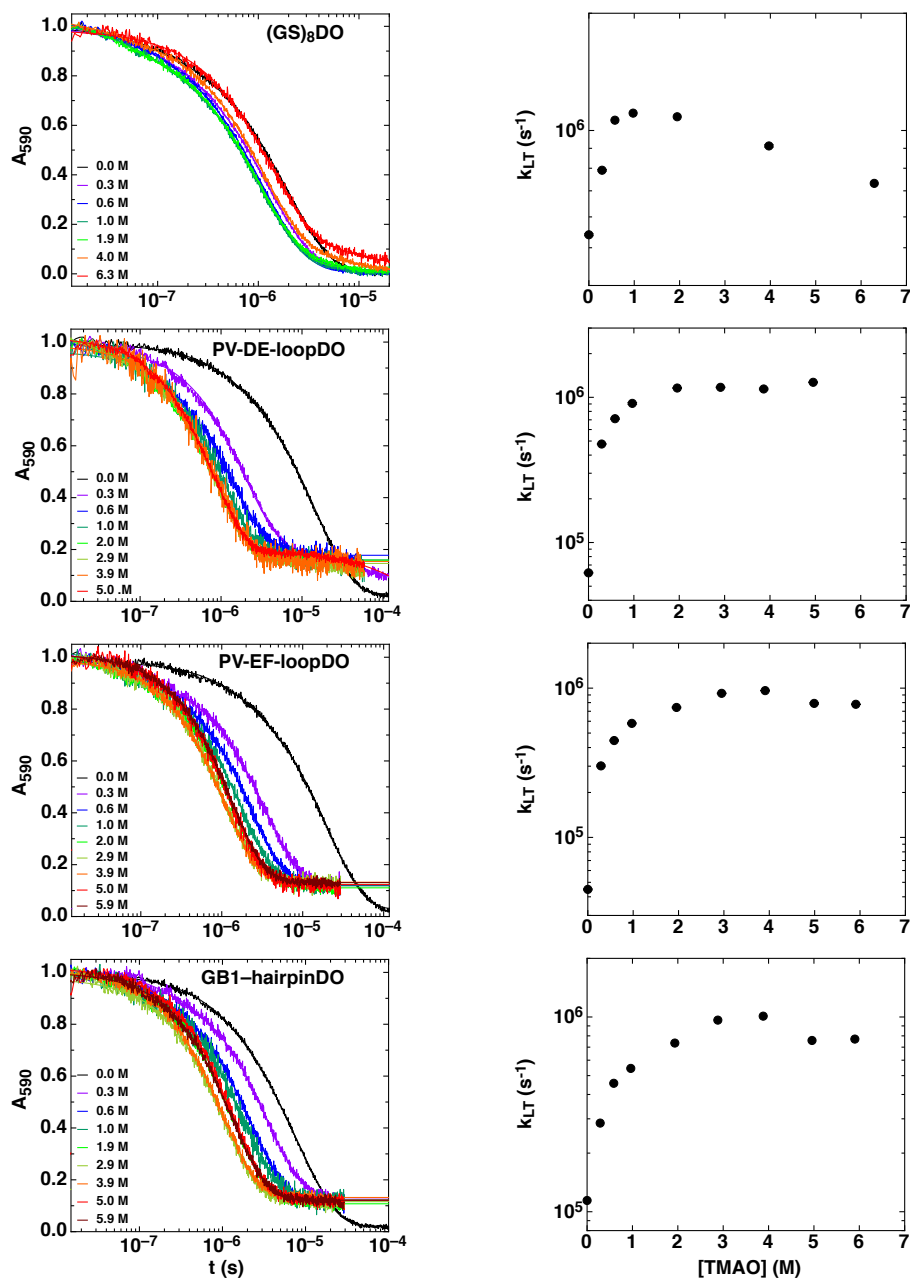
In the presence of protectants, the signal to noise ratio of the normalized kinetic traces is lower than in the presence of denaturants. This is a consequence of the lower solubility of the peptides in the poor solvents formed by TMAO and sarcosine. A double exponential fit was

#### 4. Results and Discussion



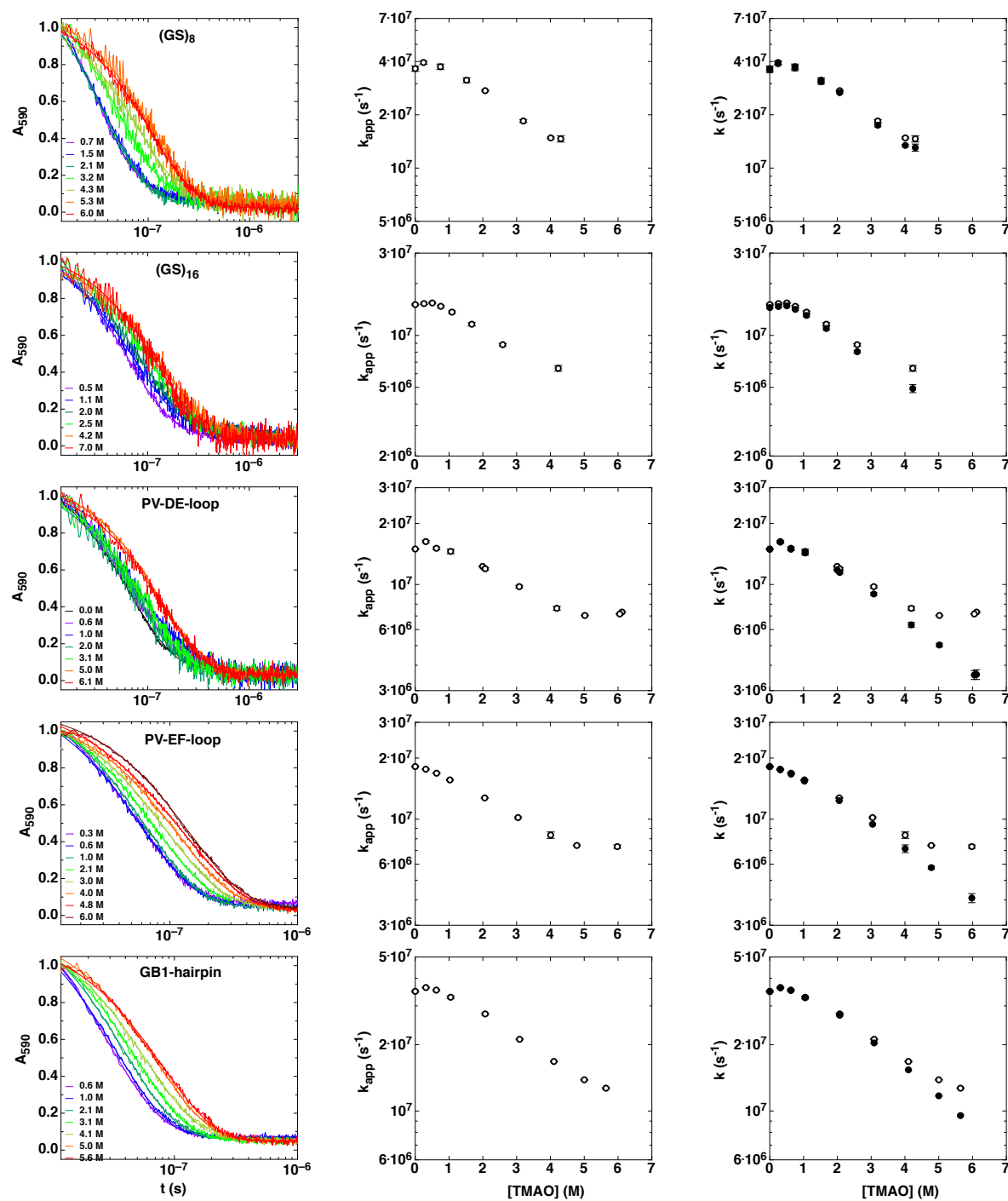
**Figure 4.15.:** Decay of the triplet state in donor-only peptides in the presence of TMAO. Left: Time based change in absorbance at 590 nm. Right: Rate constants for the triplet decay as a function of [TMAO].

## 4.2. Effect of Osmolytes and other Co-Solutes on the Dynamics of Loop Formation



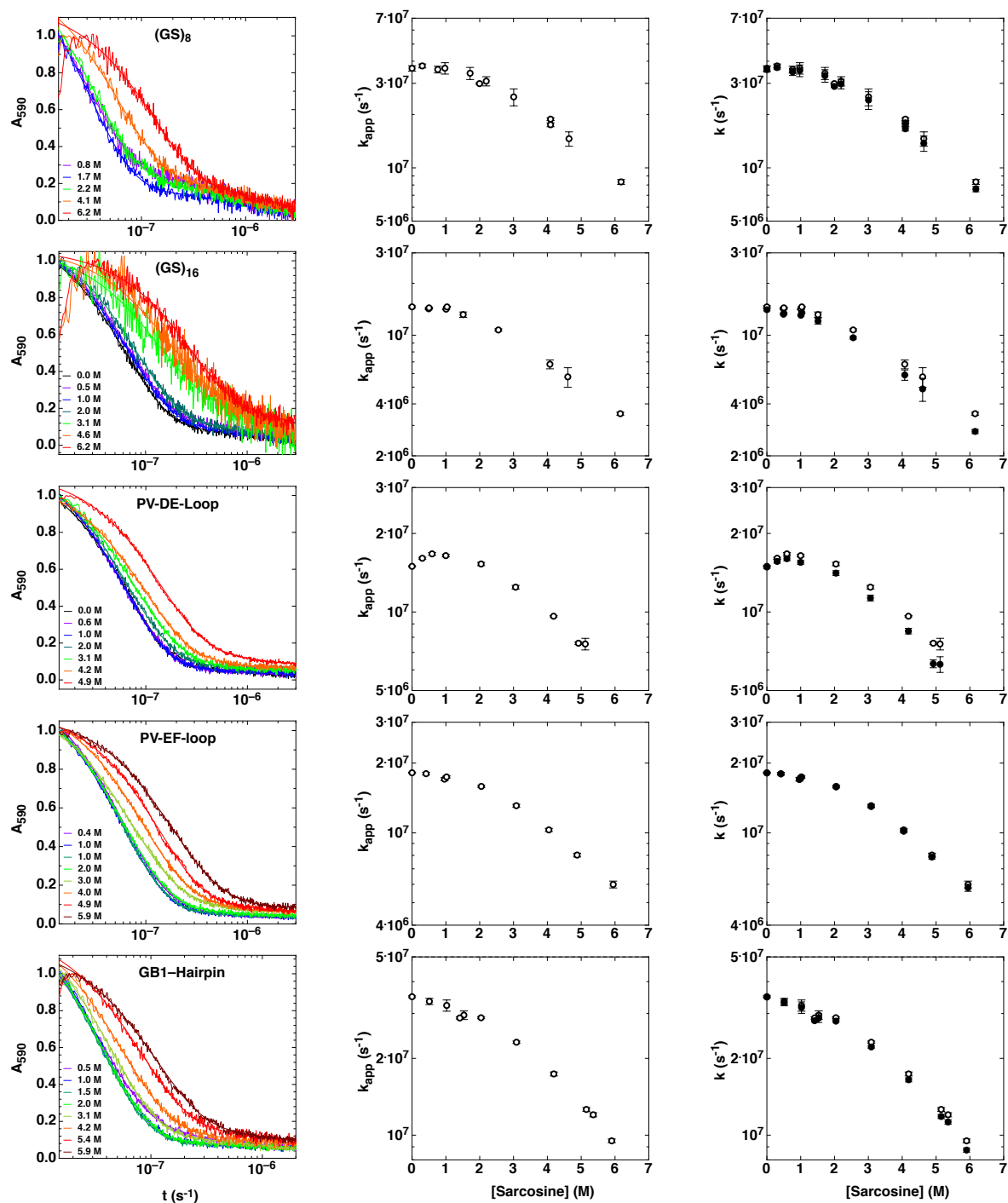
**Figure 4.16.:** Decay of the triplet state in donor-only peptides in the presence of sarcosine. Left: Time based change in absorbance at 590 nm. Right: Rate constants for the triplet decay as a function of [sarcosine].

## 4. Results and Discussion



**Figure 4.17.:** Chain dynamics at various TMAO concentrations. Left: Time-based absorbance change at 590 nm. Middle: Apparent rate constants as a function of the TMAO concentration. Right: Quenching corrected rate constants for loop formation ( $\bullet$ ) as a function of the TMAO concentration in comparison with the apparent rate constant ( $\circ$ ). All measurements were performed in 10 mM potassium phosphate at pH 7 and 22.5 °C.

## 4.2. Effect of Osmolytes and other Co-Solutes on the Dynamics of Loop Formation



**Figure 4.18.:** Chain dynamics at various sarcosine concentrations. Left: Time-based absorbance change at 590 nm. Middle: Apparent rate constants as a function of the sarcosine concentration. Right: Quenching corrected rate constants for loop formation ( $\bullet$ ) as a function of the sarcosine concentration in comparison with the apparent rate constant ( $\circ$ ). All measurements were performed in 10 mM potassium phosphate at pH 7 and 22.5 °C.

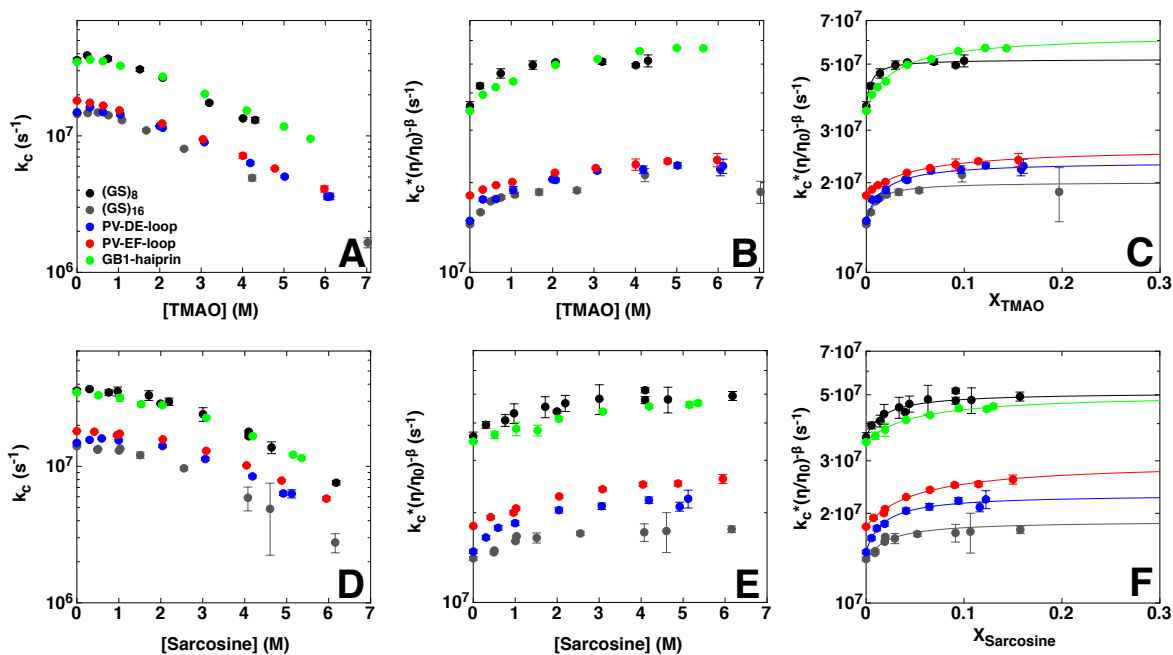
#### 4. Results and Discussion

---

necessary to describe the data for all doubly labeled peptides in sarcosine and at low TMAO concentrations ( $< 1$  M) while single exponential kinetics were observed for higher TMAO concentrations. The main or single phase corresponds to TTET from Xan to Nal and yields an apparent rate constant. For both stabilizing osmolytes, the apparent rate constants display a weak concentration dependence at low TMAO or sarcosine concentrations. While  $k_{\text{app}}$  is decreasing only slightly in most peptides with increasing protectant concentration, it even increases slightly in the PV-DE-loop. The concentration dependence of the rate constant gets more pronounced for intermediate protectant concentrations. At high TMAO concentration, the effect on  $k_{\text{app}}$  decreases again due to the quenching of Xan. The rate constants for loop formation  $k_c$  are shown on the right in Figure 4.15 and Figure 4.16 as a function of the molar co-solute concentration, in comparison to the apparent rate constant. In contrast to denaturants, no linear relationship is found for  $\log k_c$  as a function of the co-solute concentration. At low concentrations,  $k_c$  is only weakly sensitive for changes in [TMAO] or [sarcosine]. For (GS)<sub>8</sub>, (GS)<sub>16</sub>, PV-DE-loop and GB1-hairpin in TMAO as well as the PV-DE-loop in sarcosine,  $k_c$  increases slightly and reaches a maximal value at around 0.4 M of protectant.  $k_c$  then decreases rapidly with increasing [protectant] at higher concentrations. Since both TMAO and sarcosine are known to increase the viscosity of the buffer significantly with increasing co-solute concentration, solvent viscosity is expected to be responsible for the strong decrease in  $k_c$  at high co-solute concentrations. In order to elucidate the specific effect of stabilizing osmolytes,  $k_c$  was corrected for changes in solvent viscosity as described in Chapter 4.2.1, by using Equation 4.7 and the  $\beta$ -values obtained from the viscosity dependence measurement. Non-corrected and viscosity-corrected rate constants are compared for all peptides in Figure 4.19.  $k_c$  and  $k'_c$  are shown as a function of the molar TMAO concentration in Panels A and B and as a function of the molar sarcosine concentration in panels D and E. In panels C and F,  $k'_c$  is shown as a function of the TMAO or sarcosine mole fraction, respectively. In all peptides, both TMAO and Sarcosine lead to an acceleration of the viscosity-corrected rate constants for loop formation. As for loop formation in the presence of the destabilizing co-solutes GdmCl and urea,  $k'_c$  asymptotically approaches a limiting value at high TMAO or sarcosine concentrations, respectively. This indicates, that the specific effect of stabilizing osmolytes becomes also saturated at high concentrations. The effect is similar for the (GS)<sub>n</sub> peptides and the natural sequences, pointing towards a similar mechanism of action, independent of the amino acid sequence. As for the destabilizing co-solutes, the effect of TMAO and sarcosine on  $k_c$  can be explained by the sum of the osmolyte specific effect and an increase in solvent viscosity, while the latter effect dominates  $k_c$  at high co-solute concentrations, where the co-solute specific



## 4.2. Effect of Osmolytes and other Co-Solutes on the Dynamics of Loop Formation



**Figure 4.19.:** Non-corrected (A, D) and viscosity-corrected rate constants for loop formation as a function of the molar stabilizer concentration (B, E) and the stabilizer mole fraction (C, F). The solid lines in C and F correspond to a fit of Equation 4.9 to the data.

effect is already saturated.

The protectant specific effect on loop formation could be described by Schellman's weak binding model (see Chapter 4.2.1). The fit parameters for the weak binding model are summarized in Table 4.8. In contrast to destabilizing osmolytes and co-solutes where positive  $\gamma$ -values

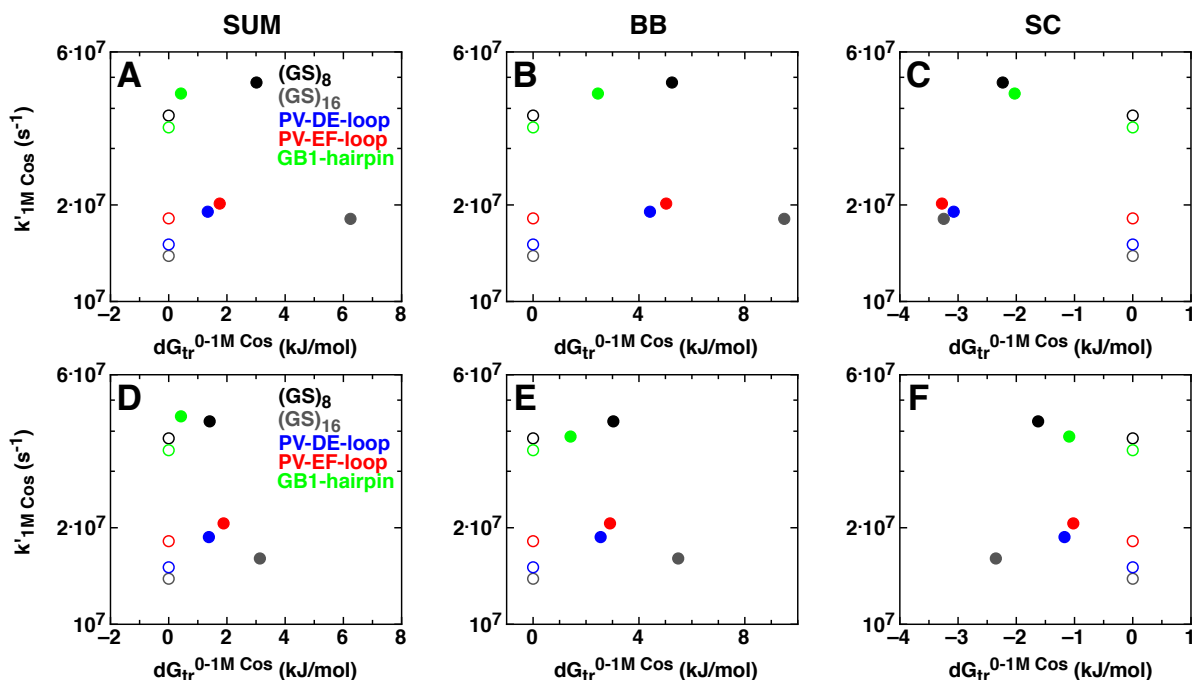
**Table 4.8.:** Parameters for Schellman's weak binding model for protecting osmolytes.

Peptide	Sarcosine		TMAO	
	$\gamma$	$K_{\text{Ex}}$	$\gamma$	$K_{\text{Ex}}$
(Gly-Ser) <sub>8</sub>	$-0.40 \pm 0.06$	$45.2 \pm 25.0$	$-0.45 \pm 0.05$	$180 \pm 55$
(Gly-Ser) <sub>16</sub>	$-0.38 \pm 0.04$	$45.6 \pm 10.3$	$-0.39 \pm 0.03$	$100 \pm 21$
PV <sup>DE</sup> -loop	$-0.56 \pm 0.04$	$47.4 \pm 9.9$	$-0.58 \pm 0.03$	$49.3 \pm 7.3$
PV <sup>EF</sup> -loop	$-0.65 \pm 0.06$	$16.1 \pm 2.9$	$-0.45 \pm 0.02$	$17.9 \pm 2.2$
GB1-hairpin	$-0.45 \pm 0.06$	$18.1 \pm 4.9$	$-0.77 \pm 0.02$	$28.0 \pm 2.3$

have been found, protectants yield negative  $\gamma$ -values in order to account for the increase in  $k'_c$  with increasing protectant concentrations. The absolute effect of denaturants and protectants on  $k'_c$  is comparable. Values for  $K_{\text{Ex}}$  in TMAO and sarcosine are positive as for denaturants,

#### 4. Results and Discussion

while the values for TMAO are generally larger than for sarcosine. The limit for  $k'_c$  is thus reached already at lower concentrations for TMAO than for sarcosine. In contrast to denaturants, which interact favorably with polypeptides, TMAO and sarcosine are expected to be preferentially excluded from the polypeptide vicinity and binding of a water molecules to a site along the polypeptide chain is more probable than binding of a stabilizing osmolyte molecule. Accordingly,  $\gamma$ , which has been positive in denaturants is negative for stabilizing osmolytes. To test, whether  $k'_c$  is correlated with the free energy for transfer from water to 1 M TMAO or 1 M sarcosine,  $\Delta G_{U;0 \rightarrow 1M}^{0;tr}$  was calculated for each peptide according to Chapter 4.2.1 using available GTFE's. Figure 4.20 shows  $k'_c$  as a function of the transfer free energy for each peptide sequence. For all peptides, the free energy of transfer of the backbone from water to



**Figure 4.20.:** Viscosity-corrected rate constant for loop formation as a function of the transfer free energy to 1 M TMAO (A-C) and 1 M sarcosine (D-F) in comparison with water (open points). BB: backbone, SC: side chain.

1 M TMAO and 1 M sarcosine is unfavorable (positive), while it is weakly favorable (negative) for the side chains, resulting in an unfavorable sum of the transfer free energy. Generally, transfer energies are larger for 1 M TMAO than 1 M sarcosine. Accordingly,  $k'_c$  is higher in 1 M TMAO than in 1 M sarcosine in most peptides.

Additionally to data from water to 1 M of osmolyte, GTFE's are also available for the transfer from water to 2, 4 and 6 M sarcosine<sup>185</sup>. The viscosity-corrected rate constant for loop for-

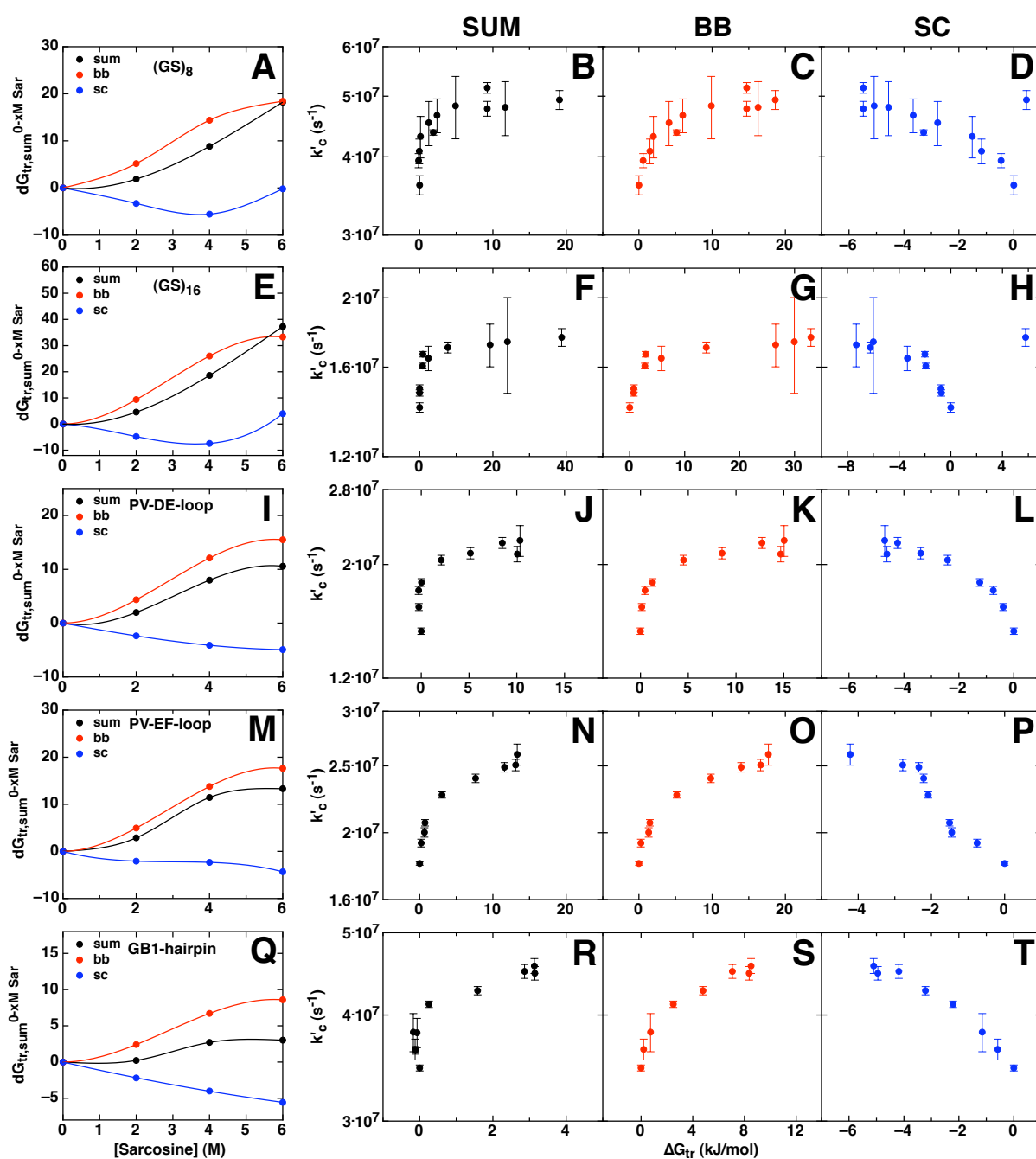
## 4.2. Effect of Osmolytes and other Co-Solutes on the Dynamics of Loop Formation

---

mation as a function of the energy for the transfer from water to the respective [sarcosine] is summarized in Figure 4.21 for all peptides. Transfer energies were calculated by interpolation of the available data (panels on the far left in Figure 4.21). The panels on the left, right and far right show  $\log k'_c$  as a function of the total, the BB or the SC transfer free energy, respectively. As seen before, the free energy associated with the transfer to medium or high concentrations of sarcosine is smaller for the SC when compared with the BB. A slightly favorable transfer free energy is found for the SC, while it is unfavorable for the BB and the sum of BB and SC. There exist two regimes for changes in  $\log k'_c$  both for the sum and the BB transfer free energy. At low concentrations,  $\log k'_c$  increases strongly, while the transfer free energy  $\Delta G^{0;tr}$  changes only marginally. At higher concentrations,  $\log k'_c$  increases only slightly with increasing  $\Delta G^{0;tr}$ . This observation is alike for the natural sequence and the (GS)-repeat peptides, revealing only a marginal effect of side chains. It is not clear, which effects lead to these two regimes at low and high  $\Delta G^{0;tr}$ . Eventually, all possible intramolecular interactions are formed already at a transfer energy of several kJ/mol. An additional decrease in solvent quality then hardly affects loop formation dynamics.

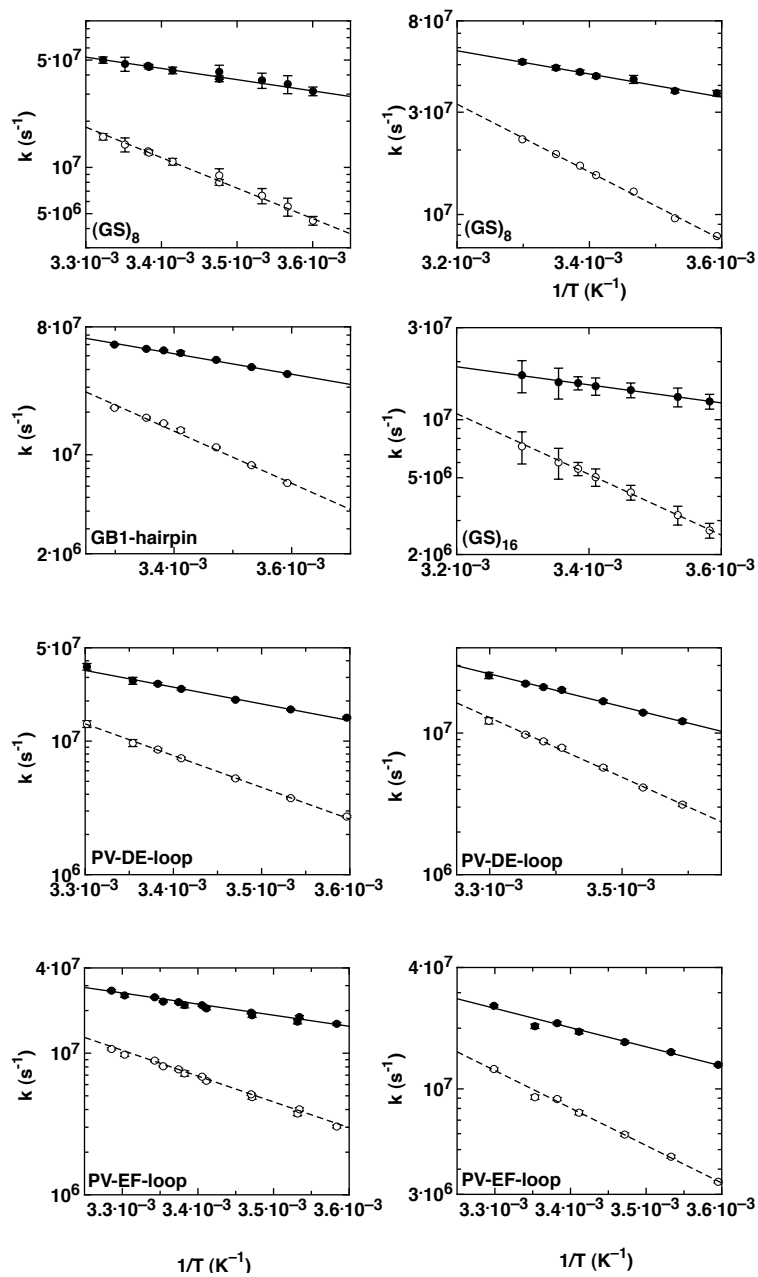
We were further interested, what effect the stabilizing osmolytes TMAO and sarcosine have on the barriers for loop formation. Loop formation dynamics were measured at different temperatures with a fixed concentration of urea or GdmCl, respectively. Loop formation was measured by TTET in the presence of 4 M TMAO or 4 M sarcosine in the temperature range between 5 °C and 30 °C. In order to calculate the rate constant for loop formation, intrinsic triplet decay in donor-only peptides was measured at identical conditions. The rate constant for loop formation was calculated from the apparent rate constant by subtracting the rate constant for the intrinsic triplet decay as described in Chapter 4.2.3. Figure 4.22 shows the calculated and viscosity-corrected rate constants for loop formation as a function of temperature. Viscosity-corrected rate constants were calculated as shown in Chapter 4.1.3. Table 4.9 summarizes the measured activation energies and pre-exponential factors for the different peptides in the presence of 4 M TMAO or 4 M sarcosine. The activation energy and the Arrhenius pre-exponential factor are increased in the presence of stabilizing osmolytes compared to water for both the (GS)-model peptides and the natural sequences (except in the GB1-hairpin.). An activation energy of  $10.4 \pm 0.8$  kJ/mol was found for (GS)<sub>8</sub> in the presence of 4 M sarcosine,  $\sim 6$  kJ/mol higher than that in sole buffer. The pre-exponential factor concomitantly increases from  $(2.4 \pm 1.2) \cdot 10^8$  s<sup>-1</sup> to  $(3.1 \pm 1.1) \cdot 10^9$  s<sup>-1</sup>.

## 4. Results and Discussion



**Figure 4.21.:** Rate constant for loop formation as a function of the transfer free energy from 0-6 M sarcosine. (A-D)  $(GS)_8$ , (E-H)  $(GS)_{16}$ , (I-L) PV-DE-loop, (M-P) PV-EF-loop and (Q-T) GB1-hairpin. Transfer free energy were calculated for the sidechains (SC), backbone (BB) and the sum of BB and SC according to Auton *et al.*<sup>101</sup>.

## 4.2. Effect of Osmolytes and other Co-Solutes on the Dynamics of Loop Formation



**Figure 4.22.:** Effect of temperature on loop formation dynamics in the presence of stabilizing osmolytes. Left: 4 M TMAO, right: 4 M sarcosine. All measurements were performed in 10 mM potassium phosphate at pH 7.

**Table 4.9.:** Arrhenius parameters for the indicated peptide in the presence of stabilizing osmolytes in comparison to water.

Peptide	Water		4 M sarcosine		4 M TMAO	
	$A$ ( $s^{-1}$ )	$E_A$ (kJ/mol)	$A$ ( $s^{-1}$ )	$E_A$ (kJ/mol)	$A$ ( $s^{-1}$ )	$E_A$ (kJ/mol)
(GS) <sub>8</sub>	$(2.4 \pm 1.2) \cdot 10^8$	$4.8 \pm 1.2$	$(3.1 \pm 1.1) \cdot 10^9$	$10.4 \pm 0.8$	$(1.4 \pm 1.1) \cdot 10^{10}$	$14.0 \pm 1.9$
(GS) <sub>16</sub>	$(6.5 \pm 1.7) \cdot 10^7$	$3.9 \pm 0.6$	$(5.9 \pm 9.0) \cdot 10^8$	$9.0 \pm 3.8$	-	-
PV-DE-loop	$(4.6 \pm 1.4) \cdot 10^9$	$14.3 \pm 0.7$	$(1.7 \pm 0.6) \cdot 10^{11}$	$22.2 \pm 0.9$	$(4.6 \pm 2.4) \cdot 10^{11}$	$24.0 \pm 1.3$
PV-EF-loop	$(1.9 \pm 0.3) \cdot 10^9$	$12.5 \pm 0.7$	$(3.5 \pm 0.7) \cdot 10^{10}$	$18.2 \pm 0.5$	$(1.1 \pm 0.4) \cdot 10^{10}$	$15.3 \pm 0.8$
GB1-hairpin	$(1.8 \pm 0.7) \cdot 10^{10}$	$15.7 \pm 0.9$	-	-	$(1.2 \pm 0.1) \cdot 10^{10}$	$13.3 \pm 0.3$

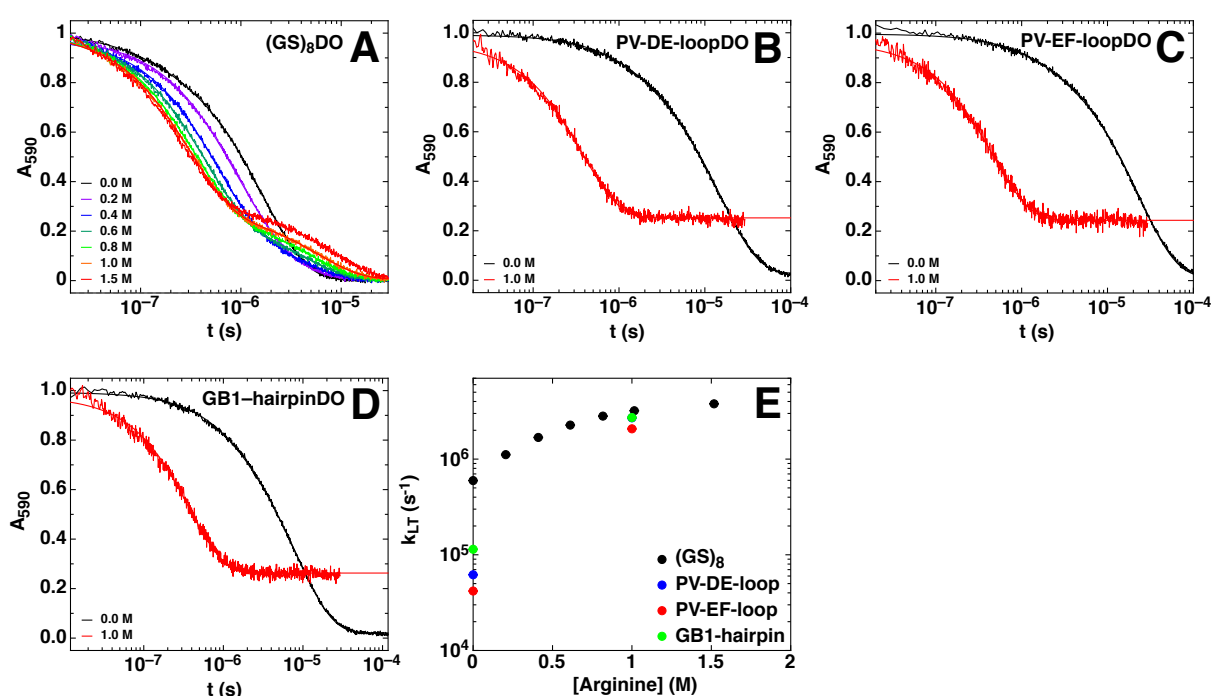
$A$  and  $E_A$  in water for (GS)<sub>1</sub> taken from Reference<sup>127</sup>

#### 4.2.4. Comparison of the Effect of Stabilizing Osmolytes and Co-Solutes on the Dynamics of Loop Formation

In all investigated peptides, viscosity corrected rate constants for loop formation increase with increasing TMAO or sarcosine concentration. The osmolyte specific effect becomes saturated at high concentrations of stabilizing co-solutes and can be described by Schellman's weak binding model. The  $K_{\text{Ex}}$  values are larger for the more potent stabilizing co-solute TMAO in all peptides, in comparison with sarcosine. Both higher activation energies and pre-exponential factors were found in the presence of stabilizers when compared with the values in water. The higher activation energy for loop formation in the presence of stabilizing osmolytes indicates, that more intramolecular enthalpic interactions in the peptide chain are present compared to water, which must be broken in order to be able to form loops. These intramolecular interactions restrict the conformational space. Despite the formation of intramolecular interactions and the restriction of the conformational space, loop formation is faster in stabilizing osmolytes, due to a higher Arrhenius pre-exponential factor than the one obtained in water. Since the effects are approximately the same in the (GS) peptides and the two PV-loops, the intramolecular interactions are believed to be hydrogen bonds, since they are the sole intramolecular interaction possible in the (GS) peptides.

### 4.2.5. Effect of Neutral Osmolytes on the Dynamics of Loop Formation

To investigate the effect of neutral osmolytes on loop formation, chain dynamics were investigated in the presence of proline or arginine. While proline was not found to interact with the triplet transfer reaction, arginine affects the intrinsic triplet lifetime of Xan. Figure 4.23 shows the time based decay in absorbance at 590 nm for the (GS)-donor-only peptide and the donor-only peptides of the natural sequences at the indicated arginine concentration. A dou-



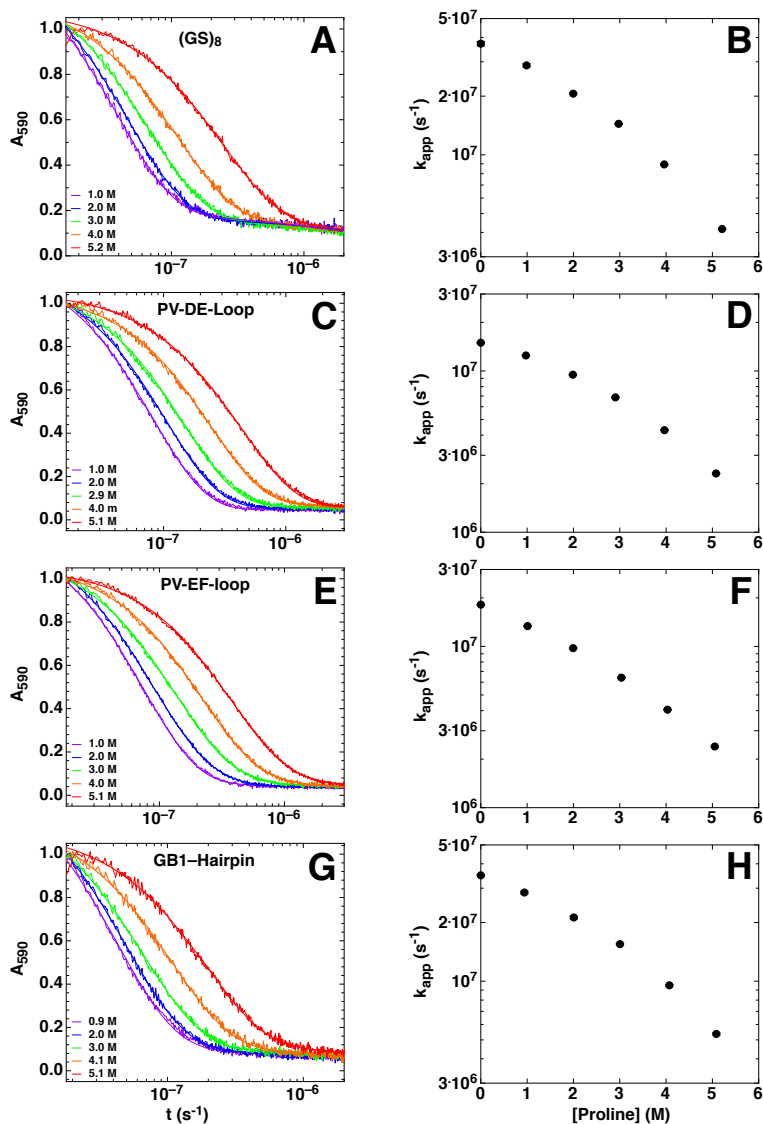
**Figure 4.23.:** Effect of arginine on the Xan triplet state. (A-D) Time based triplet decay in the presence of the indicated arginine concentration. (E) Rate constants for intrinsic triplet decay as a function of the arginine concentration.

ble exponential fit was necessary to describe the Xan triplet decay. The rate constant for the triplet decay is obtained from the main phase and is significantly increased in the presence of arginine in all donor-only peptides.

TTET was measured for the model peptides and the natural sequences bearing donor and acceptor in the presence of proline and arginine. TTET was followed by time-based decay in absorbance at 590 nm. Figure 4.24 shows the effect of proline on the kinetic traces and the rate constants for loop formation. The time based decay traces at 590 nm could be described by



## 4.2. Effect of Osmolytes and other Co-Solutes on the Dynamics of Loop Formation



**Figure 4.24.:** Chain dynamics at various proline concentrations. Left: Time-based absorbance change at 590 nm. Right: Rate constants for loop formation as a function of the proline concentration. All measurements were performed in 10 mM potassium phosphate at pH 7 and 22.5 °C.

#### 4. Results and Discussion

---

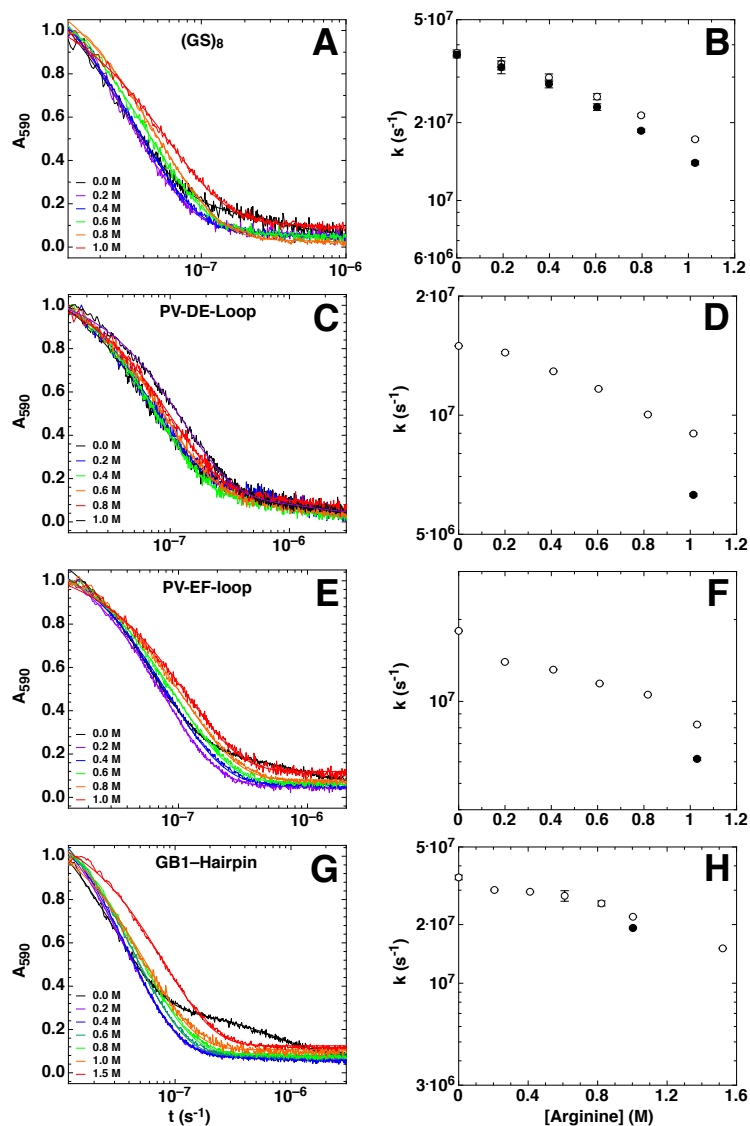
a double exponential fit, directly yielding the rate constant for loop formation from the main phase. Panels B, D F and H of Figure 4.24 show the obtained rate constant for loop formation as a function of the proline concentration. It decreases with increasing proline concentration. At high proline concentrations, the effect on  $k_c$  becomes even more significant.

Results from TTET measurements in the presence of arginine are shown in Figure 4.25. The time base absorbance decay at 590 nm could be described by single exponential kinetics (except for (GS)<sub>8</sub> where a double exponential fit was necessary), yielding an apparent rate constant. Similar to TTET measurements in TMAO and sarcosine, the rate constant for loop formation could be obtained by subtracting the triplet decay rate constant of the donor-only peptide from the apparent rate constant at identical conditions. Apparent ( $\circ$ ) and loop formation ( $\bullet$ ) rate constants are shown in Figure 4.25 on the right. Due to limitations in arginine solubility in aqueous buffer, the effect of arginine on chain dynamics could only be studied for [arginine] up to about 1 M. However, even at these low concentrations, arginine was found to strongly decrease  $k_c$ . For all peptides,  $k_c$  is decreased by a factor of 2-3 from water to 1 M arginine. Arginine thus decreases  $k_c$  even stronger than GdmCl does at identical co-solute concentrations.

The rate constant for loop formation  $k_c$  was corrected for changes in solvent viscosity as described in Chapter 4.1.3, by using Equation 4.7 and the  $\beta$ -values obtained from the viscosity dependence measurement in glycerol. Non-corrected and viscosity-corrected rate constants are summarized in Figure 4.26.  $k_c$  and  $k'_c$  are shown as a function of the molar proline concentration in Panels A and B and as a function of the molar arginine concentration in panels D and E. In panels C and F,  $k'_c$  is shown as a function of the proline or arginine mole fraction, respectively. Proline has little effect on the viscosity-corrected loop formation rate constant. In all peptides except the PV-EF-loop,  $k'_c$  increases and reaches a maximum at around 3 M proline. At higher proline concentrations,  $k'_c$  decreases again. In the PV-EF-loop, the corrected loop formation rate constant decreases with increasing proline concentration and asymptotically approaches a limiting value at high proline concentrations. Only in this peptide, the data could be described by the weak binding model from Schellman (Equation 4.9). The fit yields  $k_c^0 = (1.80 \pm 0.01) \cdot 10^7 \text{ s}^{-1}$ ,  $\gamma = 0.41 \pm 0.04$  and  $K_{\text{Ex}} = 6.5 \pm 1.2$ , the smallest value for the exchange equilibrium constant found for the investigated osmolytes and co-solutes. The effect of arginine could not be described by the weak binding model in any of the investigated peptides.

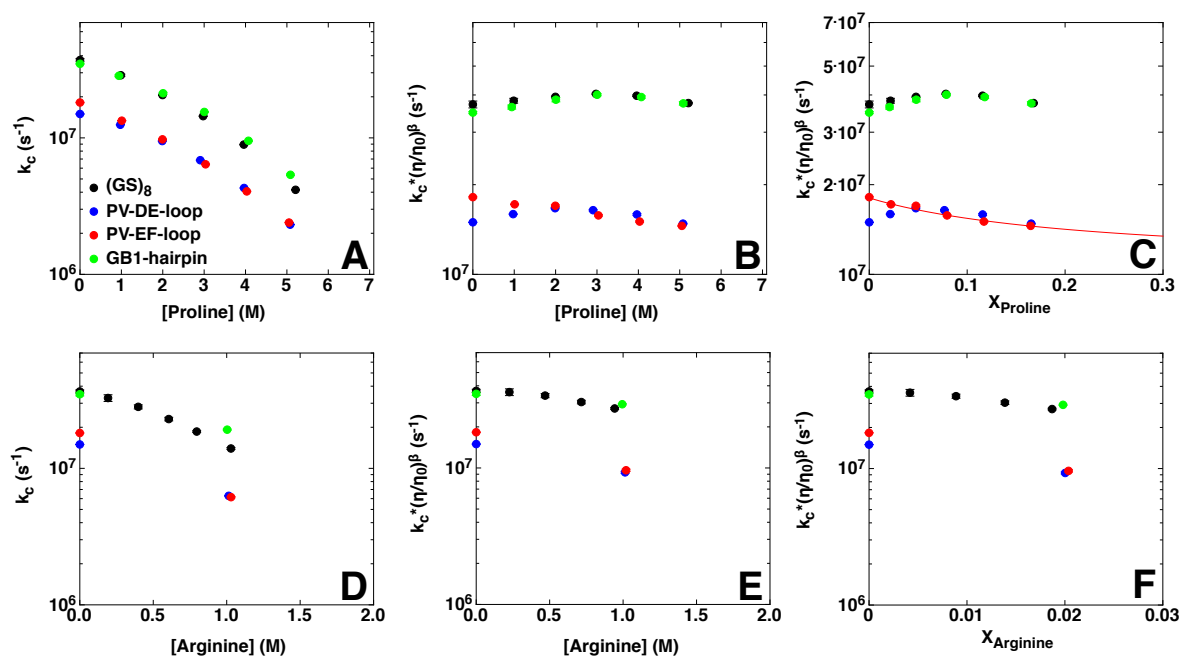
To test, whether the transfer free energy could describe the effect of neutral osmolytes on  $k'_c$ , the free energy of transfer from water to 1 M proline or 1 M arginine was calculated for each

## 4.2. Effect of Osmolytes and other Co-Solutes on the Dynamics of Loop Formation



**Figure 4.25.:** Chain dynamics at various arginine concentrations. Left: Time-based absorbance change at 590 nm. Right: apparent rate constants ( $\circ$ ) and loop formation rate constants ( $\bullet$ ) as a function of the arginine concentration. All measurements were performed in 10 mM potassium phosphate at pH 7 and 22.5 °C.

#### 4. Results and Discussion

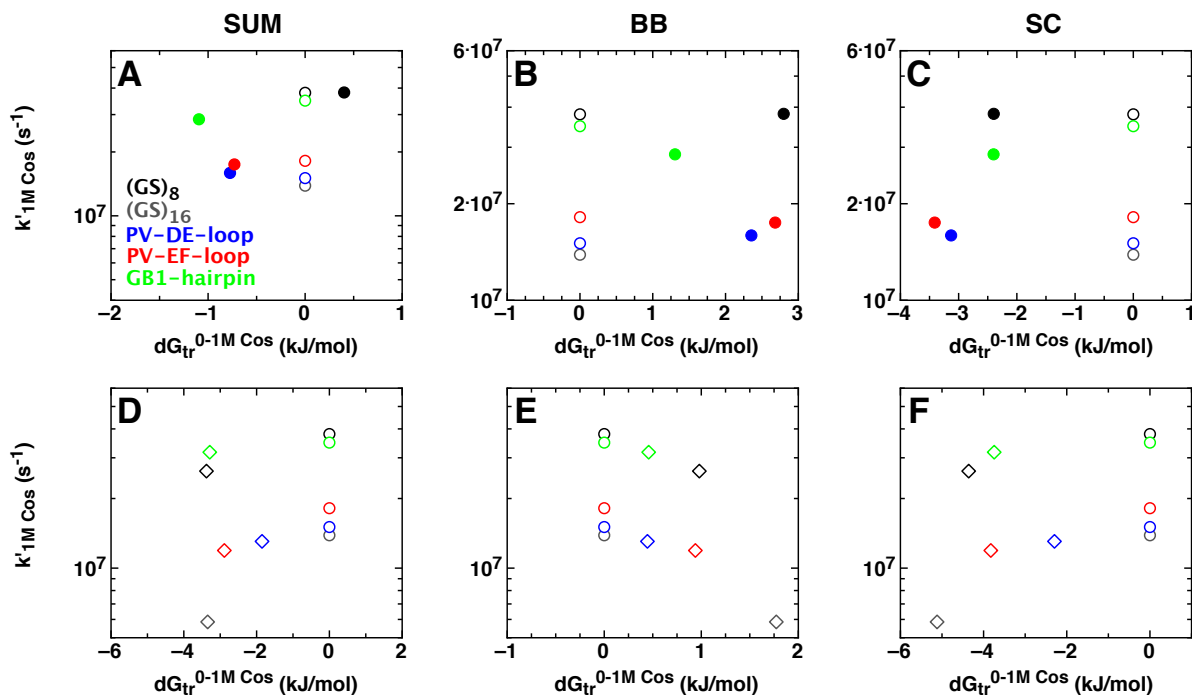


**Figure 4.26.:** Loop formation rate constants in different arginine and proline concentrations. Non-corrected (A, D) and viscosity-corrected (B, E) rate constants for loop formation as a function of the molar concentration and the mole fraction (C, F) of neutral osmolytes. The solid line in C corresponds to a fit of Equation 4.9 to the data.

peptide. This was done according to Chapter 1.6.1, using available GTFE's. Figure 4.27 shows  $k_c'$  as a function of the transfer free energy for each peptide sequence. For proline, small transfer free energies are found for all investigated peptides. The unfavorable transfer free energy of the backbone is compensated by a favorable transfer free energy of the side chains. Accordingly, the viscosity-corrected rate constants for loop formation are almost identical to those measured in water. For proline, Figure 4.27 thus yields the expected result, that osmolytes with a small transfer free energy have only a little effect on the loop formation kinetics. For arginine, the calculated transfer free energies are also small for the investigated peptides. A small unfavorable transfer free energy of the backbone is compensated by a larger favorable transfer free energy of the side chains. Although this results in slightly favorable transfer free energies, a large effect can be observed for  $k_c'$  in arginine. Aside from the transfer free energy, arginine seems to affect loop formation in the investigated peptides in an additional way, probably due to specific interactions of arginine molecules with several amino acid side chains.

The influence of the neutral osmolyte arginine on the barrier for loop formation was investi-

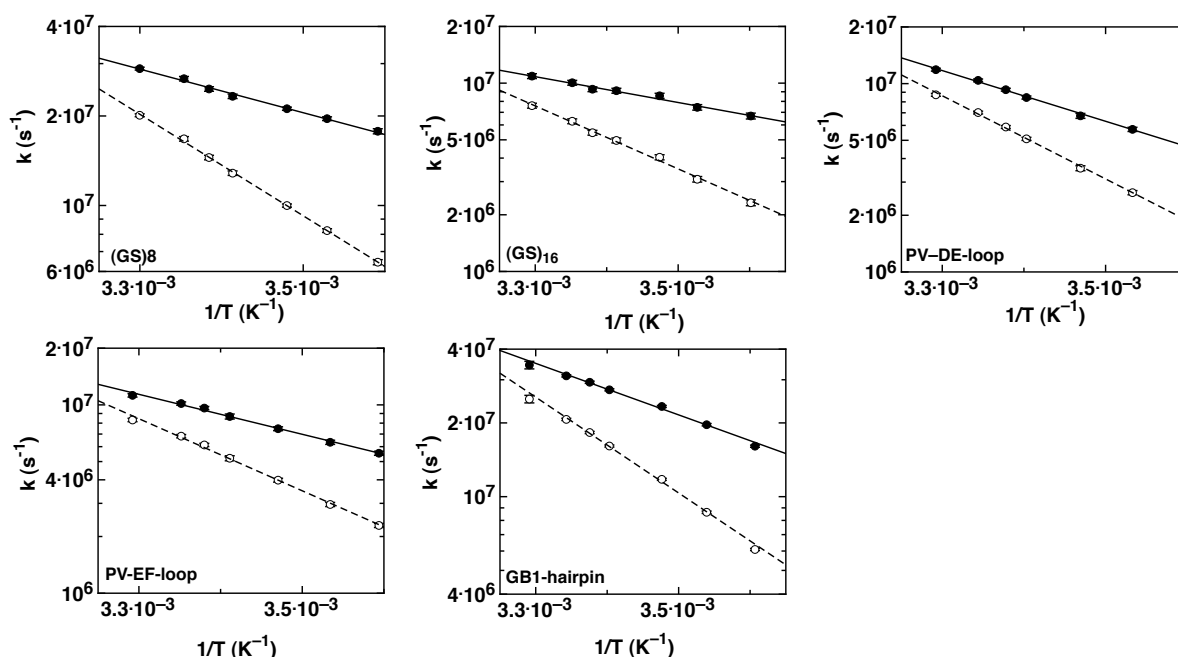
## 4.2. Effect of Osmolytes and other Co-Solutes on the Dynamics of Loop Formation



**Figure 4.27.:** Viscosity-corrected rate constant for loop formation as a function of the transfer free energy from water to (A-C) 1 M proline and (D-F) 1 M arginine in comparison to water ( $\circ$ ).

gated for the model peptides and the natural sequences. In order to calculate the rate constant for loop formation, we measured the rate constant for the intrinsic triplet decay in donor-only peptides in buffer containing 1 M arginine at various temperatures.  $k_c$  was calculated from the apparent rate constant by subtracting the rate constant for the intrinsic triplet decay measured at identical temperatures (see Chapter 4.2.3). Figure 4.28 shows the calculated and viscosity-corrected rate constants for loop formation as a function of temperature. Viscosity-corrected rate constants were calculated as shown in Chapter 4.1.3. Table 4.9 summarizes the activation energies and pre-exponential factors for the different peptides in the presence of 1 M arginine obtained from the Arrhenius Equation to the data. The activation energy and the Arrhenius pre-exponential factor are significantly increased in the presence of 1 M of the neutral osmolyte arginine in all peptides when compared with the values obtained in water. The effect is identical for  $(GS)_8$  and  $(GS)_{16}$ , where the activation energy is increased by 9 kJ/mol and the pre-exponential factor is increased by a factor of 30 compared to water. In the natural sequences, the activation energy is increased by 5-9 kJ/mol. An even larger effect on  $A$  was observed for the natural sequences compared to the  $(GS)$ -peptides. Amongst the natural sequences, the strongest effect for  $A$  and  $E_A$  was observed for the PV-DE-loop.

## 4. Results and Discussion



**Figure 4.28.:** Effect of temperature on loop formation dynamics in the presence of the neutral osmolyte arginine. All measurements were performed in 10 mM potassium phosphate, 1 M arginine at pH 7.

### 4.2.6. Comparison of the Effect of Neutral Osmolytes on the Dynamics of Loop Formation

The neutral osmolytes arginine and proline were expected to have only a small influence on loop formation dynamics, since the calculated transfer free energy for all investigated peptides from water to 1 M of the respective osmolyte is small. Proline indeed had only a negligible effect on loop formation dynamics with values for  $k'_c$  similar to those in water. The effect of proline on  $k'_c$  could not be described by Schellman's weak binding model, except for the PV-EF-loop. In the presence of arginine, an unexpectedly strong deceleration in loop formation was found, with the magnitude of the effect depending on the peptide's amino acid sequence. This effect cannot be described by the transfer free energy but might arise from specific binding of arginine to several amino acid side chains. Additionally, the presence of arginine leads to a higher Arrhenius pre-exponential factor and a higher activation energy associated with loop formation. Arginine thus leads to the restriction of the conformational space, due to interaction with the side chains of polypeptides.

## 4.2. Effect of Osmolytes and other Co-Solutes on the Dynamics of Loop Formation

**Table 4.10.:** Arrhenius parameters for the indicated peptide under different solvent compositions.

Peptide	Water		1 M Arginine	
	$A$ ( $s^{-1}$ )	$E_A$ (kJ/mol)	$A$ ( $s^{-1}$ )	$E_A$ (kJ/mol)
(GS) <sub>8</sub>	$(2.4 \pm 1.2) \cdot 10^8$	$4.8 \pm 1.2$	$(7.5 \pm 1.6) \cdot 10^9$	$14.0 \pm 0.5$
(GS) <sub>16</sub>	$(6.5 \pm 1.7) \cdot 10^7$	$3.9 \pm 0.6$	$(2.0 \pm 0.3) \cdot 10^9$	$13.1 \pm 0.4$
PV-DE-loop	$(4.6 \pm 1.4) \cdot 10^9$	$14.3 \pm 0.7$	$(3.2 \pm 1.2) \cdot 10^{11}$	$25.7 \pm 0.9$
PV-EF-loop	$(1.9 \pm 0.3) \cdot 10^9$	$12.5 \pm 0.7$	$(3.6 \pm 1.2) \cdot 10^{10}$	$20.3 \pm 0.8$
GB1-hairpin	$(1.8 \pm 0.7) \cdot 10^{10}$	$15.7 \pm 0.9$	$(1.0 \pm 0.2) \cdot 10^{11}$	$20.1 \pm 0.6$

$A$  and  $E_A$  in water for (GS)<sub>1</sub> taken from Reference<sup>127</sup>

### 4.2.7. Summary of the Effect of Osmolytes and Co-Solutes on the Dynamics of Loop Formation

This chapter summarizes the effect of destabilizing, stabilizing and neutral osmolytes and co-solutes on loop formation dynamics and barriers, obtained for the investigated peptides. The data is described with Schellman's weak binding model and can be correlated to the peptides transfer free energy calculated with Tanford's Transfer Model.

#### 4.2.7.1. Schellman's Weak Binding Model

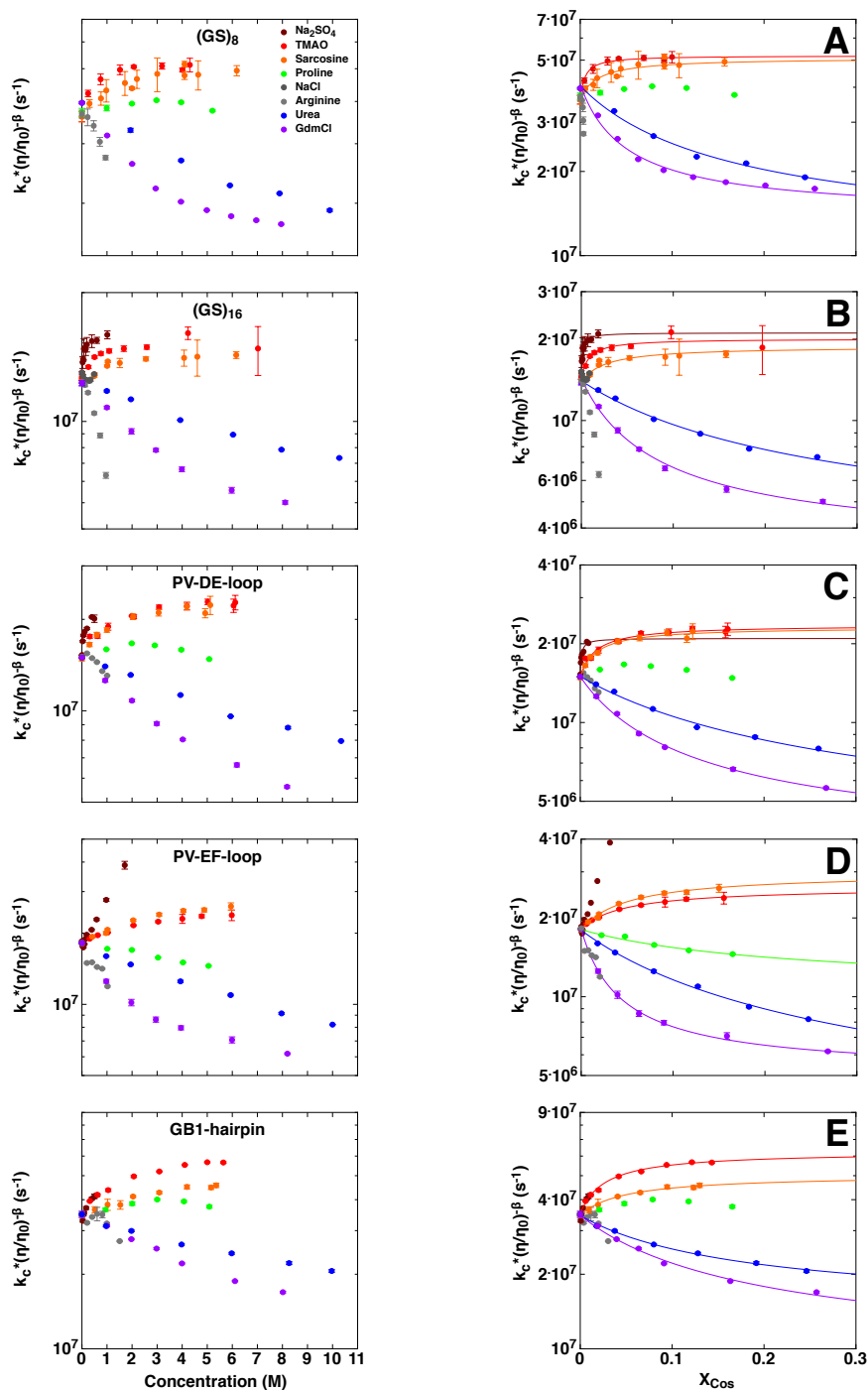
Viscosity-corrected rate constants for loop formation are summarized for all peptides in the presence of the different osmolytes in Figure 4.29. While the viscosity-corrected rate constants for loop formation  $k'_c$  become slower in the presence of destabilizing co-solutes, they increase with increasing concentrations of protecting co-solutes and are almost unaffected in the presence of the neutral co-solute proline. For the model peptides as well as the natural sequences,  $k'_c$  asymptotically approaches a limiting value at high TMAO, sarcosine, urea and GdmCl concentrations. This indicates, that the co-solute specific effect becomes saturated at high co-solute concentrations. The effect of all co-solutes on  $k_c$  can be explained by the sum of the co-solute specific effect and an increase in solvent viscosity, while the latter effect dominates  $k_c$  at high co-solute concentrations, where the co-solute specific effect is already saturated.

When plotting the viscosity-corrected rate constant as a function of the mole fraction concentration scale, the Schellman weak binding model can be used to describe the data in aqueous buffers containing denaturants but also in the presence of stabilizing osmolytes and for the neutral osmolyte proline in the PV-EF-loop. Values of  $k_c$  corrected for changes in solvent viscosity are shown as a function of the co-solute molar (Figure 4.29 left) and mole fraction (right) concentration scale for all investigated peptides. The solid lines represent a fit of the weak binding model (Equation 4.9) to the measured data. The parameters  $\gamma$  and  $K_{Ex}$  from this model were shown before and are summarized here for all investigated osmolytes and co-solutes (Table 4.11). Positive  $\gamma$ -values are found for denaturants, which interact favorably with the polypeptide chain and thus lead to a decrease of  $k'_c$  with increasing denaturant concentration. Negative  $\gamma$ -values can be found however for the stabilizing osmolytes TMAO and sarcosine, which lead to an increase of  $k'_c$  with increasing concentration.

The effect of the co-solutes  $\text{Na}_2\text{SO}_4$  (for the PV-EF-loop and the GB1-hairpin), proline (for



## 4.2. Effect of Osmolytes and other Co-Solutes on the Dynamics of Loop Formation



**Figure 4.29.:** Viscosity-corrected rate constant for loop formation as a function of the co-solute concentration. Left:  $k_c^*$  as a function of the molar co-solute concentration. Right:  $k_c^*$  as a function of the co-solute mole fraction  $X_{\text{Cos}}$ . The solid lines represent a fit of Equation 4.9 to the data.

#### 4. Results and Discussion

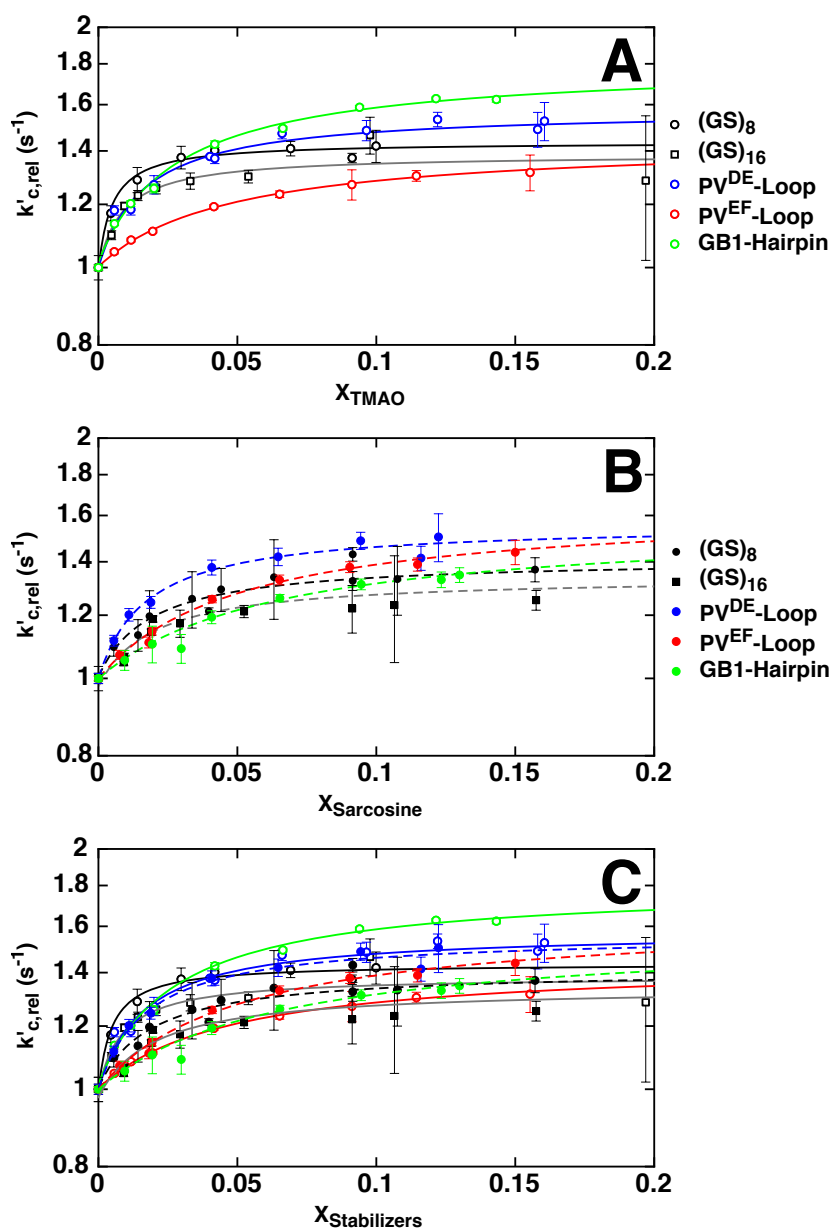
**Table 4.11.:** Parameters for Schellman's weak binding model

Peptide	GdmCl		urea	
	$\gamma$	$K_{\text{Ex}}$	$\gamma$	$K_{\text{Ex}}$
(Gly-Ser) <sub>8</sub>	$0.66 \pm 0.01$	$33.8 \pm 0.9$	$0.69 \pm 0.01$	$13.3 \pm 0.6$
(Gly-Ser) <sub>16</sub>	$0.77 \pm 0.01$	$23.7 \pm 2.0$	$0.78 \pm 0.06$	$7.7 \pm 1.3$
PV <sup>DE</sup> -loop	$0.78 \pm 0.01$	$16.2 \pm 0.7$	$0.78 \pm 0.01$	$7.1 \pm 0.2$
PV <sup>EF</sup> -loop	$0.72 \pm 0.01$	$41.8 \pm 2.8$	$0.83 \pm 0.02$	$8.6 \pm 0.6$
GB1-hairpin	$0.74 \pm 0.02$	$10.4 \pm 0.5$	$0.59 \pm 0.03$	$9.4 \pm 0.9$
	Sarcosine		TMAO	
	$\gamma$	$K_{\text{Ex}}$	$\gamma$	$K_{\text{Ex}}$
(Gly-Ser) <sub>8</sub>	$-0.40 \pm 0.06$	$45.2 \pm 25.0$	$-0.45 \pm 0.05$	$180 \pm 55$
(Gly-Ser) <sub>16</sub>	$-0.38 \pm 0.04$	$45.6 \pm 10.3$	$-0.39 \pm 0.03$	$100 \pm 21$
PV <sup>DE</sup> -loop	$-0.56 \pm 0.04$	$47.4 \pm 9.9$	$-0.58 \pm 0.03$	$49.3 \pm 7.3$
PV <sup>EF</sup> -loop	$-0.65 \pm 0.06$	$16.1 \pm 2.9$	$-0.45 \pm 0.02$	$17.9 \pm 2.2$
GB1-hairpin	$-0.45 \pm 0.06$	$18.1 \pm 4.9$	$-0.77 \pm 0.02$	$28.0 \pm 2.3$

the (GS)<sub>8</sub>, PV-DE-loop and GB1-hairpin), NaCl (all peptides) and arginine (all peptides) on  $k'_c$  could not be described by the weak binding model. In the presence of arginine, no limiting value for  $k'_c$  was reached. This was also the case for the PV-EF-loop in the presence of Na<sub>2</sub>SO<sub>4</sub>. In the presence of NaCl ((GS)<sub>8</sub>) and Na<sub>2</sub>SO<sub>4</sub> (GB1-hairpin),  $k'_c$  decreases at low co-solute concentration but increases at high co-solute concentration. This behavior is expected to be due to a salt effect of these compounds, since charge interactions were found to dominate the dimensions of unstructured polypeptides<sup>186</sup> and thus might also affect end-to-end loop formation.

In order to compare the strength of the effect, stabilizers exert on  $k'_c$  between the different peptides, relative  $k'_c$ -values were calculated. They were obtained by dividing  $k'_c$  at each co-solute concentration by the value for  $k_c$  of the same peptide in sole buffer. Figure 4.30 shows the relative effect of the co-solutes on the loop formation rate constant for the different peptides. It can be seen, that the relative effect of the different co-solutes is approximately identical in all peptides. This indicates, that the amino acid sequence has only a marginal influence on end-to-end loop formation in the presence of TMAO and sarcosine. Therefore, it seems evident, that the stabilizers interact mainly with the backbone of the polypeptide chains. They lead to a  $\sim 40\%$  increase in  $k'_c(\text{rel})$  in the high co-solute concentration limit with TMAO being slightly more efficient than sarcosine in speeding up the rate constant for loop formation already at lower concentrations. Generally, larger  $K_{\text{Ex}}$  values are found in TMAO compared

## 4.2. Effect of Osmolytes and other Co-Solutes on the Dynamics of Loop Formation



**Figure 4.30.:** Comparison of the relative rate constants for loop formation  $k'_{c,rel}$  in the presence of stabilizing co-solutes. (A) TMAO, (B) sarcosine and (C) TMAO (circles) and sarcosine (bullets).

## 4. Results and Discussion

---

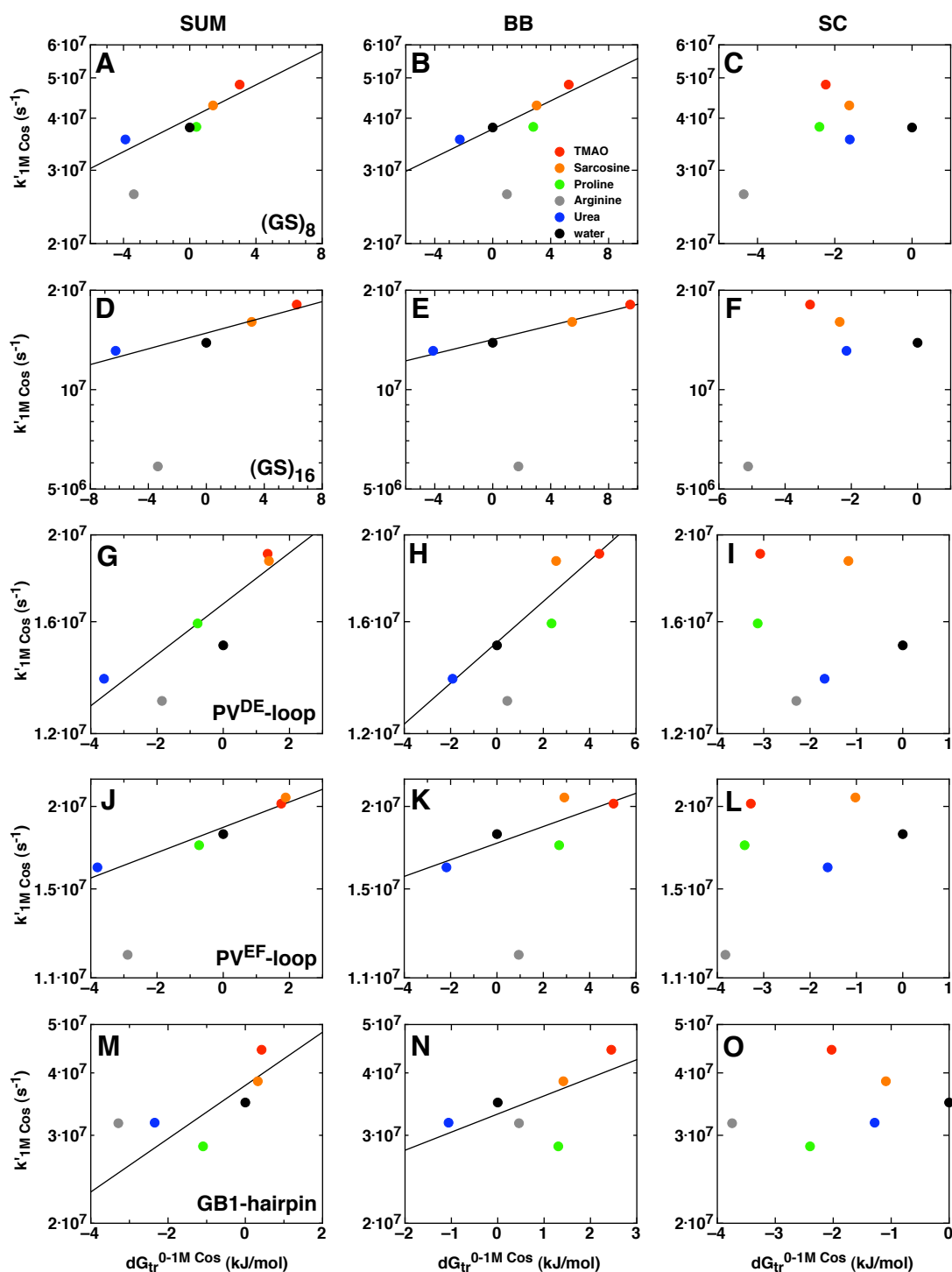
to those found in sarcosine. Accordingly, larger values for  $K_{\text{Ex}}$  are found for the more efficient denaturant GdmCl compared with those in urea. This can be seen most prominently when comparing the  $(\text{GS})_n$  peptides in sarcosine and TMAO. Since these peptides represent the bare backbone, the isolated effect of the osmolytes on the backbone is observed, where the stronger effect of TMAO on  $k'_c$  is most apparent and agrees with TMAO being found to be the most effective organic osmolyte<sup>101</sup>.

### 4.2.7.2. Using Tanford's Transfer Model to Describe the Effect of Osmolytes and other Co-Solutes on Loop Formation

Figure 4.31 shows the rate constant for loop formation at 1 M co-solute ( $k'_{1\text{M Cos}}$ ) as a function of the polypeptides transfer free energy from water to 1 M osmolyte or co-solute. The transfer free energy from water to the respective co-solute is indicated by the different colors. In the panels on the left,  $k'_{1\text{M Cos}}$  is plotted as a function of the transfer free energy of the backbone (BB), the sidechains (SC) and the sum of the two. The effect of the BB and the SC alone on  $k'_{1\text{M Cos}}$  are shown in the middle and right panels of Figure 4.31, respectively.

A direct correlation between  $\log k'_{1\text{M Cos}}$  and the sum of the transfer free energy  $\Delta G_{U;0 \rightarrow 1\text{M}}^{0;tr;sum}$  can be found for the investigated peptides, except for arginine, whose  $\log k'_{1\text{M Cos}}$  are generally lower than would be expected. A correlation can be also found for  $\log k'_{1\text{M Cos}}$  with the BB transfer free energy. This correlation is most distinct for the  $(\text{GS})_n$  peptides, which represent the backbone and where the effect of SC is absent compared to the natural sequences with their large and complex SC. In all peptides but the short GB1-hairpin, the main contribution to  $\Delta G_{U;0 \rightarrow 1\text{M}}^{0;tr;sum}$  for TMAO, sarcosine, GdmCl and urea arises from the backbone transfer free energy. In contrast to the BB, no correlation for  $\log k'_{1\text{M Cos}}$  with the SC transfer free energy was found. As a consequence, in arginine and the GB1-hairpin, where the side chain transfer free energies are dominating, the correlation with  $\Delta G_{U;0 \rightarrow 1\text{M}}^{0;tr;sum}$  is ill defined. Arginine might bind or tightly interact with the amino acid SC, which results in a strong deceleration end-to-end loop formation. In the other cases however, Tanford's Transfer Model is well suited to describe the effect of the different co-solutes on the rate constant for end-to-end loop formation, where stabilizers lead to faster kinetics and denaturants to slower kinetics, respectively.

## 4.2. Effect of Osmolytes and other Co-Solutes on the Dynamics of Loop Formation



**Figure 4.31.:** Viscosity-corrected rate constant for loop formation  $k'_c$  as a function of transfer free energy  $\Delta G_{U:0 \rightarrow 1M}^{0:tr}$  from water to 1 M co-solute. (A-C)  $(GS)_8$ , (D-F)  $(GS)_{16}$ , (G-I) PV-DE-loop, (J-L) PV-EF-loop and (M-O) GB1-hairpin. SUM: sum of backbone (BB) and sidechain (SC) transfer free energies.

##### 4.2.7.3. Summary of the Effect of Osmolytes and Co-Solutes on the Barriers for Loop Formation

In order to unravel in detail, what influences barriers for intrachain loop formation and get insight into their origin, TTET was applied to measure end-to-end loop formation in peptides with different length and amino acid sequence at varying temperatures. Further, the influence of the solvent quality on the barrier height was tested for the various peptides, in order to understand, how solvent quality influences the properties of peptides with different length amino acid composition. Solvent quality was modulated by the addition of GdmCl, urea, sarcosine, TMAO, arginine and Na<sub>2</sub>SO<sub>4</sub>, thereby rendering a good or poor solvent for the polypeptide chain. The influence of temperature on loop formation was measured for the three natural sequences (Figure 4.6) as well as (GS)<sub>8</sub> and (GS)<sub>16</sub>. Table 4.12 summarizes the measured activation energies and pre-exponential factors for the different peptides for all investigated conditions. For better visualization, the values are represented graphically in Figure 4.32.

$E_A$  is shown in red (left scale) and  $A$  in blue (right scale). At all conditions, the natural sequences exhibit higher values for  $E_A$  and  $A$  than (GS)<sub>8</sub> and (GS)<sub>16</sub>. Denaturants lead to a slight increase in both  $A$  and  $E_A$  in all peptides when compared to the values in water. The increase is stronger in the presence of protectants, and analogue to the findings in the (GS) peptides. The higher barrier and pre-exponential factor in the presence of protectants indicates, that more enthalpic interactions must be broken in order to form end-to-end loop formation. Thus, a larger number of enthalpic intramolecular interactions exists in poor solvents for the polypeptide chain. Since the effects are approximately the same in the (GS) peptides and the two PV-loops, these interactions are believed to be hydrogen bonds, since they are the sole intramolecular interaction possible in the (GS) peptides. In the presence of protectants, they are more favorable than peptide-solvent hydrogen bonds and therefore lead to a more compact chain with a higher barrier for end-to-end loop formation. This is in good agreement with earlier findings<sup>36</sup> and with faster rate constants for loop formation at high protectant concentration, where the effect can be addressed to a compactness of the polypeptide chain. This is in agreement with higher values for  $A$  in poor solvents compared to those in water.

It was shown above, that loop formation in short peptides is not solely diffusion controlled. For these peptides,  $\beta$ -values  $\neq -1$  are obtained. Concomitantly, high activation barriers for loop formation were found for these peptides. This prompted us to plot  $E_A$  as a function of  $\beta$ . A linear correlation between the two values could be found for (GS)<sub>n</sub>- and (S)<sub>n</sub>-peptides of

4.2. Effect of Osmolytes and other Co-Solutes on the Dynamics of Loop Formation

**Table 4.12.:** Arrhenius parameters for the indicated peptide under different solvent compositions.

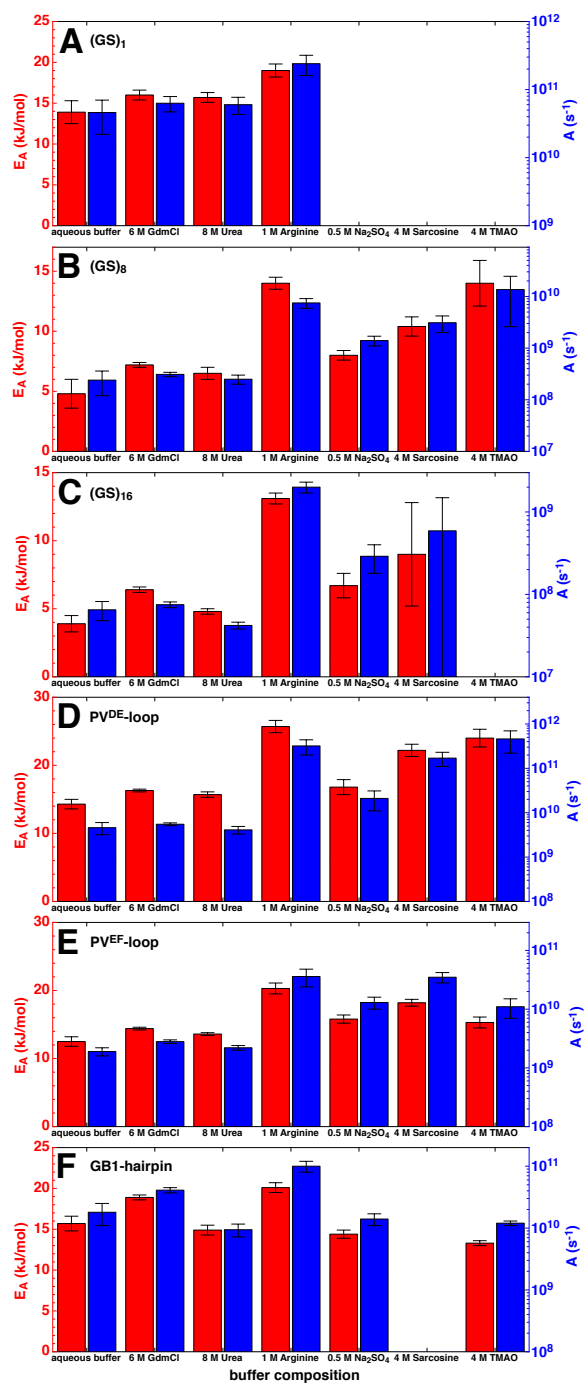
Peptide	Water			6 M GdmCl			8 M urea		
	$A$ ( $s^{-1}$ )	$E_A$ (kJ/mol)	$A$ ( $s^{-1}$ )	$E_A$ (kJ/mol)	$A$ ( $s^{-1}$ )	$E_A$ (kJ/mol)	$A$ ( $s^{-1}$ )	$E_A$ (kJ/mol)	
(GS) <sub>1</sub>	$(4.1 \pm 2.4) \cdot 10^{10}$	$13.6 \pm 1.4$	$(6.3 \pm 1.6) \cdot 10^{10}$	$16.0 \pm 0.6$	$(6.0 \pm 1.7) \cdot 10^{10}$	$15.7 \pm 0.6$			
(GS) <sub>8</sub>	$(2.4 \pm 1.2) \cdot 10^8$	$4.8 \pm 1.2$	$(3.1 \pm 0.3) \cdot 10^8$	$7.2 \pm 0.2$	$(2.5 \pm 0.5) \cdot 10^8$	$6.5 \pm 0.5$			
(GS) <sub>16</sub>	$(6.5 \pm 1.7) \cdot 10^7$	$3.9 \pm 0.6$	$(7.5 \pm 0.6) \cdot 10^7$	$6.4 \pm 0.2$	$(4.2 \pm 0.4) \cdot 10^7$	$4.8 \pm 0.2$			
PV-DE-loop	$(4.6 \pm 1.4) \cdot 10^9$	$14.3 \pm 0.7$	$(5.5 \pm 0.4) \cdot 10^9$	$16.3 \pm 0.2$	$(4.1 \pm 0.8) \cdot 10^9$	$15.7 \pm 0.4$			
PV-EF-loop	$(1.9 \pm 0.3) \cdot 10^9$	$12.5 \pm 0.7$	$(2.8 \pm 0.2) \cdot 10^9$	$14.4 \pm 0.2$	$(2.2 \pm 0.2) \cdot 10^9$	$13.6 \pm 0.2$			
GB1-hairpin	$(1.8 \pm 0.7) \cdot 10^{10}$	$15.7 \pm 0.9$	$(4.1 \pm 0.4) \cdot 10^{10}$	$18.9 \pm 0.3$	$(9.4 \pm 2.2) \cdot 10^9$	$14.9 \pm 0.6$			

Peptide	0.5 M Na <sub>2</sub> SO <sub>4</sub>			4 M sarcosine			4 M TMAO		
	$A$ ( $s^{-1}$ )	$E_A$ (kJ/mol)	$A$ ( $s^{-1}$ )	$E_A$ (kJ/mol)	$A$ ( $s^{-1}$ )	$E_A$ (kJ/mol)	$A$ ( $s^{-1}$ )	$E_A$ (kJ/mol)	
(GS) <sub>1</sub>	-	-	-	-	-	-	-	-	
(GS) <sub>8</sub>	$(1.4 \pm 0.3) \cdot 10^9$	$8.0 \pm 0.4$	$(3.1 \pm 1.1) \cdot 10^9$	$10.4 \pm 0.8$	$(1.4 \pm 1.1) \cdot 10^{10}$	$14.0 \pm 1.9$			
(GS) <sub>16</sub>	$(2.9 \pm 1.1) \cdot 10^8$	$6.7 \pm 0.9$	$(5.9 \pm 9.0) \cdot 10^8$	$9.0 \pm 3.8$	-	-			
PV-DE-loop	$(2.1 \pm 1.0) \cdot 10^{10}$	$16.8 \pm 1.1$	$(1.7 \pm 0.6) \cdot 10^{11}$	$22.2 \pm 0.9$	$(4.6 \pm 2.4) \cdot 10^{11}$	$24.0 \pm 1.3$			
PV-EF-loop	$(1.3 \pm 0.3) \cdot 10^{10}$	$15.8 \pm 0.6$	$(3.5 \pm 0.7) \cdot 10^{10}$	$18.2 \pm 0.5$	$(1.1 \pm 0.4) \cdot 10^{10}$	$15.3 \pm 0.8$			
GB1-hairpin	$(1.4 \pm 0.3) \cdot 10^{10}$	$14.4 \pm 0.5$	-	-	$(1.2 \pm 0.1) \cdot 10^{10}$	$13.3 \pm 0.3$			

$A$  and  $E_A$  in water for (GS)<sub>1</sub>, PV-DE-loop, PV-EF-loop and GB1-hairpin taken from Reference <sup>127</sup>

## 4. Results and Discussion

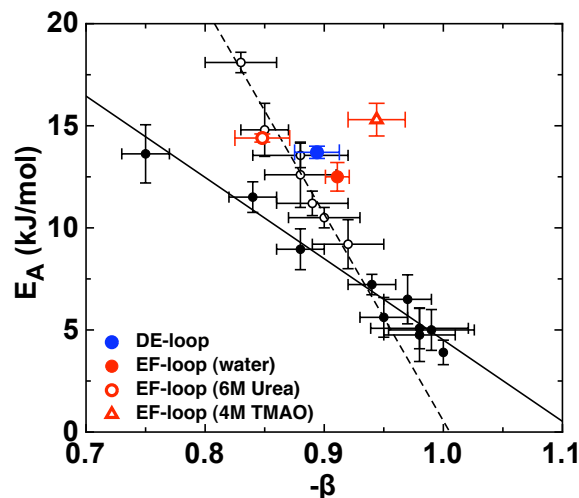


**Figure 4.32.:** Summarized Arrhenius parameters for loop formation in the presence of the indicated co-solute. Data for (GS)<sub>1</sub> (A), (GS)<sub>8</sub> (B), (GS)<sub>16</sub> (C), PV<sup>DE</sup>-loop (D), PV<sup>EF</sup>-loop (E) and GB1-hairpin (F). Activation energies are colored in red, Arrhenius pre-exponential factors in blue. All measurements were performed in 10 mM potassium phosphate buffer at pH 7 in the temperature range between 5 °C and 30 °C.



## 4.2. Effect of Osmolytes and other Co-Solutes on the Dynamics of Loop Formation

different length (Figure 4.33). In the stiffer  $(S)_n$ -peptides,  $E_A$  increases faster with increasing



**Figure 4.33.:** Activation energy  $E_A$  for loop formation as a function of the solvent viscosity sensitivity  $\beta$ . Data for  $(GS)_n$  ( $\bullet$ ) and  $(S)_n$  ( $\circ$ ) taken from References<sup>127,130</sup>.

$\beta$  than in  $(GS)_n$ . Long  $(GS)_n$  peptides with  $n \geq 8$ , which behave like statistical chains, cluster around  $\beta$ -values of -1 and activation energies of 4-5 kJ/mol. Short chains do not behave like statistical chains, since enthalpic contributions have a significant influence for the loop formation reaction. Loop formation in these peptides is in contrast to long chains not predominantly controlled by entropy. Steric or enthalpic effects are assumed to be the reason for the high activation energy in chains with a large  $A$ , which indicates a compact conformation.

The natural sequences exhibit higher activation energies for end-to-end loop formation in water than that expected for poly- $(GS)$  chains with identical  $\beta$ . The correlation of  $E_A$  with  $\beta$  for the PV-DE-loop and the PV-EF-loop in water and for the PV-EF-loop in 6 M urea is comparable to that of the  $(S)_n$  peptides. In the presence of 4 M TMAO, the difference of the activation energy of the PV-EF-loop and the poly-Ser with identical  $\beta$  is significant. The  $\beta$  value of the PV-EF-loop increases with decreasing solvent quality. This feature might be unexpected on the first glance. A possible explanation for this observation is, that the polypeptide chain is more sensitive to the viscous effect of glycerol, since TMAO is excluded from the peptide vicinity. Under these conditions, the polypeptide chain interacts mainly with the viscosifier. In the presence of urea, the peptide experiences less of the macroscopic solvent viscosity, since urea is expected to bind weakly to the polypeptide and thereby occupying interaction sites of the peptide with the viscosifier.

### 4.2.7.4. Effect of Solvent Quality on the Dynamics of Loop Formation

In good solvents like urea, the end-to-end loop formation rate constant is lower than in water, while  $k'_c$  becomes higher in poor solvents. The kinetics thus depend on solvent quality. We assume the polypeptide chain in poor solvents to be more compact due to the formation of intramolecular interactions at the expense of peptide-solvent interactions in agreement with earlier findings<sup>187</sup>. The more compact structure is responsible for the faster end-to-end loop formation. In good solvents, the situation is inverse. The polypeptide chain is less compact, since peptide-solvent interactions are formed at the expense of intramolecular interactions within the polypeptide and the rate constant for end-to-end loop formation is lower. The question remains, which intramolecular interactions a polypeptide can form. To answer it, we measured chain dynamics in (GS)-peptides mimicking the bare backbone. In these peptides, only hydrogen bonds between backbone amide and backbone carbonyl groups or between backbone carbonyl groups and the serine side chain can be formed. We thus believe that in poor solvents additional hydrogen bonds are formed (compared to water) while intramolecular hydrogen bonds are broken in good solvents, which was also found earlier<sup>36</sup>. Since the effects on co-solutes are alike for the backbone mimicking peptide and the natural sequences, a similar effect is expected to happen in the natural sequences. Modulation of the solvent quality by co-solutes thus influences chain dynamics and chain compactness as intramolecular interactions are formed or broken.

The effects osmolytes and several co-solutes exerted on polypeptides and proteins are mainly due to interactions with the peptide backbone and are thus independent of the protein's amino acid sequence. In nature, osmolytes have to act on all proteins within a cell, independent of their sequence. This can only be achieved, if osmolytes mainly affect the peptide backbone. Our results agree with the desired universal way of action of osmolytes in organisms.

### 4.3. Contributions of Internal Friction and Solvent Friction to the Dynamics of Loop Formation in Unfolded Polypeptide Chains

In Kramers theory, a chemical reaction can be described by the crossing of a hypothetical particle over a barrier of height  $E$ , controlled by Brownian motion<sup>76</sup>. The particle arrives at the final state with the rate constant  $k$ , given by

$$k = \frac{\omega_0 \omega_{\ddagger}}{2\pi \cdot \gamma} \cdot \exp\left(-\frac{E}{RT}\right) \quad (4.11)$$

The rate constant  $k$  is subject to the temperature  $T$  and the friction coefficient  $\gamma$ , which accounts for solvent damping of the rate constant. When  $\gamma$  is assumed to be proportional to the solvent dynamic viscosity  $\eta_{\text{solv}}$ ,  $k$  then depends on  $\eta_{\text{solv}}^{-1}$ . Equation 4.11 predicts that  $k$  would become infinite for  $\eta_{\text{solv}} \rightarrow 0$ . This appears implausible, since dissipative interactions within the polypeptide chain are expected to take control of the dynamics<sup>77</sup>. Intrachain interactions or internal friction effects<sup>78</sup> will eventually set an upper limit for  $k$ , leading to a deviation from the proportionality of  $k$  with  $\eta_{\text{solv}}^{-1}$  at low solvent viscosities. De Gennes proposed, that the relaxation time  $\tau$  of a polymer towards its equilibrium could be expressed as the sum of the solvent-controlled relaxation time  $\tau_{\text{solv}}$  ( $\tau_{\text{solv}} \propto \eta$ ) and a solvent-independent timescale  $\tau_{\text{int}}$ <sup>78</sup>.

$$\frac{1}{k} = \tau = \tau_{\text{int}} + \tau_{\text{solv}} \cdot \left(\frac{\eta}{\eta_0}\right) \quad (4.12)$$

It can be seen from Equation 4.12, that an eventually existing term for the internal friction  $\tau_{\text{int}}$  can be found by extrapolation of  $\tau$  for  $\eta \rightarrow 0$  from  $\tau$ . This formalism was used by several groups to determine  $\tau_{\text{int}}$  for protein folding reactions<sup>77,79–83,86,188–191</sup>. A more common way to write Equation 4.12 can be achieved by writing it as a power law according to Equation 4.13.

$$\frac{1}{k} = \tau_{\text{int}} + \tau_{\text{solv}} \cdot \left(\frac{\eta}{\eta_0}\right)^{-\beta} \quad (4.13)$$

This power law deviates from Equation 4.12 by the exponent  $\beta$ , which causes a downward curvature in case  $\beta \neq -1$  when plotting  $1/k$  as a function of  $\eta$ . The curvature is more pronounced, the stronger  $\beta$  deviates from -1. For  $\beta = -1$ ,  $1/k_c$  depends linearly on  $\eta_{\text{solv}}$  and Equation 4.13 simplifies to Equation 4.12. For protein folding reactions,  $\beta$  is unknown and a value of -1 is

#### 4. Results and Discussion

---

used often. For loop formation reactions, the  $\beta$ -values for each peptide can be determined by measuring loop formation at different solvent viscosities. In two-state systems, viscosifiers might affect the solvent quality and thus change the free energy surface by stabilizing one state relative to the other. This makes it difficult to separate thermodynamic from dynamic effects. In order to obtain solely the dynamical effect, changes in the free energy surface evoked by the viscogenic agent need to be compensated by the addition of stabilizing or denaturing co-solutes<sup>189</sup>. However in the case of loop formation, viscosifiers do not alter the free energy surface since dynamics within a single state are observed.

When using the linear relationship (Figure 4.12) or the power law (Figure 4.13), the internal friction and solvent friction in a polypeptide are expected to depend on chain length or amino acid sequence. Identical values for  $\tau_{\text{int}}$  are expected for a polypeptide, independent of the viscosifier used. Values for  $\beta$  are expected to depend on the amino acid sequence, the peptide length and the buffer compositions.

An other model to determine the internal friction was developed by Schulz *et al.*<sup>192</sup>, which found internal friction for end-to-end loop formation in explicit water molecular dynamics simulations. Based on their findings, they developed a rouse model for the mean passage time  $\tau_{\text{mp}}$  for end-to-end loop formation.

$$\frac{\tau_{\text{mp}}}{C_{\text{ee}}} = C_1\tau_{\text{b}} + C_2N\sqrt{\tau_{\text{b}}\tau_{\text{m}}^0 \cdot \frac{\eta}{\eta_0}} + C_3N\tau_{\text{m}}^0 \cdot \frac{\eta}{\eta_0} \quad (4.14)$$

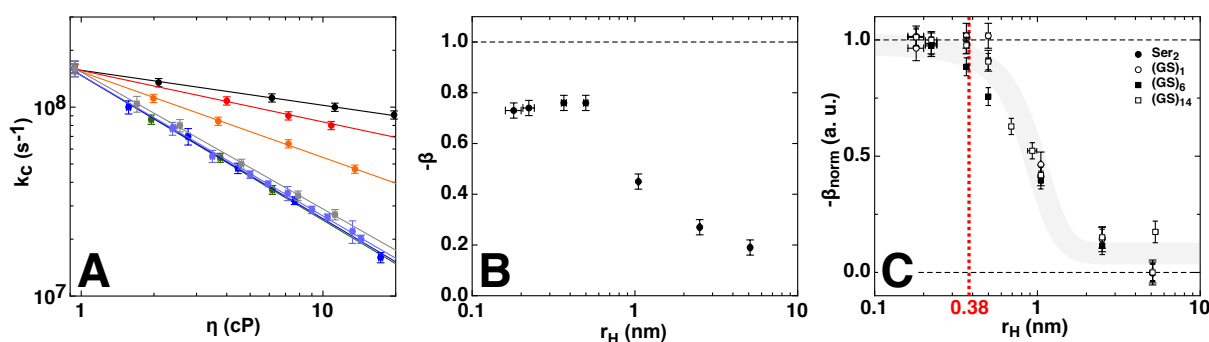
where  $\tau_{\text{m}}^0$  denotes the monomer relaxation time at the reference solvent viscosity  $\eta = \eta_0$ ,  $N$  is the number of inter-chromophore amino acids and  $\tau_{\text{b}}$  is the bond relaxation time. It corresponds to the internal friction parameter  $\tau_{\text{int}}$  in Equation 4.13 and is the only non-vanishing term for  $\eta=0$ . At intermediate viscosities, the second term of Equation 4.14 is dominating  $\tau_{\text{mp}}$ , while the last term dominates  $\tau_{\text{mp}}$  at high viscosities.  $C_{\text{ee}}$ ,  $C_1$ ,  $C_2$  and  $C_3$  are coefficients, which are weighting the effect of the different terms in Equation 4.14 and were estimated to be 0.01, 1.0,  $\sim 1$  and  $\sim 2$ , respectively<sup>192</sup>. In contrast to Equation 4.12 and Equation 4.13, this model yields length independent fit parameters for  $\tau_{\text{m}}^0$  and  $\tau_{\text{b}}$ .  $\tau_{\text{m}}^0$  and  $\tau_{\text{b}}$  might depend on the amino acid sequence but should be identical for homopolypeptides, which only differ in length.

Up to date, it is not clear, whether internal friction in polypeptide chains exists, neither how large its contributions are for chain dynamics nor which model should be used for its determination. Several cases were reported, where the internal friction vanishes<sup>81–83</sup>. In other cases,

### 4.3. Internal Friction and Solvent Friction in Loop Formation Dynamics

a substantial viscosity-independent component upon extrapolation to zero solvent viscosity could be found<sup>77,79,80,86,188–191</sup>. The reported opposing findings might be due to internal friction being small, probably in the order of microseconds or nanoseconds. For slow reactions, the internal friction might appear to be zero, where it is actually not. To accurately determine the internal friction is thus only possible for very fast reactions. For observing the fast kinetics, methods with a high time resolution like temperature-jump or TTET have to be used.

We wanted to test, which of the above models should be used to assess the magnitude of the contributions of internal friction or solvent friction for loop formation in U. We used the unstructured model peptide (GS)<sub>8</sub> and the natural sequences PV-DE-loop and PV-EF-loop (see Chapter 4.1), serving as models for the unfolded state. An advantage in using unfolded polypeptide chains is, that they only populate a single state and therefore are insensitive to thermodynamical effects of the applied viscosifiers. By comparing the model peptides and the natural sequences, we want to determine the influence of side chains on the internal friction and the solvent friction. The determination of internal and solvent friction depends in all models on measuring the time constant for loop formation by TTET at different solvent viscosities. To adjust the solvent viscosity, a multitude of viscosifiers with different sizes could be used. To find the most suitable viscosifier, the influence of their size on the rate constant for loop formation  $k_c$  was studied for several peptides<sup>130</sup>. Figure 4.34 A shows  $k_c$  for (GS)<sub>1</sub> as a function of solvent viscosity, induced by the addition of the viscosogens ethylene glycol, glycerol, glucose, sucrose, polyethylene glycol (PEG) 1500, PEG 6000 and PEG 20000. In



**Figure 4.34.:** Effect of viscosifier size on chain dynamics in (GS)<sub>1</sub>. A) End-to-end loop formation rate constant  $k_c$  as a function of solvent viscosity, which was adjusted by the addition of ethylene glycol (●), glycerol (●), glucose (●), sucrose (●), polyethylene glycol (PEG) 1500 (●), PEG 6000 (●) and PEG 20000 (●). Absolute (B) and normalized (C) sensitivity ( $\beta$ ) of  $k_c$  to solvent viscosity as a function of the co-solute hydrodynamic radius  $r_H$ . Data taken from Reference<sup>130</sup>.

the presence of all co-solutes, the logarithm of the rate constant for loop formation decreases

#### 4. Results and Discussion

---

linearly with  $\log \eta$  and could be described by the empirical Equation 4.15.

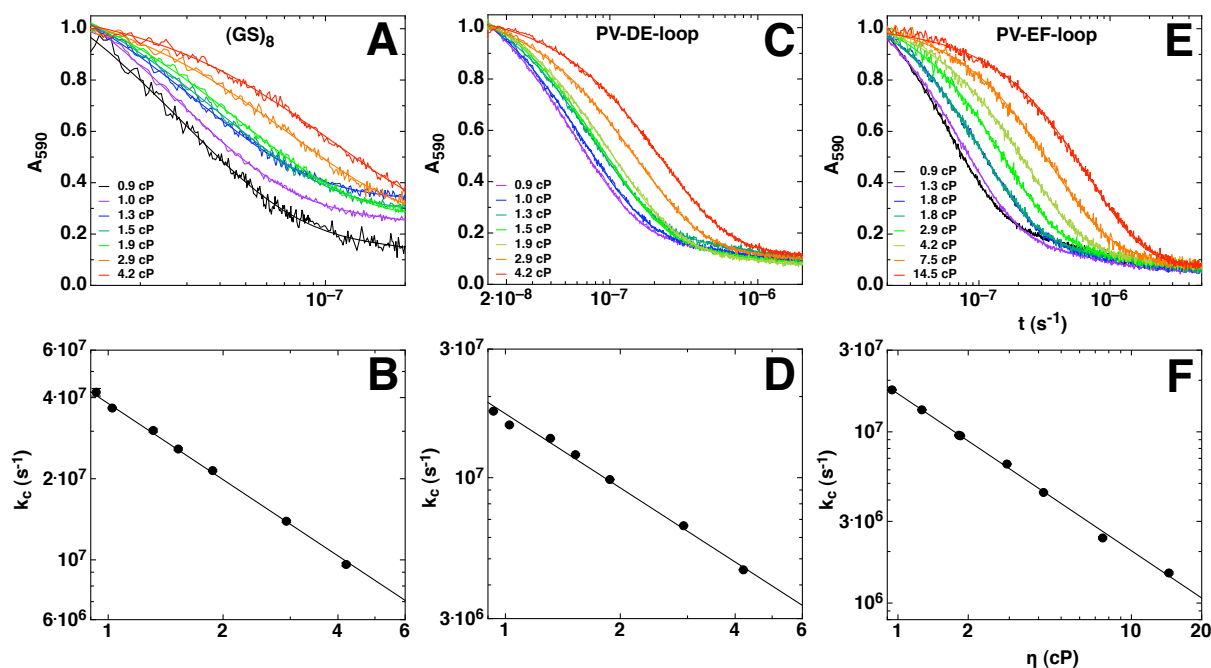
$$k_c = k_0 \cdot \left( \frac{\eta}{\eta_0} \right)^\beta \quad (4.15)$$

The sensitivity  $\beta$  of  $k_c$  to solvent viscosity can be extracted from the slope when plotting  $\log k_c$  as a function of  $\log \eta_{\text{solv}}$ . Peptides with a  $\beta$ -value of -1 experience the full viscous effect, while a  $\beta$  larger than -1 indicates only a fractional dependence of chain dynamics on solvent viscosity. For  $\beta=0$ ,  $k_c$  would be independent of solvent viscosity.  $k_0$  describes the rate constant at the reference solvent viscosity  $\eta_0$ . The magnitude of the effect of solvent viscosity on chain dynamics was found to depend on the co-solute size<sup>130</sup>. A limiting value for  $\beta$  is reached for peptides in the presence of the co-solutes glycerol and glucose, which have hydrodynamic radii ( $r_H$ ) less or equal to the size of the peptide  $C_\alpha$ - $C_\alpha$  distance of 0.38 nm<sup>42,193</sup> (Figure 4.34 C). While in the short (GS)<sub>1</sub> a limiting  $\beta$ -value of -0.75 is found,  $\beta=-1$  is reached for small viscosifiers in long (GS)<sub>n</sub> peptides. The more the co-solute size exceeds the  $C_\alpha$ - $C_\alpha$  length, the more decreases the sensitivity towards  $\eta$  for all peptides (Figure 4.34 C). Peptides thus only sense the full macroscopic viscous effect for co-solutes with  $r_H \leq 0.4$  nm. According to these findings, we chose the small viscogenic agent glycerol to adjust solvent viscosity. Loop formation was determined by TTET in glycerol/water mixtures up to 10 cP in the temperature range between 5 °C and 30 °C. The macroscopic solvent viscosity  $\eta$  was determined by viscosimetry at each temperature. Figure 4.35 top shows the time based changes in absorbance at 590 nm for the indicated solvent viscosities and  $T = 22.5$  °C. The effect of solvent viscosity on the loop formation rate constant is shown on the right, where  $k_c$  decreases linearly with  $\eta$  in a double logarithmic plot. The data can be described by Equation 4.15, yielding the  $\beta$ -value necessary for determining the internal friction when using Equation 4.13. For all peptides  $\beta \neq -1$  (Table 4.3). Chain dynamics therefore experience only a fractional macroscopic viscosity, *i.e.* loop formation is not solely diffusion controlled.

The influence of temperature on  $\beta$  is shown in Figure 4.36. In (GS)<sub>8</sub> (Figure 4.36 A, B) and the PV-EF-loop (C, D),  $\beta$  approaches a value of -1 with increasing temperature. This behavior can be explained by more ideal-like characteristics of the polypeptide chain at higher temperatures. Interestingly,  $\beta$  does not change significantly in the PV-DE-loop (E, F) with temperature. The limiting  $\beta$ -value in this case might already be reached, even at the lowest investigated temperature. The  $\beta$ -values for glycerol at 22.5 °C are summarized in Table 4.13.

To test, which of the three described models describe the data for loop formation in (GS)<sub>8</sub>,

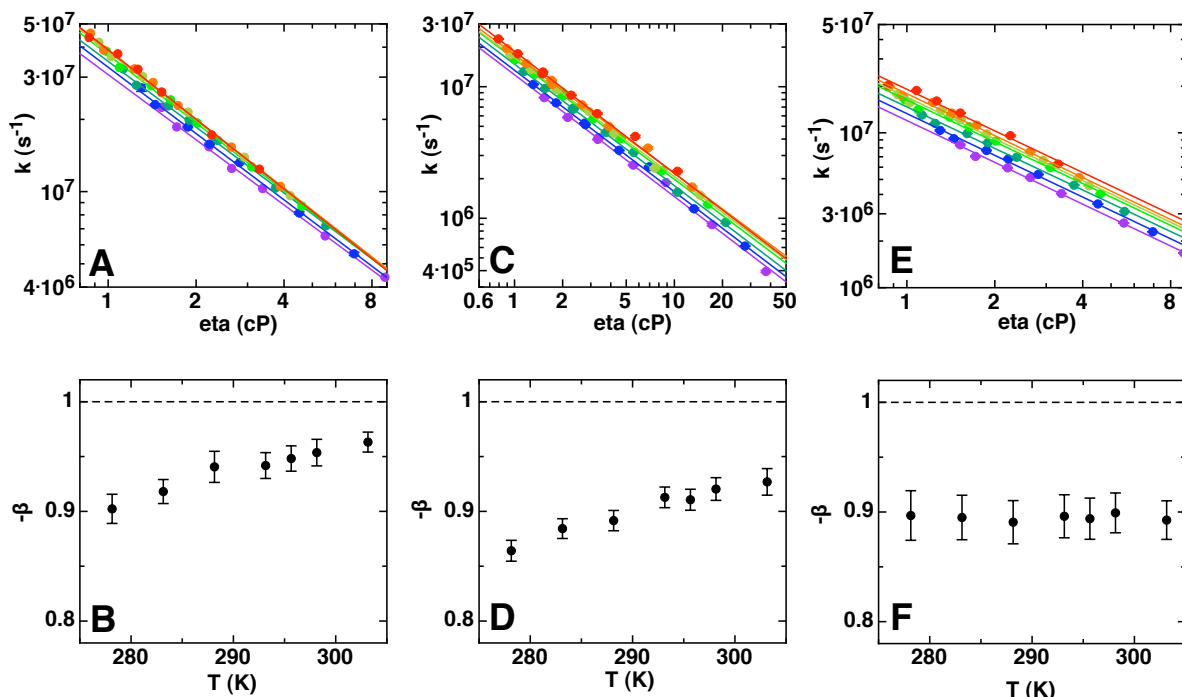
### 4.3. Internal Friction and Solvent Friction in Loop Formation Dynamics



**Figure 4.35.:** Effect of viscosity on chain dynamics. (A, C, E) Time based change in absorbance at 590 nm for the different viscosities. (D, B, F) loop formation rate constants as a function of solvent viscosity. All measurements were performed in different glycerol/water mixtures buffered with 10 mM potassium phosphate pH 7 and 22.5 °C.

the PV-DE-loop and the PV-EF-loop at different solvent viscosities best,  $1/k_c$  was plotted as a function of  $\eta$  (Figure 4.37). The solid line in Figure 4.37 A describes a linear fit to the data according to Equation 4.12, while the broken line represents the fit of the power law (Equation 4.13). From Figure 4.37 A it is not obvious, whether the linear or power law relationship is better suited to describe the time constants for loop formation as a function of  $\eta$ . Loop formation at varying solvent viscosity has been measured earlier for the model peptides  $(GS)_1$  (B, C),  $(GS)_6$  (D),  $(GS)_{14}$  (E) and  $(Ser)_2^{130}$ . Loop formation kinetics of these peptides are shown in Figure 4.37 B-F. The  $(GS)_1$  data could be well described by the power law (Figure 4.37 C) but only poorly by the linear law (Figure 4.37 B). A curvature of the data can be found in all peptides, but it is most pronounced for the short  $(GS)_1$  with its  $\beta$ -value of -0.75 in the limit of small co-solutes. Large viscosifier like polyethylene glycole (PEG) 1500, 6000 or 20000 lead to even more pronounced curvatures (Figure 4.37). For the longer polypeptides  $(GS)_6$  and  $(GS)_{14}$  (Figure 4.37 D and E),  $\beta$  is close to -1 for small viscosifiers but deviates strongly from -1 in the presence of large viscosifiers. The validity of a model can be tested by using different sized viscosifiers, which should all yield an identical value for  $\tau$ . Due to the

#### 4. Results and Discussion



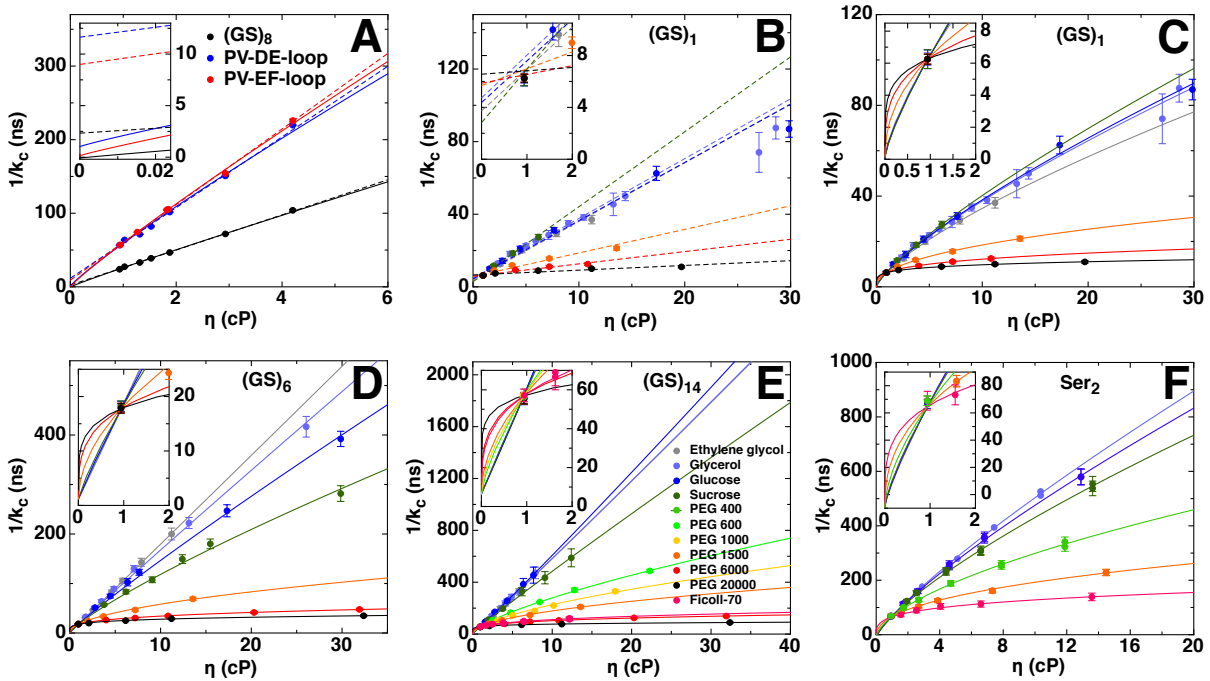
**Figure 4.36.:** Effect of temperature on  $\beta$ . Solid lines represent a fit of Equation 4.15 to the data. A/B)  $(GS)_8$ , C/D) PV-EF-loop, E,F) PV-DE-loop. 5 °C (●), 10 °C (●), 15 °C (●), 20 °C (●), 22.5 °C (●), 25 °C (●) and 30 °C (●).

observed downward curvature and varying results for  $\tau_{\text{int}}$  depending on the viscosifier size, the linear function (Equation 4.12) is not suited to describe the loop formation time constant at different solvent viscosities. Therefore, our findings suggest to use the power law (Equation 4.13) to determine the internal friction time constant for loop formation. As a consequence, published values for  $\tau_{\text{int}}$  determined by linear extrapolation<sup>194</sup> are too high when  $\beta$  is expected to deviate from -1. For the peptides  $(GS)_n$  with  $n = \{1, 6, 14\}$  and  $\text{Ser}_2$ ,  $\tau_{\text{int}}$  and  $\tau_{\text{solv}}$  were fitted globally for the various viscosifiers, since they were found to be independent of the viscosifier size. The fit yields  $\tau_{\text{int}} = 0.15 \pm 1.71$  ns and  $\tau_{\text{solv}} = 6.11 \pm 1.40$  ns with  $\beta$  fitted individually for each co-solute. Internal friction in  $(GS)_1$  is not significantly different from zero when determined with Equation 4.13. Also for the longer  $(GS)_n$  peptides with  $n = \{6 \& 14\}$ , the global fit yields  $\tau_{\text{int}} = 0$ . In all three peptides, internal friction for loop formation is thus absent at room temperature. The results of the fit are summarized in Table 4.13 for all investigated peptides.

We studied the effect of varying temperature on  $\tau_{\text{int}}$  and  $\tau_{\text{solv}}$  to find out, whether loop formation at other temperatures is associated with a nonzero internal friction. We investigated



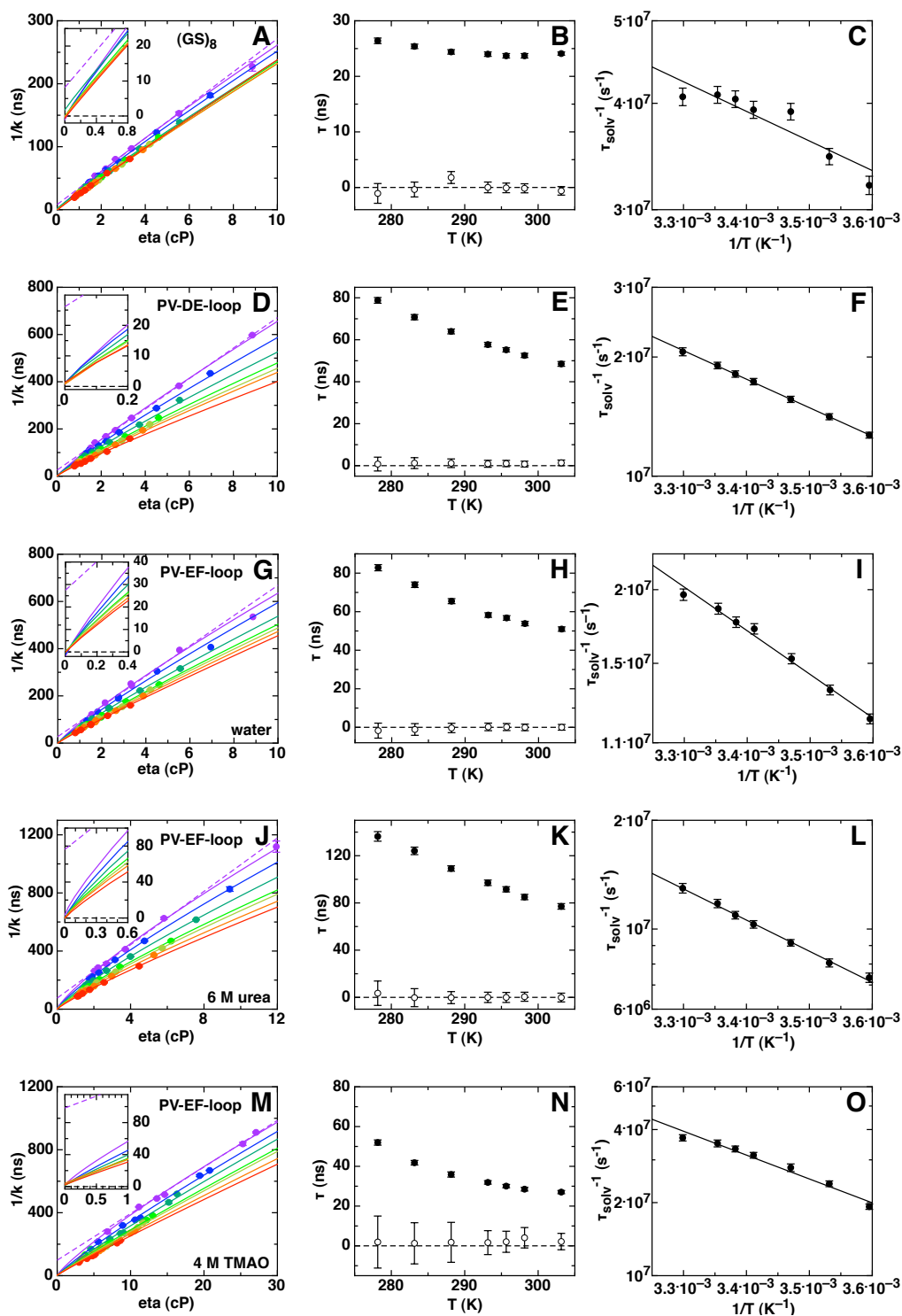
### 4.3. Internal Friction and Solvent Friction in Loop Formation Dynamics



**Figure 4.37.:** Comparison between linear fit (dashed line) and power law (solid line). In panels C-F,  $\tau_{\text{int}}$  and  $\tau_{\text{solv}}$  were fitted globally for all viscogens while  $\beta$  was fitted independently for each viscogen. (B-F) Data taken from Reference <sup>130</sup>.

three peptides, the  $(\text{GS})_8$  and the two PV-loops. To test, if the formation or breaking of intramolecular interactions affects the  $\tau_{\text{int}}$  for loop formation in the PV-EF-loop, measurements were performed in the presence of 4 M TMAO or 6 M urea, to render a poor or good solvent for the polypeptide chain, respectively. All measurements were performed in the temperature range from 5 °C - 30 °C in the presence of 10 mM potassium phosphate buffer at pH 7. Solvent viscosity was adjusted by adding glycerol to the sample buffer. The results of the measurements at different temperatures are summarized in Figure 4.38. The solid lines correspond to a fit of the power law (Equation 4.13). The values for  $\tau_{\text{int}}$  and  $\tau_{\text{solv}}$  obtained from the fit at 22.5 °C are summarized for the model peptides and the natural sequences in Table 4.13. The  $\beta$ -values were determined for each peptide at the respective temperature from the slopes of Figure 4.36. The values for  $\tau_{\text{int}}$  and  $\tau_{\text{solv}}$  at different temperatures are plotted as a function of temperature in Figure 4.38 B, E, H, K and N. In all peptides,  $\tau_{\text{int}}$  is zero within error and independent of temperature, pointing out a low or non-existing internal friction for loop formation, even at low temperatures. In both the good and the poor solvent,  $\tau_{\text{int}} = 0$ . While an inexistent or low internal friction is expected in the presence of urea, it is surprising in the presence of

## 4. Results and Discussion



**Figure 4.38.:** Temperature dependent determination of  $\tau_{int}$  ( $\circ$ ) and  $\tau_{solv}$  ( $\bullet$ ). All measurements were performed in 10mM sodium phosphate buffered glycerol/water mixtures at pH 7. Temperatures: 5 °C ( $\bullet$ ), 10 °C ( $\bullet$ ), 15 °C ( $\bullet$ ), 20 °C ( $\bullet$ ), 22.5 °C ( $\bullet$ ), 25 °C ( $\bullet$ ) and 30 °C ( $\bullet$ ). In panels A, D, G, J and M the dashed lines represent a fit to the data at 5 °C with  $\beta = -1$ .

TMAO, since the formation of intramolecular interactions leads to stiffer chain. The number or strength of the intramolecular interactions thus does not lead to a significant internal friction. We therefore assume, that the chain dynamics are mainly governed by entropic rather than enthalpic effects in the low viscosity regime.

The effect of temperature on  $1/\tau_{\text{solv}}$  can be described by the Arrhenius Equation (Figure 4.38 C, F, I, L, O).  $\tau_{\text{solv}}$  is related with  $k_c$  by

$$\tau_{\text{solv}} = \frac{1}{k_c(\text{H}_2\text{O})} - \tau_{\text{int}} \left( \frac{\eta}{\eta_0} \right)^\beta. \quad (4.16)$$

For  $\tau_{\text{int}} = 0$ , Equation 4.16 yields  $\tau_{\text{solv}} = 1/k_c(\text{H}_2\text{O})$ . In contrast to  $\tau_{\text{int}}$ ,  $\tau_{\text{solv}}$  exhibits a strong temperature dependence. It decreases with increasing temperature in all peptides but depends on the solvent composition and the amino acid sequence. The Arrhenius parameters for the temperature dependence of  $\tau_{\text{solv}}$  from Figure 4.38 are summarized in Table 4.13. All values for  $E_A$  lie in the range around 14 - 19 kJ/mol (except in the (GS)8 peptide, where the data quality is not good enough to be described by the Arrhenius equation) and correspond to an apparent activation energy whose origin is the temperature based change in solvent viscosity. To test, whether the length independent internal friction can be described globally for (GS)-peptides of different length, we used Equation 4.14 to describe the loop formation time constant as a function of  $\eta$  for (GS)<sub>n</sub> with  $n = \{1, 2, 3, 4, 5, 6, 8, 12, 14\}$ . The global fit for  $\tau_b$  and  $\tau_m^0$  yields  $0.68 \pm 0.06$  ns and  $0.35 \pm 0.01$  ns, respectively with the coefficients  $C_{\text{ee}^*} = 1$ ,  $C_1 = 1$ ,  $C_2 = 5$  and  $C_3 = 2.5$ .  $\tau_b$  appears to be near zero. From the fit in Figure 4.39 can be seen, that the model does not describe the data sufficiently, especially in case of the longer peptides at low and intermediate solvent viscosities. This points out, that the first and middle term are not suited to describe the data very well and the results from the fit need to be interpreted carefully. Additionally, the last term suggests a linear relationship at high viscosities. However, curvatures have been found in measured data even under these conditions. Due to the infirmities, Equation 4.14 is not well suited to describe the internal friction for loop formation.

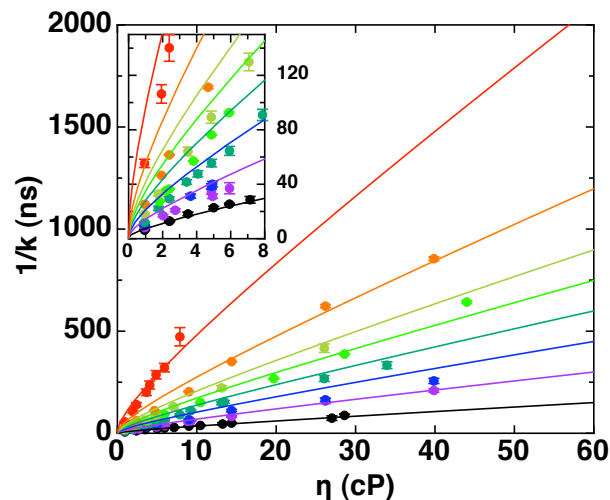
### 4.3.1. Internal Friction in the Dynamics for Loop Formation

To test, which model should be used to describe the effect of solvent viscosity on the time constant for loop formation, we measured TTET in different glycerol/water mixtures. Our results show a downward curvature for  $1/k_c$  as a function of  $\eta$  and could only be described by a power law (Equation 4.13), where the exponent  $\beta$  accounts for the curvature of the data.

**Table 4.13.:** Arrhenius parameters for the temperature dependence of  $1/\tau_{\text{solv}}$ .

Peptide	Condition	$\tau_{\text{int}}^a$ (s)	$\tau_{\text{solv}}^a$ (s)	$-\beta_{\text{Glycerol}}^a$	$E_A$ (kJ/mol)	$A$ (s <sup>-1</sup> )
(GS) <sub>1</sub>	buffer	0.15 ± 1.71	6.1 ± 1.4	0.77 ± 0.07	-	-
(GS) <sub>6</sub>	buffer	1.16 ± 2.76	16.6 ± 2.4	0.98 ± 0.05	-	-
(GS) <sub>8</sub>	buffer	-0.08 ± 0.86	23.7 ± 0.4	0.95 ± 0.01	-	-
(GS) <sub>14</sub>	buffer	6.45 ± 18.38	50.8 ± 16.9	1.03 ± 0.17	-	-
(Ser) <sub>2</sub>	buffer	3.27 ± 2.65	9.3 ± 2.4	0.93 ± 0.08	-	-
PV-DE-loop	buffer	1.02 ± 1.51	55.2 ± 1.0	0.89 ± 0.02	13.8 ± 0.6	(4.9 ± 1.3)·10 <sup>9</sup>
PV-EF-loop	buffer	0.15 ± 1.83	56.8 ± 1.2	0.91 ± 0.01	14.2 ± 0.7	(5.6 ± 1.5)·10 <sup>9</sup>
PV-EF-loop	6 M urea	-0.07 ± 4.25	91.5 ± 2.2	0.85 ± 0.02	16.5 ± 0.9	(9.0 ± 3.5)·10 <sup>9</sup>
PV-EF-loop	4 M TMAO	2.05 ± 5.30	30.0 ± 0.8	0.94 ± 0.02	18.8 ± 5.3	(6.9 ± 15.0)·10 <sup>10</sup>

<sup>a</sup>for T = 22.5 °C<sup>b</sup>Globally fitted for all viscogenic agents.



**Figure 4.39.:** Influence of peptide length on the time constant for loop formation as a function of solvent viscosity. The solid line describes a global fit of Equation 4.14 to the data with  $C_{ee}^* = 1$ ,  $C_1 = 1$ ,  $C_2 = 5$  and  $C_3 = 2.5$ .  $(GS)_1$  (●)  $(GS)_2$  (●),  $(GS)_3$  (●),  $(GS)_4$  (●),  $(GS)_5$  (●),  $(GS)_6$  (●),  $(GS)_8$  (●)  $(GS)_{14}$  (●). Data taken from Reference<sup>130</sup>.

In short peptides and in the presence of large viscosifiers, the curvature is most pronounced. When determining the internal friction time constant by extrapolation to zero solvent viscosity, we find  $\tau_{int}$ -values of approximately zero for all peptides, independent of temperature or solvent composition. Even in poor solvents, where intramolecular interactions are enhanced at the cost of peptide-solvent interactions, the internal friction was still insignificant. Internal friction thus seems to be absent or very low, not exceeding some hundreds of picoseconds. Internal friction therefore has only a small contribution to loop formation dynamics.

In literature,  $\tau_{int}$  was determined by a linear relationship, a power law or a Rouse model, in which internal friction is included. When using the linear extrapolation, an internal friction of several microseconds was observed in several cases. Compared with our findings, a value of several microseconds for the internal friction is too large and therefore unlikely. The difference can be attributed to the usage of the linear law to determine the internal friction, since generally a smaller internal friction is found when determined with the power law. In other cases, a negligible contribution from internal friction was observed. This is in accordance with our results, where internal friction is small. Based on our findings, the internal friction should be determined using the power law with the respective value for  $\beta$ . Only in fast reactions, where the internal friction time constant is close to the time constant of the observed reaction the internal friction can be determined accurately.

## 4.4. Dynamics of Loop Formation in the EF-Hand and Full Length Carp $\beta$ -Parvalbumin in the Unfolded and an Intermediate State

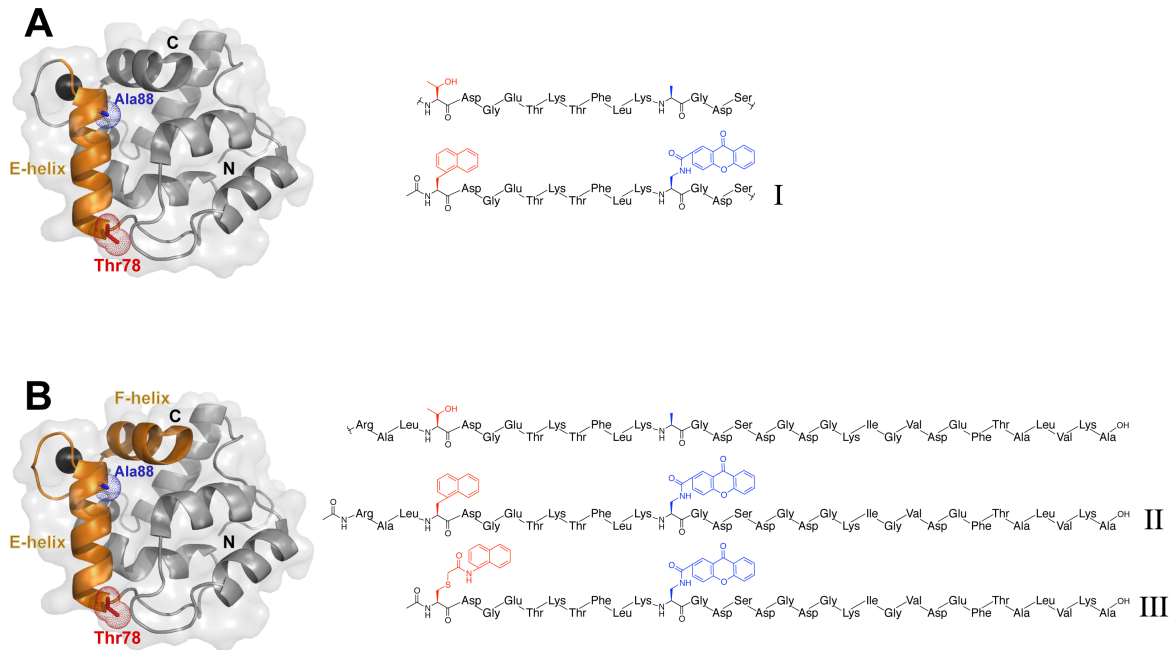
In proteins, long range interactions stabilize the native state and partially folded intermediate states. It is not known, what effect long range interactions have on chain dynamics in the different states of a protein. We wanted to test their influence on loop formation kinetics in different states of carp  $\beta$ -parvalbumin (PV) by comparing EF-hand fragments with the full length protein.

### 4.4.1. Loop Formation in Different Fragments from the EF-Hand Motif of Carp $\beta$ -Parvalbumin

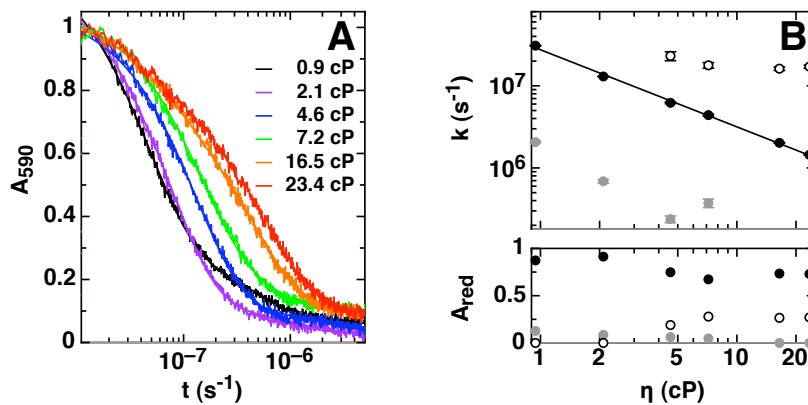
Loop formation dynamics were measured by TTET in an E-helical fragment and the EF-hand. Figure 4.40 shows the chemical structure of EF-hand fragments, which are all labeled at identical positions ( $i/i+10$ ). Fragment I represents the PV residues 78 to 91, which form the E-helix in the full length protein. Fragments II and III represent the EF-hand of PV (residues 78-108). In all three fragments, Thr78 was replaced by the triplet acceptor, while Ala88 was replaced by the triplet acceptor. In the EF-hand fragment II, Nal was used as acceptor. In Fragment III, the naphthalene moiety was attached to the Cys side chain using the thiol reactive and water soluble compound 2-bromo-*N*-(naphthalen-1-yl)acetamide (BNAA). The labeling reaction results in the Cys side-chain modified by *N*-(naphthalen-1-yl)acetamide (NAA), with bromine serving as leaving group. The properties of NAA as triplet acceptor in TTET were tested in bimolecular TTET experiments yielding a diffusion controlled bimolecular rate constant of  $(4.1 \pm 0.3) \cdot 10^9 \text{ M}^{-1} \text{ s}^{-1}$ <sup>162</sup>.

Loop formation was measured in the E-helix (Fragment I) at different solvent viscosities. Viscosity was adjusted by the addition of glycerol to the buffer solution. Figure 4.41 shows the time based absorbance decay at the indicated solvent viscosities. The kinetic traces could be described by double exponential kinetics for low and high solvent viscosities. A three exponential function was necessary to describe the kinetics at intermediate solvent viscosities. The rate constants and relative amplitudes are summarized in Figure 4.41 B. The rate constant of the main phase with an amplitude of 70-90 % could be described by Equation 4.1 (solid line), yielding  $k_c(\text{H}_2\text{O}) = (2.90 \pm 0.05) \cdot 10^7 \text{ s}^{-1}$  and  $\beta = -0.94 \pm 0.01$ . A slow phase with low

#### 4.4. Dynamics of Loop Formation in the EF-Hand and Carp $\beta$ -Parvalbumin



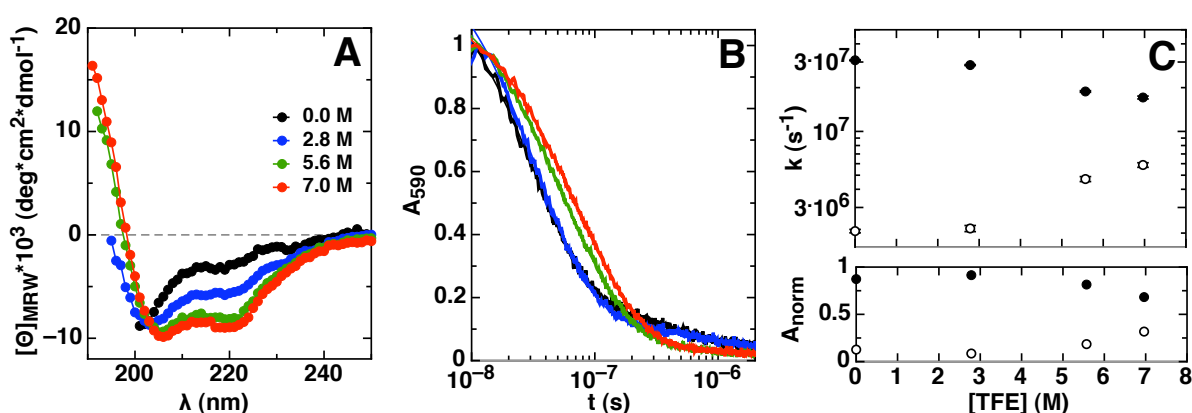
**Figure 4.40.:** Sequence of C-terminal PV fragments with their location indicated in orange in the full length protein. (A) E-helix (residues 78-91). (B) EF-hand (residues 78-108). In all peptides, Thr78 was replaced by the triplet acceptor (red) and Ala88 by the triplet donor (blue). The figure was prepared using PyMOL<sup>173</sup> and the pdb<sup>146</sup> files 4CPV<sup>145</sup> for PV.



**Figure 4.41.:** Chain dynamics of the PV E-helix as a function of solvent viscosity. (A) Time-based change in absorbance at 590 nm at the viscosities indicated. (B) Rate constants with their corresponding amplitude as a function of solvent viscosity  $\eta$ . The solid line corresponds to Equation 4.2. All measurements were performed at 10 mM potassium phosphate at pH 7.

#### 4. Results and Discussion

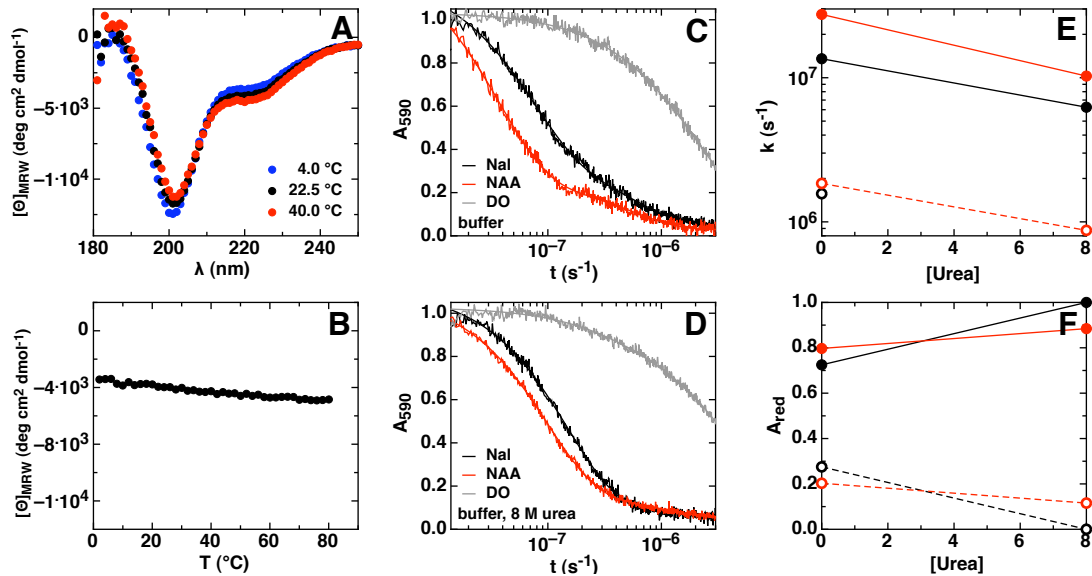
amplitude can be observed for  $\eta \leq 7$  cP, corresponding to the intrinsic lifetime of xanthone. At viscosities  $\geq 4$  cP, a fast phase with an amplitude of 20-30% can be found. This fast phase with a rate constant of about  $2 \cdot 10^7 \text{ s}^{-1}$  is independent of solvent viscosity. It is assumed, that residual helical structure exists, which is stabilized by high concentrations of glycerol. When the helix is formed, the labels are close together and are able to form contact by minor conformational rearrangements. Since the E-helix seems to have a latent tendency to form a helical structure, we used 2,2,2-trifluoroethanol (TFE) as co-solute, to induce secondary structure<sup>195,196</sup> and measure TTET in a helical state (Figure 4.42). From the CD-measurement,



**Figure 4.42.:** Chain dynamics in the PV E-helix as a function of the TFE concentration. (A) CD spectrum at the indicated [TFE]. (B) Time-based decay of the absorbance at 590 nm under identical conditions. (C) Rate constants and corresponding amplitudes.

it can be seen that in the absence and at low TFE concentrations, helix content is negligible. It can be gradually increased with increasing TFE concentrations (the helix content of the E-helical fragment cannot be calculated due to the contribution of the chromophores to the CD-signal). The isosbestic point at 202 nm indicates, that the peptide contains only of two secondary structure elements, which are  $\alpha$ -helix and coil. The results show, that the helical content in the E-helix can be adjusted easily by the addition of TFE. In the EF-hand, no secondary structure can be observed by CD (Figure 4.43 A, B). Loop formation was measured by TTET in the EF-hand fragments II and III in water and 8 M urea (Figure 4.43). The time-based decay at 590 nm could be described for the NAA-labeled EF-hand in both conditions by double exponential kinetics. For the Nal labeled EF-hand peptide, double exponential kinetics were necessary in the absence of urea while a single exponential was sufficient to describe the data in 8 M urea. The rate constants corresponding to the main or single kinetic phase are shown in Figure 4.43 E and F. At both conditions, the kinetics differ for the two





**Figure 4.43.:** Loop formation dynamics in EF-hand peptides. (A) Far UV CD spectrum of the EF-hand. (B) Thermal transition of the EF-hand at 222 nm. Comparison of the absorbance decay of the EF-hand with the acceptor Nal (black), NAA (red) and donor-only (grey) in dilute buffer (C) and 8 M urea (D). Rate constants (E) and amplitudes (F) from the kinetics in panels C and D, using identical colors.

alternatively labeled EF-hand peptides. The rate constant for loop formation is slower in the Nal labeled peptide compared to the NAA labeled peptide. This might be due to additional N-terminal amino acids in Fragment II, which were introduced to prevent the triplet acceptor from interactions with the N-terminus, or due to the longer linker in Fragment III, which allows loop formation without the formation of a very tight loop. To compare loop formation dynamics of the EF-hand with the full length protein, results from the NAA-labeled version were used. The kinetics for loop formation are identical in Fragment I and III, suggesting also the presence of residual structure in the EF-hand as in the E-helix, in which the labels are close and loop formation can occur fast and where the additional tail in the EF-hand does not affect loop formation.

#### 4.4.2. Loop Formation in the Unfolded and in an Intermediate State of Full Length Carp $\beta$ -Parvalbumin

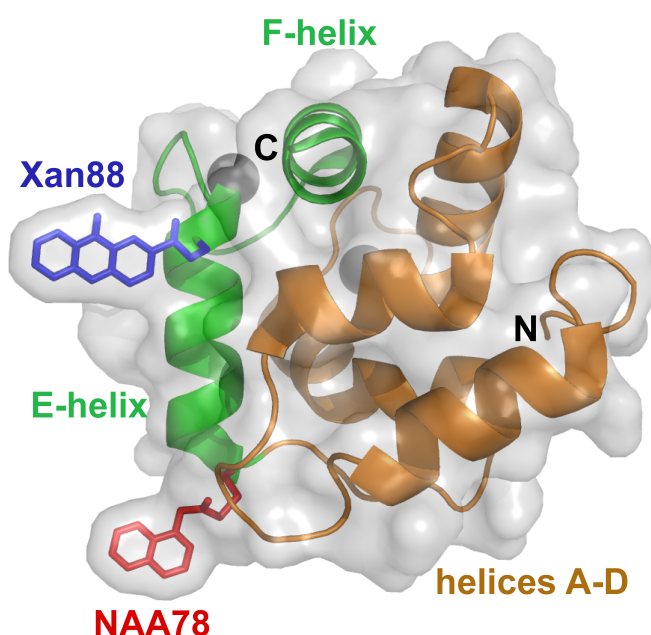
To measure loop formation in different states of a full length protein,  $\beta$ -parvalbumin from *cyprinus carpio* represents an ideal model, since a thermodynamic intermediate exists, which can be populated in the absence of the ligand calcium<sup>149</sup>. Furthermore, the PV wild-type

## 4. Results and Discussion

contains only two amino acids, namely Cys18 and His26, which interfere with the triplet state of xanthone. They can be replaced by Ser and Phe, respectively without affecting the protein stability<sup>150</sup>.

### 4.4.2.1. Semisynthesis of Carp $\beta$ -Parvalbumin for TTET Measurements

In order to measure loop formation in full-length PV by TTET, the moieties xanthone and naphthalene need to be introduced at specific positions along the protein. To gain insight into the local dynamics of a helix in the context of a full-length protein, we chose to label the amino acid side chains at positions 78 and 88 with naphthalene and xanthone, respectively, which reside at the ends of the E-helix and are solvent exposed. The structure of PV with the triplet labels is shown in Figure 4.44 (PV<sup>TTET</sup>). Xanthone is colored in blue, naphthalene in red and the EF-hand structural motif is represented in green. In short peptides, Xan and Nal



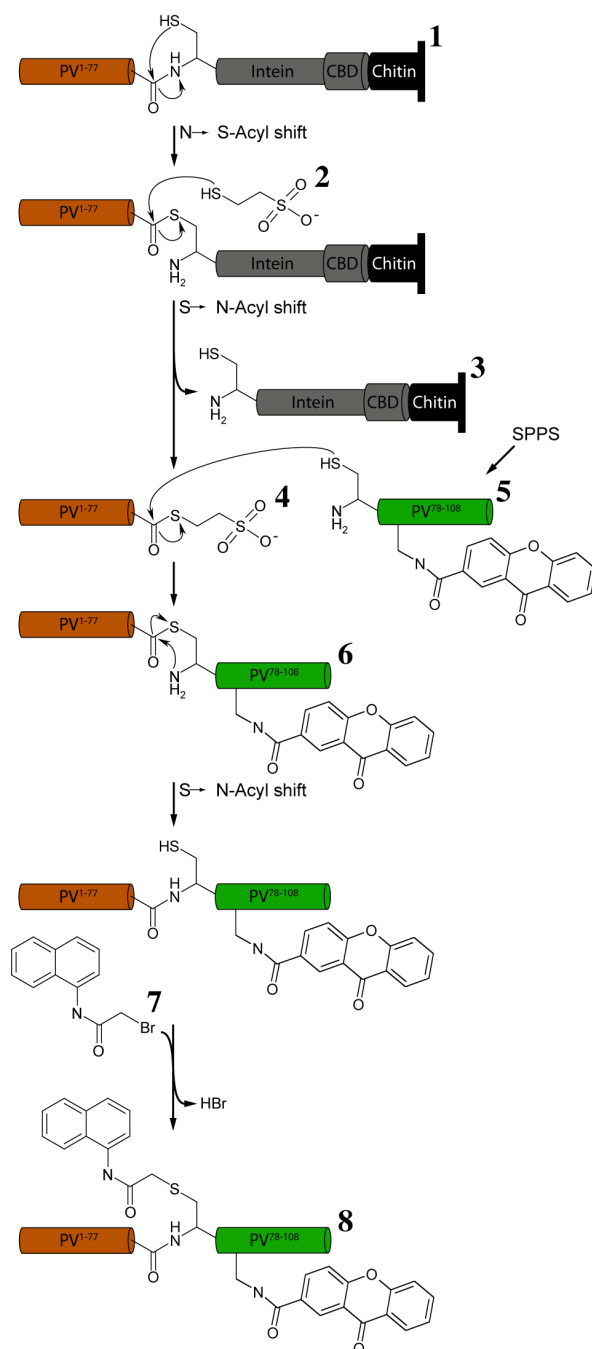
**Figure 4.44.:** PV variant for TTET measurements (PV<sup>TTET</sup>). Xanthone attached at position 88 is shown in blue, naphthalene at position 78 in red. The EF-hand structural motif is represented in green, residues 1-77 in orange. The N- and C-termini are indicated in black.

can be introduced in the course of SPPS, allowing for the introduction of nonnatural amino

acids. However, full-length proteins like PV exceed the accessible length of peptides, which can be fully chemically synthesized. Long polypeptides and proteins are usually expressed in recombinant organisms. However, it is not possible to introduce chromophores directly during protein expression. In order to prepare a full length, doubly labeled version of PV, we wanted to combine the strength of both methods by using native chemical ligation (NCL) to condense an expressed long fragment with a chemically synthesized fragment bearing the triplet donor. The strategy for the production of full-length doubly labeled PV is shown in detail in Figure 4.45. First, a fusion protein with the N-terminal protein fragment (PV residues 1-77) with an in frame C-terminal intein and a chitin binding site needs to be expressed. The fusion protein can be bound to a chitin column, purified and cleaved on column with  $\beta$ -mercaptosulfonic acid (MESNA), which leads to an  $\alpha$ -carboxyl MESNA thioester. The thioester can be eluted and mixed with a Xan bearing C-terminal protein fragment, synthesized by SPPS. An N-terminal cysteine of this fragment is crucial for the ligation reaction with the C-thioester of the expressed N-terminal protein fragment. The sulfur of the Cys residue attacks the  $\alpha$ -carboxyl MESNA thioester in a chemo selective nucleophile reaction. A following S $\rightarrow$ N acyl shift results in a native peptide bond, linking the two protein fragments irreversibly together and yielding the full length protein. Unfortunately, the sulfur of the cysteine, necessary for the ligation reaction, quenches the triplet state of xanthone<sup>32</sup>. This drawback can be circumvented using a thiol reactive naphthyl derivative to attach the triplet acceptor to the side chain of this Cys residue.

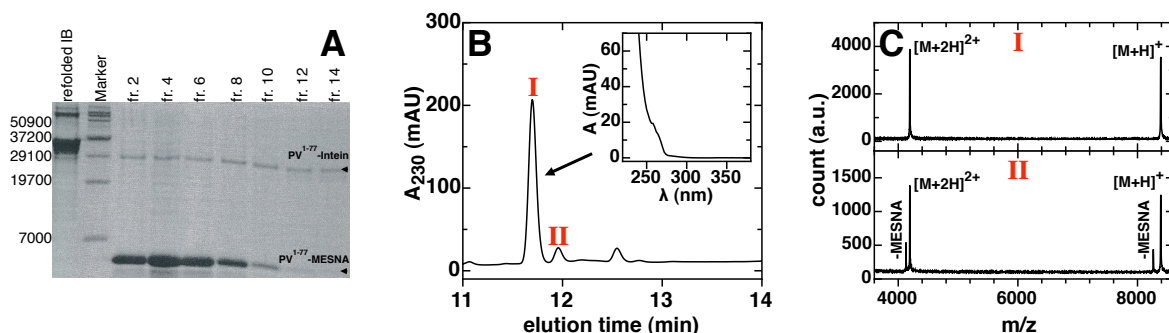
To generate the N-terminal PV fragment (residues 1-77) with the  $\alpha$ -carboxyl thioester, we used the high expression plasmid pTXB3, where the PV fragment was cloned upstream in frame with the *Mxe* GyrA-intein, using the restriction sites NcoI and SapI. Protein expression was induced by the addition of  $\beta$ -D-1-thiogalactopyranose (IPTG) at an OD<sub>600</sub> of 0.6 in the logarithmic growth phase. The fusion protein was expressed as inclusion bodies. They were washed, disaggregated in high concentrations of urea and the protein renatured after removal of urea by dialysis. To prepare the  $\alpha$ -carboxyl thioester, the renatured fusion protein was loaded onto a chitin column. The column was washed, flushed with MESNA, the flow stopped and the mixture incubated over night at room temperature. The PV<sup>1-77</sup>-MESNA thioester was eluted the next day from the chitin column. Purity and identity were verified by SDS PAGE, HPLC and mass spectrometry (Figure 4.46). PV<sup>1-77</sup> containing the reactive  $\alpha$ -carboxyl MESNA thioester (PV<sup>1-77</sup>-MESNA) corresponds to the lower band in the gel. The faint upper band with a size of  $\sim$  30000 Da corresponds to the uncleaved intein fusion protein. A theoretical mass of 8386 Da was calculated for the single protonated PV<sup>1-77</sup>  $\alpha$ -carboxyl

#### 4. Results and Discussion



**Figure 4.45.:** Native chemical ligation (NCL) of a protein thioester fragment with a synthesized fragment. The C-terminal thioester of the N-terminal protein fragment was planned to be generated with the help of an in frame C-terminal intein via an N→S acyl shift followed by cleavage with sodium 2-sulfanylethanesulfonate (MESNA). The so-formed thioester can be ligated in a chemo selective step to a synthesized C-terminal protein fragment, bearing the triplet donor and an N-terminal Cys, necessary for the transthioesterification. A following S→N acyl shift would yield the native peptide bond at the ligation site. The chromophore moiety naphthalene serving as triplet acceptor can be introduced in a thiol reactive step using 2-bromo-N-(naphthalen-1-yl)acetamide (BNA).

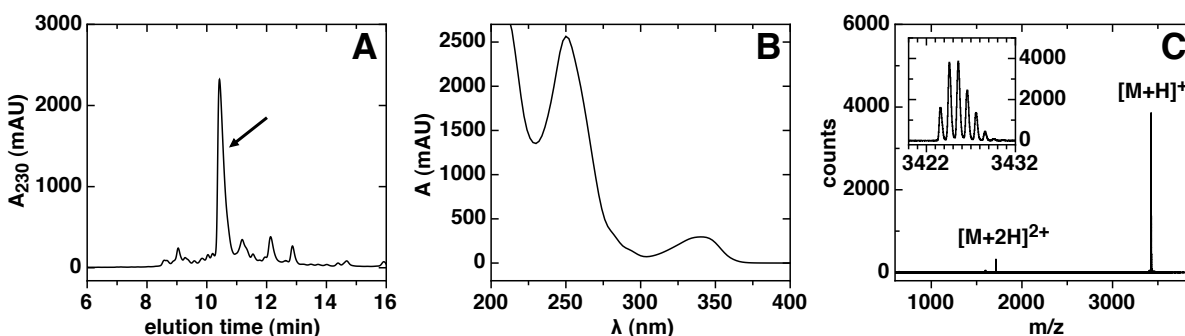
#### 4.4. Dynamics of Loop Formation in the EF-Hand and Carp $\beta$ -Parvalbumin



**Figure 4.46.:** Analysis of prepared PV<sup>1-77</sup>  $\alpha$ -carboxyl MESNA thioester. SDS-PAGE of eluted fractions from the chitin column after intein cleavage with MESNA (A). Verification of the desired product by HPLC and mass spectrometry (B). Calculated average mass for the single protonated PV<sup>1-77</sup>  $\alpha$ -carboxyl MESNA thioester: 8386 Da.

MESNA thioester. Mass spectrometry was performed for the main (I) and a minor (II) peak, eluted by analytical HPLC. For the main peak, a single protonated mass of 8386 Da can be found. The minor peak corresponds to the hydrolyzed thioester, resulting in a mass shift of 123 Da. The hydrolysis of the thioester can be slowed down at low temperatures but it is relatively long-lived, even PV at room temperature. The thioester was stored at  $-20^{\circ}\text{C}$  for subsequent ligation steps.

The C-terminal fragment corresponding to the PV residues 78-108 (PV<sup>78-108</sup>) bearing the chromophore xanthone was synthesized by standard 9-fluorenylmethoxycarbonyl (Fmoc) SPPS. Xanthone was attached to the side chain of the nonnatural amino acid  $\alpha$ -L-diaminopropionic acid (Dpr) at position 88 in the fully protected peptide during peptide synthesis. The N-terminal amino acid Thr 78 of PV<sup>78-108</sup> was replaced by Cys. The cleavage product (Figure

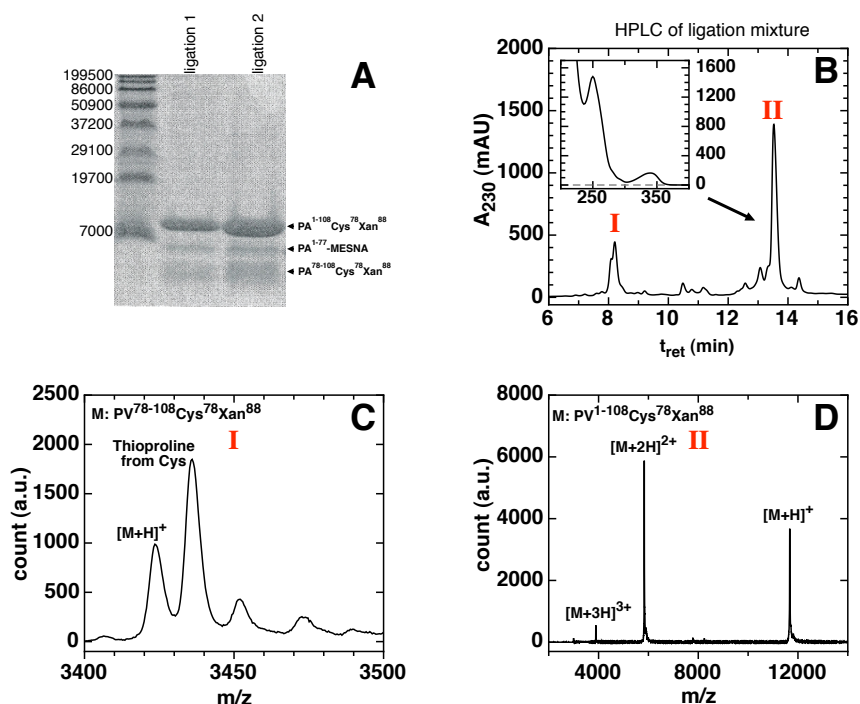


**Figure 4.47.:** Analysis of synthesized PV<sup>78-108</sup>. (A) Analytical HPLC of crude product. (B) Absorption spectrum of the main product (arrow). (C) Mass spectrometry measurement of the main product (arrow).

## 4. Results and Discussion

4.47 A) was purified by HPLC. The purity and identity of PV<sup>78–108</sup> was confirmed by analytical HPLC and mass spectrometry. The desired product can be identified by the characteristic absorbance band of the attached Xan at a wavelength of 343 nm (Figure 4.47 B). The calculated average mass of 3426 Da could be confirmed by MALDI-TOF (C).

In order to obtain the full length PV<sup>1–108</sup>, the PV<sup>1–77</sup>  $\alpha$ -carboxyl MESNA thioester was mixed 1:1 with PV<sup>78–108</sup>. Urea was added to a final concentration of 1.5 M, to increase solubility and remove residual structure in the fragments. The mixture was incubated while shaking at room temperature over night (see Materials and Methods, Chapter 3.6.2). The success of the ligation reaction was checked by SDS PAGE, analytical HPLC and mass spectrometry (Figure 4.48). A mass of 11670 Da was measured for the ligated product, identical with the expected mass of

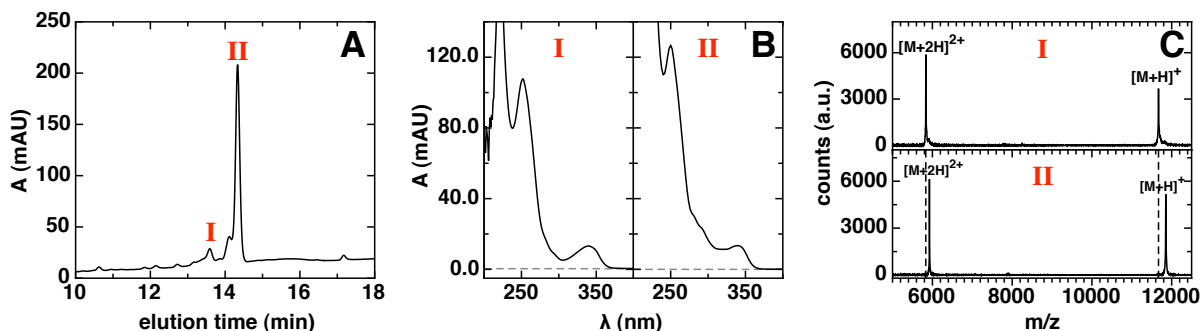


**Figure 4.48.:** Analysis of the native chemical ligation. (A) SDS PAGE of ligation mixtures after 12 h incubation. (B) Analytical HPLC after the ligation reaction. Ligated, full length PV is indicated by the arrow. Mass spectrometry of HPLC peaks I (C) and II (D).

the full length protein. Full length PV<sup>1–108</sup>Cys<sup>78</sup>Xan<sup>88</sup> (species II, Figure 4.48 B and D) was purified by semi preparative HPLC, freeze dried and stored at -20 °C. Eventually oxidized Cys residues (*e.g.* by MESNA) were reduced by dithiothreitol (DTT) prior to purification. The non-ligated educt PV<sup>78–108</sup>Cys<sup>78</sup>Xan<sup>88</sup> (species I, Figure 4.48 B, C) was recycled for later ligation reactions after being purified and lyophilized.

#### 4.4. Dynamics of Loop Formation in the EF-Hand and Carp $\beta$ -Parvalbumin

BNAA was used to attach the naphthyl moiety NAA to Cys 78 in the full length protein. Labeling was performed for 40 minutes in aqueous buffer at pH 8.48 and its success confirmed by analytical HPLC and mass spectrometry (Figure 4.49). An increase in the mass of 183 Da



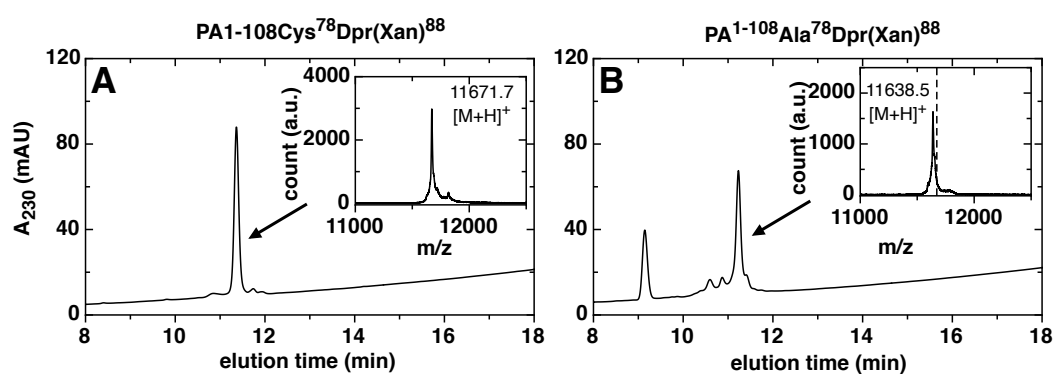
**Figure 4.49.:** Labeling of the Cys sidechain with the thiol reactive naphthyl derivative BNAA. (A) HPLC profile with unlabeled (I) and labeled (II) PV. Absorption spectra (B) and mass spectra (C) of the two species.

corresponds to the attachment of NAA to the Cys side-chain. Excess labeling reagent was removed by semi preparative HPLC. Purified labeled PV<sup>T<sup>T</sup>ET</sup> was freeze dried and stored at -20 °C.

To test the lifetime of the Xan triplet state in PV, a donor-only variant with xanthone at position 88 was prepared by NCL, as described for the doubly labeled PV. The quenching Cys sulfur, necessary for the ligation reaction, was removed with a selective free radical desulfurization method<sup>167</sup> (see Materials and Methods), resulting in an alanyl-side chain. Desulfurization of the full length protein could be followed by analytical HPLC and resulted in a mass shift of 32 Da (Figure 4.50).

#### 4. Results and Discussion

---

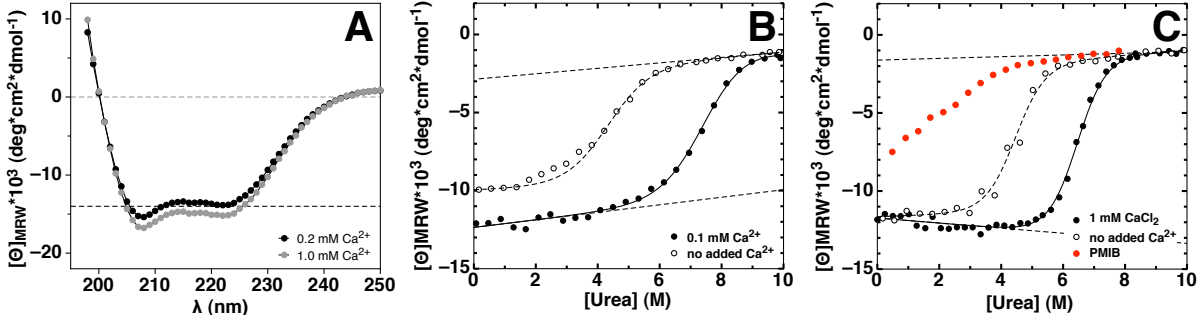


**Figure 4.50.:** Desulfurization of donor-only PV. (A) Analytical HPLC profile of the educt PV<sup>1-108</sup>Cys<sup>78</sup>Xan<sup>88</sup>. The insert shows the mass of the main peak (arrow). (B) Analytical HPLC profile after the desulfurization reaction, yielding PV<sup>1-108</sup>Ala<sup>78</sup>Xan<sup>88</sup>. Mass spectra of the main peak (arrow) shown as insert.



## 4.4.2.2. Loop Formation in the Unfolded and an Intermediate State of PV

The effect of the triplet labels Xan and NAA on the stability of PV<sup>TTET</sup> was determined by an equilibrium urea transition (Figure 4.51). PV<sup>TTET</sup> unfolds cooperatively between 5 and



**Figure 4.51.:** Characterization of the secondary structure and stability of PV<sup>TTET</sup>. (A) Far UV CD spectrum at the indicated Ca $^{2+}$  concentrations. (B) Urea equilibrium transition at the indicated Ca $^{2+}$  concentrations. (C) Urea equilibrium transition in the PV double mutant at the indicated Ca $^{2+}$  concentrations. PMIB: polyvalent ion metal binding resin. All measurements were performed in 10 mM sodium cacodylic acid at pH 7 and 25  $^{\circ}$ C.

9 M urea at a calcium concentration of 100 mM ( $\circ$ ). The transition could be described by Equation 4.17 from Santoro and Bolen<sup>197</sup> on the basis of the linear extrapolation method<sup>39</sup>, which relates the mean ellipticity per amino acid  $\Theta_{MRW}$  with the denaturant concentration [D].

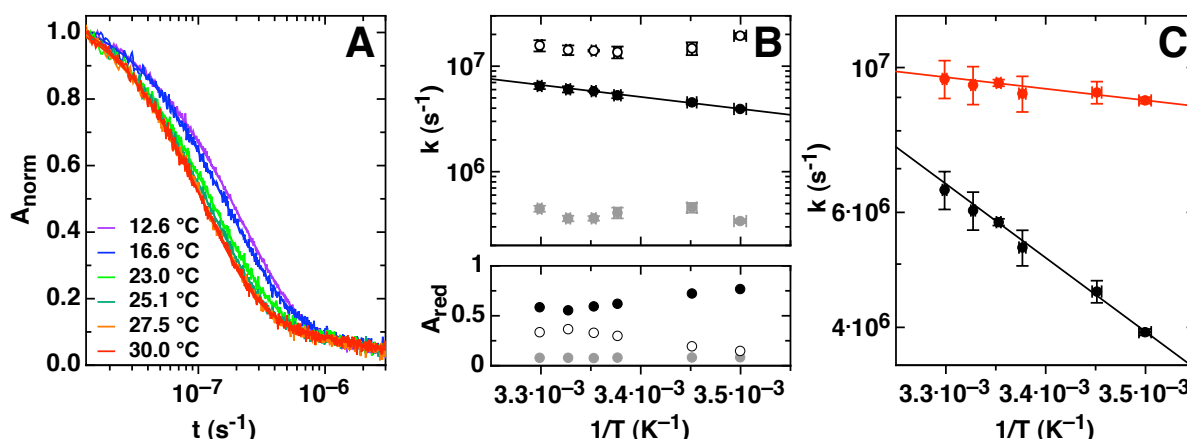
$$\Theta_{MRW} = \frac{(\Delta\Theta_N + m_N[D]) + (\Delta\Theta_U + m_U[D]) \cdot \exp(-(\Delta G_{N-U}^0/(RT) + m_{eq}[D]/(RT)))}{1 + \exp(-(\Delta G_{N-U}^0/(RT) + m_{eq}[D]/(RT)))} \quad (4.17)$$

$\Theta_N$  and  $\Theta_U$  represent the intercepts and  $m_N$  and  $m_U$  the slopes of pre- and postunfolding baselines, respectively.  $\Delta G_{N-U}^0$  and  $m_{eq}$  describe the intercept and slope of the linear extrapolation method. The fit yields  $\Delta G^0(\text{H}_2\text{O}) = 26.9 \pm 1.8$  kJ/mol and  $m_{eq} = 3.6 \pm 0.2$  kJ/mol/M with the denaturation midpoint at 7.5 M urea. In comparison to the double mutant, the values for  $m_{eq}$  and  $\Delta G^0(\text{H}_2\text{O})$  are smaller in the PV<sup>TTET</sup> variant. The triplet labels thus seem to affect the stability of the native state of PV. However, the stability of PV<sup>TTET</sup> is still high. In the pseudo wild-type, a cooperative urea transition was found even when no calcium was added to the sample solution (Figure 4.51 C). Chelating agents had to be used to remove calcium ions from PV. These findings prompted us to perform a urea transition in the PV<sup>TTET</sup>, where no calcium was added. As observed for the pseudo wild-type, PV<sup>TTET</sup> also unfolds cooperatively (Figure 4.51  $\bullet$ ). Similar to the double mutant, PV<sup>TTET</sup> binds Ca $^{2+}$  so strong, that even residual calcium traces are enough to stabilize the protein and thereby lead to a cooperative thermody-

#### 4. Results and Discussion

dynamic transition, which is in good agreement with reported strong calcium binding of PV<sup>147</sup>. In order to remove calcium from the protein and to populate an intermediate state, a polyvalent ion metal binding resin (PMIB) has to be used (Figure 4.51 C).

To investigate loop formation dynamics in the unfolded state of PV<sup>TTET</sup>, intramolecular TTET was measured in 8 M urea at different temperatures (Figure 4.52). To describe the time based



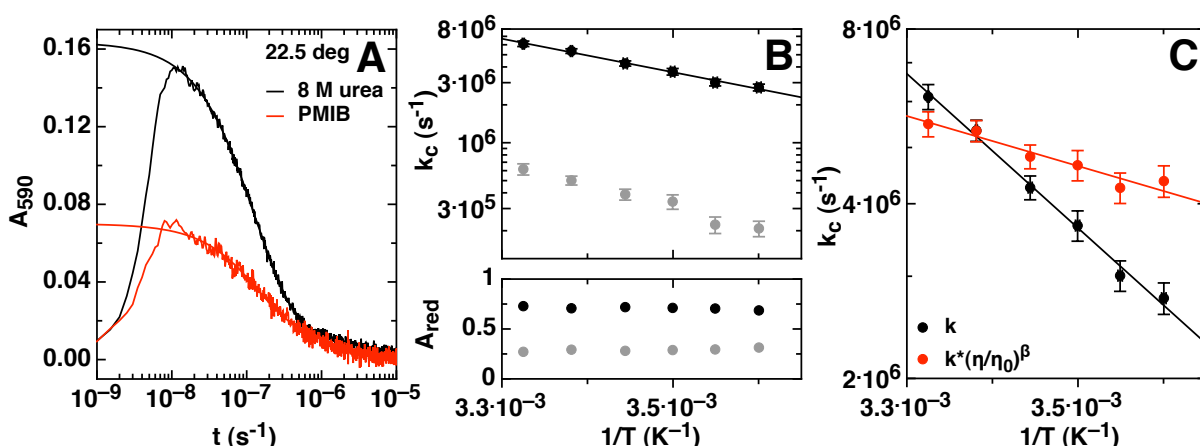
**Figure 4.52.:** Temperature dependence of chain dynamics in the unfolded state of PV. (A) Time based absorbance decay at 590 nm. (B) Rate constants and associated amplitudes as a function of temperature. (C) Non-corrected (●) and viscosity-corrected rate constants (●). The solid lines in B and C correspond to the Arrhenius Equation. All measurements were performed in 10 mM sodium cacodylic acid and 8 M urea at pH 7.

absorbance decay at 590 nm, a triple exponential fit was necessary. The rate constant of the main phase with an amplitude of 60 - 80% depends linearly on  $1/T$  and could be described by the Arrhenius equation (Figure 4.52 B, C). The viscosity-corrected rate constant was calculated with a  $\beta$ -value of -1 (Figure 4.52 C, ●) and could also be described by the Arrhenius equation (Equation 4.2) yielding  $E_A = 3.4 \pm 0.9 \text{ kJ/mol}$  and  $A = (3.7 \pm 1.4) \cdot 10^7 \text{ s}^{-1}$ . The value for the activation energy is unexpectedly small and comparable to that of long (GS) chains. The low  $E_A$  might be a hint, that the triplet labels can make contact with minor conformational rearrangements of the peptide backbone, due to the quite long linkers. An additional fast phase with 15 - 35% amplitude can be found in PV<sup>TTET</sup>, with its rate constant independent of temperature. This phase might correspond to a species, where residual structure is present and where the labels are closer together than in the stretched out chain, therefore leading to fast loop formation. We thus have evidence for residual structure in PV<sup>TTET</sup>, even under strong denaturing conditions. A third slow phase with 8% amplitude corresponds to the intrinsic decay of the xanthone triplet state, most probably due to small aggregates, in which xanthone

#### 4.4. Dynamics of Loop Formation in the EF-Hand and Carp $\beta$ -Parvalbumin

and naphthalene are unable to form contact.

In addition to kinetics in the unfolded state, we investigated the chain dynamics in a partially folded intermediate state of PV. According to observations from Kuwajima *et al.*<sup>149</sup>, this state was populated in aqueous buffer when calcium was removed by EDTA. Figure 4.53 shows the dynamics in the presence of resin bound EDTA (polyvalent metal ion binding resin). As

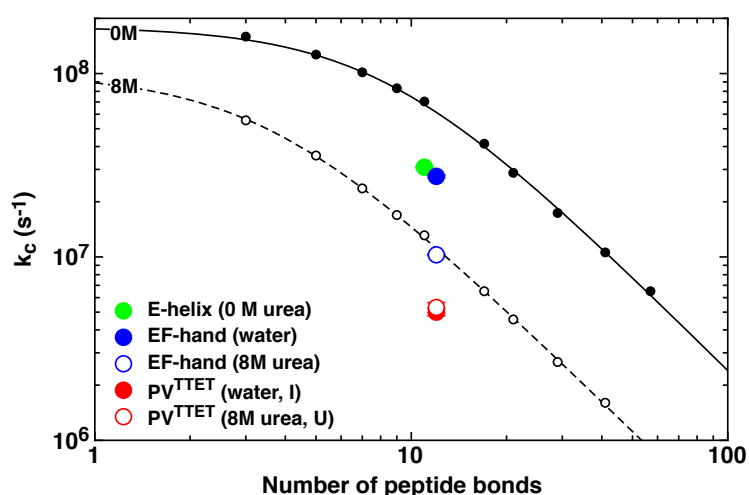


**Figure 4.53.:** Chain dynamics in an intermediate state of PV as a function of temperature. The intermediate state was populated by removing calcium from PV by chelators. The solid lines represent a fit of the Arrhenius Equation to the data. All measurements were performed in 10 mM sodium cacodylic acid pH 7.

for dynamics in the unfolded state of  $PV^{TTET}$ , the viscosity-corrected rate constants were calculated with a  $\beta$ -value of -1 and could be described with the Arrhenius equation (Equation 4.2) yielding  $E_A = 8.2 \pm 2.0$  kJ/mol and  $A = (1.5 \pm 1.2) \cdot 10^8$  s $^{-1}$ . Both parameters exhibit higher values in the intermediate state than in the unfolded state. The higher pre-exponential factor indicates a more compact conformation in I compared to U. This can be explained by a higher number of native intramolecular interactions in the intermediate state compared with the unfolded state. An additional fast phase was not found in the intermediate state of  $PV^{TTET}$ , although the amount of residual structure is expected to be higher in the intermediate state than in the unfolded state. However, a loss in amplitude was found for the dynamics in I compared to U. Therefore, a fast reaction occurs in the dead-time of the kinetic measurement. This reaction is addressed to the population of molecules, in which the labels are able to form contact by only a few bond rotations.

### 4.4.3. Comparison of Loop Formation Dynamics in PV Fragments and Full Length PV

Loop formation kinetics in the E-helix, the EF-hand and  $PV^{TTET}$  are compared, to assess the contributions of long range interactions to chain dynamics in proteins. The rate constants for loop formation from TTET in presence or absence of 8 M urea are shown in Figure 4.54, in comparison with the dynamics of  $(GS)_n$ -peptides in water ( $\bullet$ ) and in the presence of 8 M denaturant ( $\circ$ ). The fastest loop formation kinetics are observed for the E-helix and the EF-



**Figure 4.54.:** Loop formation dynamics in the E-helix, the EF-hand and  $PV^{TTET}$  as a function of interchromophore peptide bonds at different conditions.

hand in water. In  $PV^{TTET}$ , the loop formation is significantly slowed down by a factor of 5 in comparison to the EF-hand. The additional tail in  $PV^{TTET}$  is predicted to slow down loop formation only by a factor of 1.7. Therefore, long range interactions in the intermediate state are held responsible for the additional effect. In the presence of denaturants, the kinetics for loop formation in  $PV^{TTET}$  are slowed down only by a factor of 1.9. This value is close to the predicted value of 1.7, accounting for the additional tail. In the unfolded state, long range interactions are absent and therefore do not influence loop formation.

In  $PV^{TTET}$ , the rate constant for loop formation in the intermediate was found to be identical with the rate constant of the main kinetic phase in U. It is assumed, that this rate constant arises from a mainly unfolded subpopulation in the intermediate state. The additional fast phase in U and the loss in amplitude for loop formation in the intermediate state indicate, that a partially folded species exists, where the E-helix is formed and the labels may form contact by a few

#### 4.4. Dynamics of Loop Formation in the EF-Hand and Carp $\beta$ -Parvalbumin

---

bond rotations.

Using TTET to monitor the dynamics in  $PV^{TTET}$  is shown to be perfectly suited to identify sub-states, determine their chain dynamics and thereby assess the contributions of long range interactions in the different states.



## 5. Summary

During protein folding, the polypeptide chain explores the conformational space to form intramolecular native-like interactions, which result in loop formation. In this work, we used triplet-triplet energy transfer (TTET) to monitor loop formation kinetics in model peptides, natural sequences and a protein. In TTET, the triplet energy from the donor xanthone is transferred to the acceptor naphthalene in a diffusion controlled process, when the chromophores are in Van der Waals contact. This allowed us to measure absolute rate constants for loop formation in polypeptides on a timescale between hundreds of picoseconds and several microseconds.

TTET was used earlier to measure loop formation in poly(Gly-Ser) homopolymers ((GS)-peptides). Different regimes for loop formation have been found in short and long (GS)-peptides. In short peptides, chain stiffness is high, while long peptides behave like statistical chains. In this work, we investigated loop formation in (GS)<sub>n</sub>-peptides with  $n = 1, 8$  and  $16$ , which serve as simple models for the polypeptide backbone. To determine the effect of amino acid side chains on the loop formation dynamics and the associated activation energy, we compared the (GS)-peptides with fragments from naturally occurring proteins. These fragments were derived from carp  $\beta$ -parvalbumin (PV) and from the B1 domain of protein G. In the two PV fragments representing the loops between helices D-E (PV-DE-loop) and helices E-F (PV-EF-loop) as well as in the B1 hairpin (GB1-hairpin), loop formation is slower and the activation energy higher when compared with (GS)-peptides of identical length. We attribute this effect to the lower Gly content and the presence of large side chains in the natural sequences to be responsible for the observed effects.

The properties of a solvent determine the dimensions and dynamics of a polypeptide chain. The solvent can be either good or poor, depending on the strength of interactions within the peptide in comparison with the strength of peptide-solvent interactions. In good solvents, peptide-solvent interactions are more favorable, while intramolecular interactions are more

## 5. Summary

---

favorable in poor solvents. Osmolytes and other co-solutes exist, which are able to modulate solvent quality, to tune the strength of peptide intramolecular interactions and thereby stabilize, destabilize or do not affect the native state. To test the influence of solvent quality on barriers and dynamics for loop formation within a single state, we measured TTET in polypeptide chains in the presence of the destabilizing co-solutes urea and guanidinium chloride (GdmCl), the stabilizing osmolytes trimethylamine *N*-oxide (TMAO) and sarcosine, and the neutral co-solutes arginine and proline. The effect of amino acid side chains on loop formation in good, poor and neutral solvents was determined by comparing (GS)<sub>1</sub>, (GS)<sub>8</sub> and (GS)<sub>16</sub> with the natural sequences in the respective co-solute. The results show, that all investigated osmolytes and co-solutes slow down loop formation both in the model peptides and in the natural sequences. The addition of co-solutes significantly increases solvent viscosity, which additionally affects the chain dynamics. When correcting for the effect of increased solvent viscosity caused by the addition of co-solutes, a co-solute specific effect was found which is different for stabilizers, denaturants and neutral osmolytes. Independent of the amino acid sequence, loop formation is accelerated in all peptides in the presence of stabilizers. At high TMAO or sarcosine concentrations, the viscosity-corrected rate constant for contact formation asymptotically approaches a limiting value. In the presence of GdmCl or urea, the viscosity-corrected rate constant for loop formation is decelerated with increasing denaturant concentration and approaches a limiting value at high concentrations. This indicates, that the co-solute specific effect becomes saturated at high concentration of stabilizing and destabilizing agents. In the presence of the neutral osmolyte proline, the viscosity-corrected rate constant for loop formation only changes slightly compared to water. The magnitude of the effects of the different co-solutes on loop formation were similar in the (GS)-peptides and the natural sequences. Thus, the co-solute specific effect is exerted mainly on the peptide backbone. Additionally, we measured the effect of different osmolytes and co-solutes on barriers for loop formation. Loop formation was measured in the temperature range between 5 °C and 30 °C, to determine the activation parameters in the presence of stabilizing or destabilizing co-solutes. The effect of temperature on the corrected rate constant can be described by the Arrhenius Equation. The activation energy and the Arrhenius pre-exponential factor associated with loop formation are hardly affected in the presence of denaturants when compared with water. Stabilizing co-solutes however increase both parameters significantly, independent of the peptide sequence.

Two models were used to describe the effect of the various co-solutes on loop formation kinetics. In the exchange formalism from Schellman, co-solutes are assumed to bind weakly



---

to the polypeptide chain. Stabilizing co-solutes yield negative  $\gamma$ -values, while destabilizing co-solutes yield  $\gamma \geq 0$ . An equilibrium constant for the exchange of a water molecule with a co-solute molecule at a binding site was determined. The exchange constant is larger for the more potent denaturant GdmCl, when compared with urea and larger in the stronger stabilizer TMAO, when compared with sarcosine. Using Tanford's Transfer model, a correlation between the calculated transfer free energy for the peptides to the respective co-solute and the viscosity-corrected rate constant for loop formation at the respective condition can be found. For denaturants, favorable transfer free energies were calculated for all peptides and decelerated loop formation was observed. For stabilizing osmolytes, an unfavorable free energy for the transfer of the peptides from water to 1 M osmolyte was calculated. Accordingly, loop formation is accelerated at these conditions. For the neutral osmolyte proline, the transfer free energy is small and the viscosity-corrected rate constant for contact formation changes only slightly compared to water. The correlation between transfer free energies and corrected loop formation rate constants allows us to predict the effect of different osmolytes and other co-solutes on loop formation dynamics in any amino acid sequence.

The effect of osmolytes and other co-solutes is summarized in the following. Loop formation in stabilizing osmolytes and co-solutes is accelerated, due to a higher Arrhenius pre-exponential factor in comparison with water. Additionally, the activation energy associated with loop formation is increased in poor solvents due to the formation of intramolecular interactions, which restrict the conformational space. Since the observed effect is independent of the peptide sequence, we assume that these interactions must be hydrogen bonds formed of the peptide backbone with itself or with the solvent. The kinetics for loop formation become slower in denaturants in comparison to water, where intramolecular hydrogen bonds are disrupted and the polypeptide chain is more stretched out.

Chain diffusion in polypeptides is subject to solvent viscosity ( $\eta$ ) but might be affected by contributions from the viscosity-independent internal friction. It was suggested, that in addition to solvent friction, contributions from internal friction affect the relaxation of a polymer chain towards its equilibrium and could be determined by kinetic measurements at different solvent viscosities through extrapolation to zero solvent viscosity. Up to date, it is not clear whether a linear, a power law or a Rouse model should be used for the extrapolation of  $\eta \rightarrow 0$ . Loop formation was measured in model peptides and natural sequences at different solvent viscosities to test, which model is able to describe the contributions from internal friction ( $\tau_{\text{int}}$ ) and from solvent friction ( $\tau_{\text{solv}}$ ) and what effect side chains have. Our results show, that

## 5. Summary

---

a power law with a fractional exponent  $\beta$  (with  $\beta$  describing the sensitivity of chain dynamics to solvent viscosity) is suited best to describe the time constant for loop formation in (GS)<sub>8</sub> and the natural sequences at different solvent viscosities. We used the determined  $\beta$ -value at the respective condition to determine  $\tau_{\text{int}}$  by extrapolation to zero solvent viscosity. We found, that contribution of internal friction to chain dynamics are low, on the sub-nanosecond time scale both in good and poor solvents, independent the amino acid sequence and temperature. In contrast to  $\tau_{\text{int}}$ ,  $\tau_{\text{solv}}$  depends on the amino acid sequence and temperature. Contributions from solvent friction are lower in (GS)<sub>8</sub> in comparison with the PV-DE-loop and the PV-EF-loop. To test the effect of poor or good solvents on  $\tau_{\text{solv}}$ , we measured loop formation in the PV-EF-loop in the presence of urea or TMAO at varying solvent viscosities. We find, that contributions from solvent friction to loop formation are high in good solvents, low in poor solvents and intermediate in water. Using the Arrhenius Equation, we could describe  $1/\tau_{\text{solv}}$  as a function of  $1/T$ . The activation energy associated with  $1/\tau_{\text{solv}}$  of the PV-EF-loop was found to be identical for all three solvents.

The last part of this thesis aimed at determining the effect of long range interactions on the dynamics of a helix in different states of a protein. To test the effect of long range interactions on chain dynamics, we measured loop formation of the E-helix in protein fragments derived from the calcium binding protein carp  $\beta$ -parvalbumin (PV) and in full length PV. PV was chosen to be studied by TTET, since it lacks amino acids interfering with the triplet state of xanthone and exhibits a thermodynamical intermediate in the absence of calcium. Using native chemical ligation, we successfully produced full length parvalbumin by condensing an expressed fragment 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. The dynamics for loop formation in the doubly labeled full length PV were measured by TTET in the unfolded state and an intermediate state and compared with an EF-hand peptide and the isolated E-helix of PV. In the intermediate state of PV, loop formation in the E-helix is slowed down due to long range interactions. In the unfolded state of PV, long range interactions are inexistent and therefore do not influence the kinetics for loop formation. Using TTET, we gained evidence for residual helical structure in a subpopulation of molecules in the intermediate and the unfolded state, where the E-helix is formed and loop formation can occur by only a few bond rotations.

## A. Appendix

### Parameters for the Calculation of the Transfer Free Energy

**Table A.1.:** Solvent accessibilities for amino acids<sup>102</sup> and tripeptides<sup>170</sup>

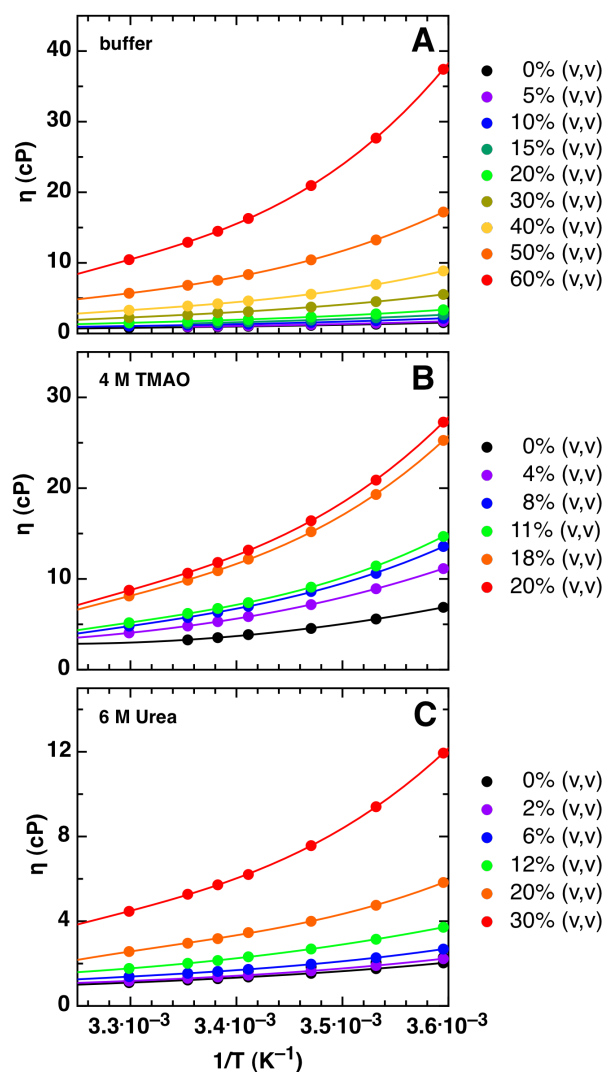
residue	ASA (Å/residue)		ASA <sub>Gly-X-Gly</sub> (Å/residue)	
	backbone	sidechain	backbone	sidechain
Ala	27.9	55.1	46.2	71.9
Arg	25.1	171.1	39.1	216.9
Asn	25.2	90.1	40.2	125.3
Asp	26.0	87.0	40.5	118.2
Cys	26.4	73.0	42.6	103.5
Gln	25.3	116.9	37.8	155.4
Glu	25.7	113.4	37.8	148.4
Gly	65.2	0.0	88.1	0.0
His	24.2	111.5	40.4	162.1
Ile	20.0	117.1	30.9	150.1
Leu	22.7	109.6	35.3	157.8
Lys	26.1	150.7	38.7	187.1
Met	25.3	122.4	38.6	164.8
Phe	24.3	128.8	38.4	184.4
Pro	22.5	87.0	35.6	111.0
Ser	29.4	66.5	44.0	85.8
Thr	24.1	84.3	37.9	114.6
Trp	23.6	156.6	37.4	228.9
Tyr	25.6	141.7	38.7	148.4
Val	20.4	96.4	36.1	128.4

**Table A.2.:** Group transfer free energies in J/mol.

Side-chain and backbone unit	TMAO	Sarcosine	Proline	Urea	L-ArgHCl
Ala	-61.25	45.65	-0.29	-19.62	107.06
ArgHCl	-457.19	-134.89	-251.79	-88.58	215.89
Asn	233.01	-171.25	-74.10	-162.30	-697.88
NaAsp	-278.95	-59.41	-378.69	14.85	-365.57
Cys	-	-	-	-	-
Gln	173.26	-42.63	-134.98	-229.33	-605.28
NaGlu	-348.32	-52.76	-373.09	2.59	-291.67
Gly	0	0	0	0	0
His	176.02	-87.03	-188.70	-211.33	-897.06
Ile	-106.40	167.28	-11.38	-160.79	-472.58
Leu	48.62	160.37	19.96	-228.32	-366.20
LysHCl	-461.20	-114.73	-250.50	-95.23	181.35
Met	-32.01	34.23	-146.94	-202.25	-508.10
Phe	-38.99	-52.89	-298.15	-347.73	-1042.89
Pro	-576.26	-143.22	-267.61	-73.85	-44.39
Ser	-163.34	-117.07	-140.12	-86.02	-102.51
Thr	14.94	-31.55	-76.69	-92.42	-314.08
Trp	-639.61	-472.92	-829.98	-591.87	-2136.26
Tyr	-478.31	-110.33	-579.11	-188.61	-1634.86
Val	-4.27	122.67	33.30	-90.58	-353.45
Backbone	376.56	217.57	200.83	-163.18	494.43

## Solvent Viscosity as a Function of Temperature

Solvent viscosity measurement as a function of temperature (Figure A.1). The data was fitted with a third degree polynomial function of the function  $y = c + a_1x + a_2x^2 + a_3x^3$ . The results from the fit are summarized in Table A.3.



**Figure A.1.:** Measured solvent viscosities as a function of Temperature. A) 10 mM phosphate buffer pH 7, B) 10 mM phosphate buffer pH 7 and 4 M TMAO, C) 10 mM phosphate buffer pH 7 and 6 M urea.

**Table A.3.:** Fit parameters of solvent viscosities as a function of temperature.

10 mM phosphate buffer; pH 7				
Glycerol (% (v,v))	c	a <sub>1</sub>	a <sub>2</sub>	a <sub>3</sub>
0	$-1.5333 \cdot 10^4$	$1.3988 \cdot 10^7$	$-4.2635 \cdot 10^9$	$4.3445 \cdot 10^{11}$
5	$-3.0898 \cdot 10^2$	$2.8127 \cdot 10^5$	$-8.5669 \cdot 10^7$	$8.7536 \cdot 10^9$
10	$-4.4653 \cdot 10^3$	$4.1250 \cdot 10^6$	$-1.2740 \cdot 10^9$	$1.3168 \cdot 10^{11}$
15	$-2.4723 \cdot 10^2$	$2.3454 \cdot 10^5$	$-7.4514 \cdot 10^7$	$7.9574 \cdot 10^9$
20	$-3.0059 \cdot 10^3$	$2.7324 \cdot 10^6$	$-8.3002 \cdot 10^8$	$8.4351 \cdot 10^{10}$
30	$-6.5260 \cdot 10^2$	$5.9582 \cdot 10^5$	$-1.8207 \cdot 10^8$	$1.8663 \cdot 10^{10}$
40	$-3.4482 \cdot 10^2$	$3.2739 \cdot 10^5$	$-1.0406 \cdot 10^8$	$1.1106 \cdot 10^{10}$
50	$-2.1874 \cdot 10^2$	$2.0540 \cdot 10^5$	$-6.4608 \cdot 10^7$	$6.8332 \cdot 10^9$
60	$-2.3418 \cdot 10^2$	$2.1243 \cdot 10^5$	$-6.4535 \cdot 10^7$	$6.5878 \cdot 10^9$
10 mM phosphate buffer; pH 7; 6 M urea				
Glycerol (% (v,v))	c	a <sub>1</sub>	a <sub>2</sub>	a <sub>3</sub>
0	-182.89	$1.6984 \cdot 10^5$	$-5.2871 \cdot 10^7$	$5.5455 \cdot 10^9$
2	-194.90	$1.8514 \cdot 10^5$	$-5.8805 \cdot 10^7$	$6.2748 \cdot 10^9$
6	-430.96	$3.9453 \cdot 10^5$	$-1.2082 \cdot 10^8$	$1.2415 \cdot 10^{10}$
12	-111.05	$1.1880 \cdot 10^5$	$-4.2123 \cdot 10^7$	$4.9950 \cdot 10^9$
20	-2080.6	$1.8630 \cdot 10^6$	$-5.5747 \cdot 10^8$	$5.5828 \cdot 10^{10}$
30	-3503.1	$3.1996 \cdot 10^6$	$-9.7689 \cdot 10^8$	$9.9818 \cdot 10^{10}$
10 mM phosphate buffer; pH 7; 4 M TMAO				
Glycerol (% (v,v))	c	a <sub>1</sub>	a <sub>2</sub>	a <sub>3</sub>
0	322.52	$-1.9777 \cdot 10^5$	$3.0590 \cdot 10^7$	-
4	-906.02	$9.0229 \cdot 10^5$	$-2.9982 \cdot 10^8$	$3.3324 \cdot 10^{10}$
8	-3830.0	$3.4888 \cdot 10^6$	$-1.0631 \cdot 10^9$	$1.0850 \cdot 10^{11}$
11	-4905.1	$4.4637 \cdot 10^6$	$-1.3577 \cdot 10^9$	$1.3818 \cdot 10^{11}$
18	-8865.2	$8.0655 \cdot 10^6$	$-2.4530 \cdot 10^9$	$2.4963 \cdot 10^{11}$
20	-9439.8	$8.5885 \cdot 10^6$	$-2.6124 \cdot 10^9$	$2.6590 \cdot 10^{11}$

## List of Figures

1.1. Structure of the peptide backbone . . . . .	4
1.2. Hypothetical free energy surface for two state folding . . . . .	7
1.3. Thermodynamic cycle of the Transfer Model. . . . .	13
1.4. Triplet-triplet energy transfer between xanthone and naphthalene . . . . .	18
1.5. Triplet-triplet energy transfer in an ensemble of unstructured peptides . . . . .	18
1.6. Length dependence of loop formation in glycine-serine repeat peptides . . . . .	20
1.7. Loop formation rate constants in host-guest peptides . . . . .	21
1.8. Probing local dynamics and stability in $\alpha$ -helical peptides . . . . .	23
1.9. Structure of carp $\beta$ -parvalbumin . . . . .	24
1.10. Native chemical ligation . . . . .	27
3.1. Cloning of PV <sup>1-77</sup> into the pTXB3 vector . . . . .	37
3.2. Labeling of cysteine with BNAA . . . . .	40
4.1. Chemical structure of poly(Gly-Ser) peptides. . . . .	46
4.2. Loop formation kinetics in poly(Gly-Ser) peptides. . . . .	46
4.3. Effect of solvent viscosity on loop formation in (GS) <sub>8</sub> . . . . .	47
4.4. Chain Dynamics of (GS) <sub>16</sub> at Different Temperatures. . . . .	48
4.5. Activation energy, Arrhenius pre-exponential factor and activation entropy in homo-polypeptides of different length. . . . .	50
4.6. Peptide fragments from naturally occurring proteins. . . . .	52
4.7. Loop formation kinetics in natural sequences. . . . .	53
4.8. Intramolecular Loop formation in natural sequences . . . . .	54
4.9. Loop formation dynamics as a function of temperature. . . . .	56
4.10. Chemical structure of co-solutes. . . . .	59
4.11. Chain dynamics at various denaturant concentrations. . . . .	61
4.12. Non-corrected and viscosity-corrected rate constants for loop formation as a function of the molar denaturant concentration and the denaturant mole fraction	63

## List of Figures

---

4.13. Viscosity-corrected rate constant for loop formation as a function of the transfer free energy from water to 1 M urea . . . . .	65
4.14. Non-corrected and viscosity-corrected rate constants for loop formation as a function of temperature in the presence of 8 M urea (left) or 6 M GdmCl (right). . . . .	66
4.15. Decay of the triplet state in donor-only peptides in the presence of TMAO. . . . .	70
4.16. Decay of the triplet state in donor-only peptides in the presence of sarcosine. . . . .	71
4.17. Chain dynamics at various TMAO concentrations. . . . .	72
4.18. Chain dynamics at various sarcosine concentrations. . . . .	73
4.19. Non-corrected and viscosity-corrected rate constants for loop formation as a function of the molar protectant concentration and the protectant mole fraction. . . . .	75
4.20. Viscosity-corrected rate constant for loop formation as a function of the transfer free energy from water to 1 M TMAO and sarcosine . . . . .	76
4.21. Rate constant for loop formation as a function of the transfer free energy from 0-6 M sarcosine. . . . .	78
4.22. Effect of temperature on loop formation dynamics in the presence of stabilizing osmolytes. . . . .	79
4.23. Effect of arginine on the Xan triplet state. . . . .	82
4.24. Chain dynamics at various proline concentrations. . . . .	83
4.25. Chain dynamics at various arginine concentrations. . . . .	85
4.26. Loop formation rate constants in different arginine and proline concentrations. . . . .	86
4.27. Viscosity-corrected rate constant for loop formation as a function of the transfer free energy from water to 1 M proline and 1 M arginine . . . . .	87
4.28. Effect of temperature on loop formation dynamics in the presence of the neutral osmolyte arginine. . . . .	88
4.29. Viscosity-corrected rate constant for loop formation as a function of the co-solute concentration. . . . .	91
4.30. Comparison of the relative rate constants for loop formation $k'_{c,rel}$ in the presence of stabilizing co-solutes. . . . .	93
4.31. Viscosity-corrected rate constant for loop formation $k'_c$ as a function of transfer free energy $\Delta G_{U;0 \rightarrow 1M}^{0;tr}$ from water to 1 M co-solute. . . . .	95
4.32. Summarized Arrhenius parameters for loop formation in the presence of the indicated co-solute. . . . .	98
4.33. Activation energy $E_A$ for loop formation as a function of the solvent viscosity sensitivity $\beta$ . . . . .	99



---

4.34. Effect of viscosifier size on chain dynamics in (GS) <sub>1</sub> . . . . .	103
4.35. Effect of viscosity on chain dynamics. . . . .	105
4.36. Effect of temperature on $\beta$ . . . . .	106
4.37. Comparison between linear fit and power law. . . . .	107
4.38. Temperature dependent determination of $\tau_{\text{int}}$ and $\tau_{\text{solv}}$ . . . . .	108
4.39. Influence of peptide length on the time constant for loop formation as a function of solvent viscosity. . . . .	111
4.40. Sequence of C-terminal PV fragments with their location indicated in the full length protein. . . . .	113
4.41. Chain dynamics of the E-helix of PV at different glycerol/water mixtures. . .	113
4.42. Chain dynamics in the PV E-helix as a function of the TFE concentration. . .	114
4.43. Loop formation dynamics in EF-hand peptides. . . . .	115
4.44. PV variant for TTET measurements. . . . .	116
4.45. Native chemical ligation (NCL) of a peptide thioester with a synthesized fragment. . . . .	118
4.46. Analysis of prepared PV <sup>1-77</sup> $\alpha$ -carboxyl MESNA thioester. . . . .	119
4.47. Analysis of synthesized PV <sup>78-108</sup> . . . . .	119
4.48. Analysis of the native chemical ligation. . . . .	120
4.49. Labeling of the Cys sidechain with the thiol reactive naphthyl derivative BNAA. .	121
4.50. Desulfurization of the donor-only PV. . . . .	122
4.51. Characterization of the secondary structure and stability of PV <sup>TTET</sup> . . . . .	123
4.52. Temperature dependence of chain dynamics in the unfolded state of PV. . . .	124
4.53. Chain dynamics in an intermediate state of PV. . . . .	125
4.54. Loop formation dynamics in the E-helix, the EF-hand and PV <sup>TTET</sup> as a function of interchromophore peptide bonds at different conditions. . . . .	126
A.1. Measured solvent viscosities. . . . .	135



# List of Tables

1.1. Amino acids interacting with the triplet state of xanthone . . . . .	19
3.1. Solid supports used for SPPS . . . . .	32
3.2. Mutagenic primer sequences . . . . .	34
3.3. PCR temperature program . . . . .	35
3.4. List of Buffers for Protein Thioester Formation . . . . .	37
3.5. SDS-PAGE gel casting . . . . .	39
4.1. Parameters for the temperature dependence of loop formation in (GS) <sub>16</sub> . . . . .	49
4.2. Amino acid sequences of investigated peptides . . . . .	53
4.3. Parameters for the viscosity dependence of chain dynamics. . . . .	55
4.4. Arrhenius parameters for the indicated peptide in water. . . . .	56
4.5. Denaturant dependent fitting parameters of Equation 4.6. . . . .	62
4.6. Denaturant Parameters for Schellman's weak binding model . . . . .	64
4.7. Arrhenius parameters for the indicated peptide in the presence of the destabilizing co-solutes GdmCl and urea. . . . .	67
4.8. Parameters for Schellman's weak binding model for protecting osmolytes . . . . .	75
4.9. Arrhenius parameters for the indicated peptide in the presence of stabilizing osmolytes in comparison to water. . . . .	80
4.10. Arrhenius parameters for the indicated peptide under different solvent compositions. . . . .	89
4.11. Parameters for Schellman's weak binding model . . . . .	92
4.12. Arrhenius parameters for the indicated peptide under different solvent compositions. . . . .	97
4.13. Arrhenius parameters for the temperature dependence of $\tau_{\text{solv}}$ . . . . .	110
A.1. Solvent accessibilities . . . . .	133
A.2. Group transfer free energies . . . . .	134

A.3. Fit parameters of solvent viscosities as a function of temperature. . . . . 136

# List of Abbreviations

**Asp** Aspartate

**BB** Backbone

**BNAA** 2-bromo-*N*-(1-naphthyl)acetamide

**Boc** Tert-butyloxycarbonyl

**Ca<sup>2+</sup>** Calcium

**CD** Circular dichroism

**Cos** Co-solute

**Cys** Cysteine

**Dpr**  $\alpha$ -L-diaminopropionic acid

**DTT** Dithiothreitol

**Fmoc** 9-fluorenylmethoxycarbonyl

**FRET** Förster resonance energy transfer

**GdmCl** Guanidinium chloride

**Gly** Glycine

## List of Abbreviations

---

**His** Histidine

**HPLC** High performance liquid chromatography

**IPTG**  $\beta$ -D-1-thiogalactopyranose

**MESNA**  $\beta$ -mercaptosulfonic acid

**Met** Methionine

**N** Native state

**NAA** *N*-(naphthalen-1-yl)acetamide

**Nal** L-1-Naphthylalanine

**NCL** Native chemical ligation

**NMR** Nuclear magnetic resonance

**PAGE** Polyacrylamide gel electrophoresis

**Phe** Phenylalanine

**pI** Isoelectric point

**PV** carp  $\beta$ -parvalbumin

**REFER** Rate equilibrium free energy relationship

**SASA** Solvent accessible surface area

**SC** Side chain

---

**SDS** Sodium dodecyl sulfate

**SPPS** Solid phase peptide synthesis

**Thr** Threonine

**TMAO** Trimethylamine *N*-oxide

**Trp** Tryptophane

**TTET** Triplet-triplet energy transfer

**Tyr** Tyrosine

**U** Unfolded state

**Xan** Xanthonic acid





## Bibliography

- [1] Fersht, A. R. *Structure and Mechanism in Protein Science* (W. H. Freeman and Company, New York, 1999). 1
- [2] Anfinsen, C. B. Principles that govern the folding of protein chains. *Science* **181**, 223–230 (1973). 1, 11
- [3] Kuwajima, K. K., Nitta, K., Yoneyama, M. & Sugai, S. Three-state denaturation of alpha-lactalbumin by guanidine hydrochloride. *Journal of Molecular Biology* **106**, 359–373 (1976). 1, 11
- [4] Hughson, F., Wright, P. & Baldwin, R. L. Structural characterization of a partly folded apomyoglobin intermediate. *Science* **249**, 1544–1548 (1990).
- [5] Pfeil, W., Noelting, B. O. & Jung, C. Apocytochrome P450cam is a native protein with some intermediate-like properties. *Biochemistry* **32**, 8856–8862 (1993). PMID: 8364032. 1
- [6] Messens, J. & Collet, J.-F. Pathways of disulfide bond formation in *Escherichia coli*. *The International Journal of Biochemistry and Cell Biology* **38**, 1050 – 1062 (2006). 1, 11
- [7] Lang, K., Schmid, F. X. & Fischer, G. Catalysis of protein folding by prolyl isomerase. *Nature* **329**, 268–270 (1987). 1, 11
- [8] Prusiner, S. B. Prions. *Proceedings of the National Academy of Sciences* **95**, 13363–13383 (1998). 1
- [9] Dyson, H. J. & Wright, P. E. Intrinsically unstructured proteins and their functions. *Nature Reviews Molecular Cell Biology* **6**, 197–208 (2005). 1
- [10] Perutz, M. *et al.* Structure of haemoglobin: a three-dimensional fourier synthesis at 5.5Å resolution, obtained by X-ray analysis. *Nature* **185** (1960). 1

- [11] Wüthrich, K. The development of nuclear magnetic resonance spectroscopy as a technique for protein structure determination. *Accounts of Chemical Research* **22**, 36–44 (1989). 2
- [12] Clore, G. M. & Gronenborn, A. M. Two-, three-, and four-dimensional NMR methods for obtaining larger and more precise three-dimensional structures of proteins in solution. *Annual Review of Biophysics and Biophysical Chemistry* **20**, 29–63 (1991).
- [13] Wüthrich, K. *NMR of Proteins and Nucleic Acids* (J. Wiley, New York, USA, 1986). 2
- [14] Pauling, L. & Corey, R. B. Configurations of polypeptide chains with favored orientations around single bonds. *Proceedings of the National Academy of Sciences* **37**, 729–740 (1951). 2
- [15] Pauling, L. & Corey, R. B. Atomic coordinates and structure factors for two helical configurations of polypeptide chains. *Proceedings of the National Academy of Sciences* **37**, 235–240 (1951). 2
- [16] Privalov, P. L. & Gill, S. J. Stability of protein structure and hydrophobic interaction. vol. 39 of *Advances in Protein Chemistry*, 191 – 234 (Academic Press, 1988). 2
- [17] Baldwin, R. L. Energetics of protein folding. *Journal of Molecular Biology* **371**, 283 – 301 (2007). 2
- [18] Wetlaufer, D. B. Nucleation, rapid folding, and globular intrachain regions in proteins. *Proceedings of the National Academy of Sciences* **70**, 697–701 (1973). 2, 6
- [19] Opitz, C. A. *et al.* Damped elastic recoil of the titin spring in myofibrils of human myocardium. *Proceedings of the National Academy of Sciences* **100**, 12688–12693 (2003). 2
- [20] Labeit, S. & Kolmerer, B. Titins: Giant proteins in charge of muscle ultrastructure and elasticity. *Science* **270**, 293–296 (1995). 2
- [21] Pain, R. H. *Mechanisms of Protein Folding* (Oxford University Press, Oxford, United Kingdom, 1994). 2
- [22] Cammon, J. A. M. & Harvey, S. C. *Dynamics of Proteins and Nucleic Acids* (Cambridge University Press, 1988). 2

- 
- [23] Ansari, A. *et al.* Protein states and proteinquakes. *Proceedings of the National Academy of Sciences* **82**, 5000–5004 (1985). 2
- [24] Creighton, T. E. (ed.) *Protein Folding* (W. H. Freeman and Company, New York, 1992). 2, 3
- [25] Tanford, C., Pain, R. H. & Otchin, N. S. Equilibrium and kinetics of the unfolding of lysozyme (muramidase) by guanidine hydrochloride. *Journal of Molecular Biology* **15**, 489–504 (1966). 3
- [26] Ginsburg, A. & Carroll, W. R. Some specific ion effects on the conformation and thermal stability of ribonuclease. *Biochemistry* **4**, 2159–2174 (1965). 3
- [27] Brown, J. E. & Klee, W. A. Helix-coil transition of the isolated amino terminus of ribonuclease. *Biochemistry* **10**, 470–476 (1971). 3
- [28] Blanco, F. J. *et al.* Tendamistat (12-26) fragment. NMR characterization of isolated beta-turn folding intermediates. *European Journal of Biochemistry* **200**, 345–351 (1991).
- [29] Prat Gay, G. d. & Fersht, A. R. Generation of a family of protein fragments for structure-folding studies. 1. folding complementation of two fragments of chymotrypsin inhibitor-2 formed by cleavage at its unique methionine residue. *Biochemistry* **33**, 7957–7963 (1994).
- [30] Alexandrescu, A. T., Abeygunawardana, C. & Shortle, D. Structure and dynamics of a denatured 131-residue fragment of staphylococcal nuclease: A heteronuclear NMR study. *Biochemistry* **33**, 1063–1072 (1994).
- [31] Tang, Y., Rigotti, D. J., Fairman, R. & Raleigh, D. P. Peptide models provide evidence for significant structure in the denatured state of a rapidly folding protein: The villin headpiece subdomain. *Biochemistry* **43**, 3264–3272 (2004).
- [32] Krieger, F. *et al.* Intrachain diffusion in a protein loop fragment from carp parvalbumin. *Chemical Physics* **307**, 209–215 (2004). 3, 19, 25, 117
- [33] Bieri, O. *et al.* The speed limit for protein folding measured by triplet-triplet energy transfer. *Proceedings of the National Academy of Sciences of the United States of America* **96**, 9597–601 (1999). 3, 15, 16

- [34] Krieger, F., Fierz, B., Bieri, O., Drewello, M. & Kiefhaber, T. Dynamics of unfolded polypeptide chains as model for the earliest steps in protein folding. *Journal of Molecular Biology* **332**, 265–274 (2003). 8, 16, 17, 18, 20, 21, 22, 45, 54, 60
- [35] Möglich, A., Krieger, F. & Kiefhaber, T. Molecular basis for the effect of urea and guanidinium chloride on the dynamics of unfolded polypeptide chains. *Journal of Molecular Biology* **345**, 153–62 (2005). 22, 60, 62, 63
- [36] Möglich, A., Joder, K. & Kiefhaber, T. End-to-end distance distributions and intrachain diffusion constants in unfolded polypeptide chains indicate intramolecular hydrogen bond formation. *Proceedings of the National Academy of Sciences of the United States of America* **103**, 12394–9 (2006). 15, 96, 100
- [37] Fierz, B. *et al.* Loop formation in unfolded polypeptide chains on the picoseconds to microseconds time scale. *Proceedings of the National Academy of Sciences of the United States of America* **104**, 2163–2168 (2007). 3
- [38] 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* **437**, 1053–1056 (2005). 3
- [39] Tanford, C. Protein denaturation. *Advances in Protein Chemistry* **23**, 121–282 (1968). 3, 12, 14, 123
- [40] Tanford, C. Protein denaturation. C. Theoretical models for the mechanism of denaturation. *Advances in Protein Chemistry* **24**, 1–95 (1970). 12, 14
- [41] Dill, K. A. & Shortle, D. Denatured states of proteins. *Annual Review of Biochemistry* **60**, 795–825 (1991). 3
- [42] Flory, P. J. *Statistical mechanics of chain molecules* (Wiley, 1969). 3, 4, 12, 104
- [43] Ramachandran, G. & Sasisekharan, V. Conformation of polypeptides and proteins. vol. 23 of *Advances in Protein Chemistry*, 283 – 437 (Academic Press, 1968). 4
- [44] Pappu, R. V., Srinivasan, R. & Rose, G. D. The Flory isolated-pair hypothesis is not valid for polypeptide chains: Implications for protein folding. *Proceedings of the National Academy of Sciences* **97**, 12565–12570 (2000). 4

- 
- [45] 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* **108**, 109–113 (2011). 4
- [46] Doi, M. & Edwards, S. F. *The Theory of Polymer Dynamics*, vol. 1 (Oxford University Press, 1986). 5
- [47] Edwards, S. F. The statistical mechanics of polymers with excluded volume. *Proceedings of the Physical Society* **85**, 613 (1965). 5
- [48] 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* **72**, 1807–1811 (1975). 5
- [49] Prince E. Rouse, J. A theory of the linear viscoelastic properties of dilute solutions of coiling polymers. *The Journal of Chemical Physics* **21**, 1272–1280 (1953). 5
- [50] Zimm, B. H. Dynamics of polymer molecules in dilute solution: Viscoelasticity, flow birefringence and dielectric loss. *The Journal of Chemical Physics* **24**, 269–278 (1956). 5
- [51] Szabo, A., Schulten, K. & Schulten, Z. First passage time approach to diffusion controlled reactions. *The Journal of Chemical Physics* **72**, 4350 (1980). 5
- [52] Jacobson, H. & Stockmayer, W. H. Intramolecular reaction in polycondensations. i. the theory of linear systems. *The Journal of Chemical Physics* **18**, 1600–1606 (1950). 5
- [53] Levinthal, C. How to fold graciously. In *Mossbauer Spectroscopy in Biological Systems: Proceedings of a meeting held at Allerton House, Monticello, Illinois, 22–24* (University of Illinois Press, 1969). 6
- [54] Jackson, S. E. & Fersht, A. R. Folding of chymotrypsin inhibitor 2. 1. Evidence for a two-state transition. *Biochemistry* **30**, 10428–10435 (1991). 6
- [55] Jackson, S. E. How do small single-domain proteins fold? *Folding and Design* **3**, R81–R91 (1998). 11
- [56] Kubelka, J., Hofrichter, J. & Eaton, W. A. The protein folding speed limit. *Current Opinion in Structural Biology* **14**, 76–88 (2004). 6

- [57] Levinthal, C. Are there pathways for protein folding? *Journal of Medical Physics* 44–45 (1968). 6
- [58] Zwanzig, R., Szabo, A. & Bagchi, B. Levinthals paradox. *Proceedings of the National Academy of Sciences of the United States of America* **89**, 20 (1992). 6
- [59] Karplus, M. & Weaver, D. L. Protein-folding dynamics. *Nature* **3**, 404–406 (1976). 6
- [60] Burton, R. E., Myers, J. K. & Oas, T. G. Protein folding dynamics: Quantitative comparison between theory and experiment. *Biochemistry* **37**, 5337–5343 (1998). 6
- [61] Fersht, A. R. Optimization of rates of protein folding: the nucleation-condensation mechanism and its implications. *Proceedings of the National Academy of Sciences* **92**, 10869–10873 (1995). 6
- [62] Itzhaki, L. S., Otzen, D. E. & Fersht, A. R. The structure of the transition state for folding of chymotrypsin inhibitor 2 analysed by protein engineering methods: Evidence for a nucleation-condensation mechanism for protein folding. *Journal of Molecular Biology* **254**, 260 – 288 (1995). 6
- [63] Ptitsyn, O. B. Molten globule and protein folding. vol. 47 of *Advances in Protein Chemistry*, 83 – 229 (Academic Press, 1995). 6
- [64] Bachmann, A. & Kiefhaber, T. Apparent two-state tandem folding is a sequential process along a defined route. *Journal of Molecular Biology* **306**, 375 – 386 (2001). 7, 9
- [65] Sanchez, I. E. & Kiefhaber, T. Evidence for sequential barriers and obligatory intermediates in apparent two-state protein folding. *Journal of Molecular Biology* **325**, 367 – 376 (2003). 7
- [66] Wagner, C. & Kiefhaber, T. Intermediates can accelerate protein folding. *Proceedings of the National Academy of Sciences* **96**, 6716–6721 (1999). 7
- [67] Ikai, A. & Tanford, C. Kinetic evidence for incorrectly folded intermediate states in the refolding of denatured proteins. *Nature* **230**, 100–102 (1971). 7

- 
- [68] Kiefhaber, T., Labhardt, A. M. & Baldwin, R. L. Direct NMR evidence for an intermediate preceding the rate-limiting step in the unfolding of ribonuclease a. *Nature* **375**, 513–515 (1995). 7
- [69] Kim, P. S. & Baldwin, R. L. Intermediates in the folding reactions of small proteins. *Annual Review of Biochemistry* **59**, 631–660 (1990). 7
- [70] Kim, P. S. & Baldwin, R. L. Specific intermediates in the folding reactions of small proteins and the mechanism of protein folding. *Annual Review of Biochemistry* **51**, 459–489 (1982). 7
- [71] Eyring, H. The activated complex in chemical reactions. *The Journal of Chemical Physics* **3**, 107–115 (1935). 7, 8
- [72] LEFFLER, J. E. Parameters for the description of transition states. *Science* **117**, 340–341 (1953). 8
- [73] Sanchez, I. E. & Kiefhaber, T. Non-linear rate-equilibrium free energy relationships and hammond behavior in protein folding. *Biophysical Chemistry* **100**, 397 – 407 (2003). 9
- [74] Sanchez, 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. *Journal of Molecular Biology* **327**, 867 – 884 (2003).
- [75] 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. *Proceedings of the National Academy of Sciences* **97**, 17–22 (2000). 9
- [76] Kramers, H. Brownian motion in a field of force and the diffusion model of chemical reactions. *Physica* **7**, 284 – 304 (1940). 9, 101
- [77] Qiu, L. & Hagen, S. J. A limiting speed for protein folding at low solvent viscosity. *Journal of the American Chemical Society* **126**, 3398–3399 (2004). PMID: 15025447. 9, 10, 101, 103
- [78] de Gennes, P. G. *Scaling concepts in polymer physics*, vol. 22 of *G - Reference, Information and Interdisciplinary Subjects Series* (Cornell University Press, 1979). 9, 101

- [79] Qiu, L. & Hagen, S. J. Internal friction in the ultrafast folding of the tryptophan cage. *Chemical Physics* **307**, 243–249 (2004). 10, 101, 103
- [80] Pabit, S. A., Roder, H. & Hagen, S. J. Internal friction controls the speed of protein folding from a compact configuration. *Biochemistry* **43**, 12532–12538 (2004). PMID: 15449942. 10, 103
- [81] Jacob, M., Schindler, T., Balbach, J. & Schmid, F. X. Diffusion control in an elementary protein folding reaction. *Proceedings of the National Academy of Sciences* **94**, 5622–5627 (1997). 10, 102
- [82] Plaxco, K. W. & Baker, D. Limited internal friction in the rate-limiting step of a two-state protein folding reaction. *Proceedings of the National Academy of Sciences of the United States of America* **95**, 13591–13596 (1998).
- [83] Jacob, M., Geeves, M., Holtermann, G. & X., S. Diffusional barrier crossing in a two-state protein folding reaction. *Nature Structural and Molecular Biology* **6**, 923–926 (1999). 101, 102
- [84] Jacob, M. & Schmid, F. X. Protein folding as a diffusional process. *Biochemistry* **38**, 13773–13779 (1999). 10
- [85] Hagen, S. J., Hofrichter, J., Szabo, A. & Eaton, W. A. Diffusion-limited contact formation in unfolded cytochrome c: Estimating the maximum rate of protein folding. *Proceedings of the National Academy of Sciences of the United States of America* **93**, 11615–11617 (1996). 10, 15
- [86] Hagen, S. J., Qiu, L. & Pabit, S. A. Diffusional limits to the speed of protein folding: fact or friction? *Journal of Physics: Condensed Matter* **17**, S1503–S1514 (2005). 10, 101, 103
- [87] Privalov, P. L. Stability of proteins: small globular proteins. *Advances in Protein Chemistry* **33**, 167–241 (1979). 11
- [88] Young, J. C., Agashe, V. R., Siegers, K. & Hartl, F. U. Pathways of chaperone-mediated protein folding in the cytosol. *Nature Reviews Molecular Cell Biology* **5**, 781–791 (2004). 11



- [89] Donato, H. & Martin, R. B. Conformations of carp muscle calcium binding parvalbumin. *Biochemistry* **13**, 4575–4579 (1974). 11, 25
- [90] Hofmeister, F. Zur Lehre von der Wirkung der Salze. *Naunyn-Schmiedeberg's Archives of Pharmacology* **24**, 247–260 (1888). 10.1007/BF01918191. 11
- [91] Baldwin, R. L. How Hofmeister ion interactions affect protein stability. *Biophysical Journal* **71**, 2056 – 2063 (1996). 12
- [92] Hochachka, P. W. & Somero, G. N. *Biochemical Adaptation: Mechanism and Process in Physiological Evolution*, vol. 30 (Oxford University Press, 2002). 12
- [93] Bolen, D. W. & Rose, G. D. Structure and energetics of the hydrogen-bonded backbone in protein folding. *Annual Review of Biochemistry* **77**, 339–62 (2008). 12, 59
- [94] Holthauzen, L. M. F., Rösgen, J. & Bolen, D. W. Hydrogen bonding progressively strengthens upon transfer of the protein urea-denatured state to water and protecting osmolytes. *Biochemistry* **49**, 1310–8 (2010). 12, 14
- [95] Schellman, J. A. Selective binding and solvent denaturation. *Biopolymers* **26**, 549–559 (1987). 12, 14
- [96] Timasheff, S. N. Control of protein stability and reactions by weakly interacting cosolvents: The simplicity of the complicated. In Frederic M. Richards, D. S. E. & Kim, P. S. (eds.) *Linkage Thermodynamics of Macromolecular Interactions*, vol. 51 of *Advances in Protein Chemistry*, 355 – 432 (Academic Press, 1998). 12
- [97] Tanford, C. Isothermal unfolding of globular proteins in aqueous urea solutions. *Journal of the American Chemical Society* **86**, 2050–2059 (1964). 12
- [98] Auton, M. & Bolen, D. W. Application of the transfer model to understand how naturally occurring osmolytes affect protein stability. *Methods Enzymol* **428**, 397–418 (2007). 13
- [99] 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 Science* **4**, 2138–2148 (1995). 14

- [100] Pace, C. N., Grimsley, G. R. & Scholtz, J. M. *Denaturation of Proteins by Urea and Guanidine Hydrochloride*, 45–69 (Wiley-VCH Verlag GmbH, 2005). 13
- [101] Auton, M. & Bolen, D. W. Predicting the energetics of osmolyte-induced protein folding/unfolding. *Proceedings of the National Academy of Sciences of the United States of America* **102**, 15065–8 (2005). 13, 14, 42, 64, 78, 94
- [102] Creamer, T. P., Srinivasan, R. & Rose, G. D. Modeling unfolded states of proteins and peptides. ii. backbone solvent accessibility. *Biochemistry* **36**, 2832–2835 (1997). 13, 42, 133
- [103] Auton, M., Rösger, J., Sinev, M., Holthauzen, L. M. F. & Bolen, D. W. Osmolyte effects on protein stability and solubility: A balancing act between backbone and side-chains. *Biophysical Chemistry* **159**, 90–99 (2011). 14, 64
- [104] Pace, C. [14] Determination and analysis of urea and guanidine hydrochloride denaturation curves. In C. H. W. Hirs, S. N. T. (ed.) *Enzyme Structure Part L*, vol. 131 of *Methods in Enzymology*, 266 – 280 (Academic Press, 1986). 14, 41
- [105] Schellman, J. A. Fifty years of solvent denaturation. *Biophysical Chemistry* **96**, 91 – 101 (2002).
- [106] Timasheff, S. N. Thermodynamic binding and site occupancy in the light of the schellman exchange concept. *Biophysical Chemistry* **101-102**, 99 – 111 (2002). Special issue in honour of John A Schellman. 14
- [107] Tanford, C. *Physical Chemistry of Macromolecules* (Wiley, 1961). 14
- [108] Baskakov, I. & Bolen, D. W. Forcing thermodynamically unfolded proteins to fold. *Journal of Biological Chemistry* **273**, 4831–4834 (1998). 14
- [109] Schellman, J. A. Macromolecular binding. *Biopolymers* **14**, 999–1018 (1975). 14
- [110] Schellman, J. A. Solvent denaturation. *Biopolymers* **17**, 1305–1322 (1978).
- [111] Schellman, J. A. The thermodynamic stability of proteins. *Annual Review of Biophysics and Biophysical Chemistry* **16**, 115–137 (1987). 14

- [112] Haas, E., Katchalski-Katzir, E. & Steinberg, I. Z. Brownian motion of the ends of oligopeptide chains in solution as estimated by energy transfer between the chain ends. *Biopolymers* **17**, 11–31 (1978). 15
- [113] Lapidus, L. J., Eaton, W. A. & Hofrichter, J. Measuring the rate of intramolecular contact formation in polypeptides. *Proceedings of the National Academy of Sciences of the United States of America* **97**, 7220–7225 (2000). 15
- [114] Bent, D. V. & Hayon, E. Excited state chemistry of aromatic amino acids and related peptides. iii. Tryptophan. *Journal of the American Chemical Society* **97**, 2612–2619 (1975). PMID: 237041. 15
- [115] Yeh, I.-C. & Hummer, G. Peptide loop-closure kinetics from microsecond molecular dynamics simulations in explicit solvent. *Journal of the American Chemical Society* **124**, 6563–6568 (2002). 15
- [116] Chang, I.-J., Lee, J. C., Winkler, J. R. & Gray, H. B. The protein-folding speed limit: intrachain diffusion times set by electron-transfer rates in denatured Ru(NH<sub>3</sub>)<sub>5</sub>(His-33)-Zn-cytochrome c. *Proceedings of the National Academy of Sciences of the United States of America* **100**, 3838–3840 (2003). 16
- [117] Hudgins, R. R., Huang, F., Gramlich, G. & Nau, W. M. A fluorescence-based method for direct measurement of submicrosecond intramolecular contact formation in biopolymers: An exploratory study with polypeptides. *Journal of the American Chemical Society* **124**, 556–564 (2002). PMID: 11804484. 16
- [118] Huang, F. & Nau, W. M. A conformational flexibility scale for amino acids in peptides. *Angewandte Chemie International Edition* **42**, 2269–2272 (2003). 16
- [119] Neuweiler, H., Schulz, A., Böhmer, M., Enderlein, J. & Sauer, M. Measurement of submicrosecond intramolecular contact formation in peptides at the single-molecule level. *Journal of the American Chemical Society* **125**, 5324–5330 (2003). 16
- [120] Fierz, B. & Kiefhaber, T. *Dynamics of Unfolded Polypeptide Chains*, 809–855 (Wiley-VCH Verlag GmbH, 2005). 16, 20
- [121] Dexter, D. L. A theory of sensitized luminescence in solids. *The Journal of Chemical Physics* **21**, 836 (1953). 16, 17

- [122] Klán, P. & Wirz, J. *Photochemistry of Organic Compounds* (Wiley-VCH, 2009). 17
- [123] Satzger, H. *et al.* Ultrafast quenching of the xanthone triplet by energy transfer: New insight into the intersystem crossing kinetics. *Journal of Physical Chemistry A* **108**, 10072–10079 (2004). 17
- [124] Heinz, B. *et al.* On the unusual fluorescence properties of xanthone in water. *Physical Chemistry Chemical Physics* **8**, 3432–3439 (2006). 17
- [125] Reiner, A., Henklein, P. & Kiefhaber, T. An unlocking/relocking barrier in conformational fluctuations of villin headpiece subdomain. *Proceedings of the National Academy of Sciences of the United States of America* **107**, 4955–60 (2010). 17, 22, 23
- [126] Wagner, P. J. & Klán, P. Intramolecular triplet energy transfer in flexible molecules: Electronic, dynamic, and structural aspects. *Journal of the American Chemical Society* **121**, 9626–9635 (1999). 17
- [127] Krieger, F. *Dynamics in Unfolded Polypeptide Chains as Model for Elementary Steps in Protein Folding*. Ph.D. thesis, Universität Basel (2004). 20, 25, 50, 51, 53, 55, 56, 67, 80, 89, 97, 99
- [128] Brant, D. A. & Flory, P. J. The configuration of random polypeptide chains. ii. Theory. *Journal of the American Chemical Society* **87**, 2791–2800 (1965). 21
- [129] Schimmel, P. R. & Flory, P. J. Conformational energy and configurational statistics of poly-l-proline. *Proceedings of the National Academy of Sciences* **58**, 52–59 (1967). 21
- [130] Joder, K. N. *Intramolecular and Intermolecular Diffusion Processes in Protein Folding and Assembly*. Ph.D. thesis, Technische Universität München, München (2011). 21, 22, 25, 50, 99, 103, 104, 105, 107, 111
- [131] Fierz, B. & Kiefhaber, T. End-to-end vs interior loop formation kinetics in unfolded polypeptide chains. *Journal of the American Chemical Society* **129**, 672–9 (2007). 21
- [132] 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* **23**, 797–803 (1990). 21

- 
- [133] Krieger, F., Möglich, A. & Kiefhaber, T. Effect of proline and glycine residues on dynamics and barriers of loop formation in polypeptide chains. *Journal of the American Chemical Society* **127**, 3346–52 (2005). 22
- [134] Dix, J. A. & Verkman, A. Crowding effects on diffusion in solutions and cells. *Annual Review of Biophysics* **37**, 247–263 (2008). PMID: 18573081. 22
- [135] Fierz, B., Reiner, A. & Kiefhaber, T. Local conformational dynamics in alpha-helices measured by fast triplet transfer. *Proceedings of the National Academy of Sciences of the United States of America* **106**, 1057–62 (2009). 22, 23
- [136] Reiner, A. Triplet-triplet energy transfer studies on conformational dynamics in peptides and a protein. *Journal of Peptide Science* **17**, 413–419 (2011). 22
- [137] Deuticke, H. J. Über die Sedimentationskonstante von Muskelproteinen. *Hoppe-Seyler's Zeitschrift für Physiologische Chemie* **224**, 216–228 (1934). 24
- [138] Jiang, Y., Johnson, J. D. & Rall, J. A. Parvalbumin relaxes frog skeletal muscle when sarcoplasmic reticulum ca(2+)-atpase is inhibited. *American Journal of Physiology - Cell Physiology* **270**, C411–C417 (1996). 24
- [139] Schwaller, B. *et al.* Prolonged contraction-relaxation cycle of fast-twitch muscles in parvalbumin knockout mice. *American Journal of Physiology - Cell Physiology* **276**, C395–C403 (1999). 24
- [140] Arif, S. H. A Ca(2+)-binding protein with numerous roles and uses: parvalbumin in molecular biology and physiology. *BioEssays news and reviews in molecular cellular and developmental biology* **31**, 410–421 (2009). 24
- [141] Stuhfauth, I., Reininghaus, J., Jockusch, H. & Heizmann, C. Calcium-binding protein, parvalbumin, is reduced in mutant mammalian muscle with abnormal contractile properties. *Proceedings of the National Academy of Sciences of the United States of America* **81**, 4814 (1984). 24
- [142] Moncrieffe, M. C. *et al.* Structure-fluorescence correlations in a single tryptophan mutant of carp parvalbumin: solution structure, backbone and side-chain dynamics. *Journal of Molecular Biology* **297**, 147–163 (2000). 24, 38

- [143] 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. *The Journal of Biological Chemistry* **248**, 3305–3312 (1973). 24
- [144] Kretsinger, R. H. & Nockolds, C. E. Carp muscle calcium-binding protein. ii. Structure determination and general description. *The Journal of Biological Chemistry* **248**, 3313–3326 (1973). 24, 25
- [145] Kumar, V. D., Lee, L. & Edwards, B. F. P. Refined crystal structure of calcium-liganded carp parvalbumin 4.25 at 1.5 Å resolution. *Biochemistry* **29**, 1404–1412 (1990). PMID: 2334704. 24, 52, 113
- [146] Bernstein, F. C. *et al.* The protein data bank: A computer-based archival file for macromolecular structures. *Journal of Molecular Biology* **112**, 535–542 (1977). 24, 52, 113
- [147] Kretsinger, R. H. Structure and evolution of calcium-modulated proteins. *CRC critical reviews in biochemistry* **8** (1980). 25, 124
- [148] Lin, L.-N. & Brandts, J. F. Further evidence suggesting that the slow phase in protein unfolding and refolding is due to proline isomerization: a kinetic study of carp parvalbumins. *Biochemistry* **17**, 4102–4110 (1978). 25
- [149] Kuwajima, K., Sakuraoka, A., Fueki, S., Yoneyama, M. & Sugai, S. Folding of carp parvalbumin studied by equilibrium and kinetic circular dichroism spectra. *Biochemistry* **27**, 7419–7428 (1988). 25, 115, 125
- [150] Langheld, S. Ph.D. thesis, Biozentrum der Universität Basel (2007). 25, 34, 116
- [151] Nyffenegger, C. *Dynamics in Different States of Parvalbumin Measured by Triplet-Triplet Energy Transfer*. Master's thesis, Universität Basel (2007). 25
- [152] Dawson, P. E., Muir, T. W., Clark-Lewis, I. & Kent, S. B. H. Synthesis of proteins by native chemical ligation. *Science* **266**, 776–779 (1994). 26, 27
- [153] David, R., Richter, M. P. & Beck-Sickinger, A. G. Expressed protein ligation. *European Journal of Biochemistry* **271**, 663–677 (2004). 27

- [154] Blanco-Canosa, J. B. & Dawson, P. E. An efficient Fmoc-SPPS approach for the generation of thioester peptide precursors for use in native chemical ligation. *Angewandte Chemie International Edition* **47**, 6851–6855 (2008). 27
- [155] Noren, C. J., Wang, J. & Perler, F. B. Protein-Spleißen: Mechanismus und Anwendungen. *Angewandte Chemie* **112**, 458–476 (2000). 27
- [156] Paulus, H. Protein splicing and related forms of protein autoprocessing. *Annual Review of Biochemistry* **69**, 447–496 (2000). 27
- [157] Muir, T. W. Semisynthesis of proteins by expressed protein ligation. *Annual Review of Biochemistry* **72**, 249–289 (2003). 27
- [158] Southworth, M. W., Benner, J. & Perler, F. B. An alternative protein splicing mechanism for inteins lacking an N-terminal nucleophile. *The EMBO Journal* **19**, 5019–5026 (2000). 27
- [159] Evans, T. C., Benner, J. & Xu, M.-Q. Semisynthesis of cytotoxic proteins using a modified protein splicing element. *Protein Science* **7**, 2256–2264 (1998). 27
- [160] Southworth, M. W., Amaya, K., Evans, T. C., Xu, M.-Q. & Perler, F. B. Purification of proteins fused to either the amino or carboxy terminus of the *Mycobacterium xenopi* gyrase a intein. *BioTechniques* **27**, 110–120 (1999). 27
- [161] Graham, R. & Lewis, J. R. Synthesis of 9-oxoxanthen-2-carboxylic acids. *Journal of the Chemical Society, Perkin Transactions 1* 876–881 (1978). 32
- [162] Reiner, A. *Conformational Dynamics and Stability of Structured Peptides and Small Proteins*. Ph.D. thesis, Universität Basel (2007). 33, 112
- [163] Saiki, R. *et al.* Primer-directed enzymatic amplification of DNA with a thermostable DNA polymerase. *Science* **239**, 487–491 (1988). 34
- [164] Cohen, S. N., Chang, A. C. Y. & Hsu, L. Nonchromosomal antibiotic resistance in bacteria: Genetic transformation of *Escherichia coli* by R-factor DNA. *Proceedings of the National Academy of Sciences* **69**, 2110–2114 (1972). 35

- [165] Bradford, M. M. A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein-dye binding. *Analytical Biochemistry* **72**, 248 – 254 (1976). 38
- [166] Laemmli, U. K. Cleavage of structural proteins during the assembly of the head of bacteriophage t4. *Nature* **227**, 680–685 (1970). 39
- [167] Wan, Q. & Danishefsky, S. J. Free-radical-based, specific desulfurization of cysteine: A powerful advance in the synthesis of polypeptides and glycopolypeptides. *Angewandte Chemie International Edition* **46**, 9248–9252 (2007). 40, 121
- [168] Schmid, F. X. *Protein Structure, A Practical Approach* (Oxford University Press, Oxford, United Kingdom, 1997). 41
- [169] Perl, D. *et al.* Thermodynamics of a diffusional protein folding reaction. *Biophysical Chemistry* **96**, 173 – 190 (2002). 41, 60
- [170] Lee, B. & Richards, F. The interpretation of protein structures: Estimation of static accessibility. *Journal of Molecular Biology* **55**, 379 – 400 (1971). 42, 133
- [171] 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* **36**, 8413–8421 (1997). 43
- [172] Reiner, A., Wildemann, D., Fischer, G. & Kiefhaber, T. Effect of thiopeptide bonds on alpha-helix structure and stability. *Journal of the American Chemical Society* **130**, 8079–84 (2008). 43
- [173] DeLano, W. The pymol molecular graphics system, version 0.99. 52, 113
- [174] Gronenborn, A. *et al.* A novel, highly stable fold of the immunoglobulin binding domain of streptococcal protein g. *Science* **253**, 657–661 (1991). 52
- [175] Pace, C. N. & Scholtz, J. M. A helix propensity scale based on experimental studies of peptides and proteins. *Biophysical Journal* **75**, 422 – 427 (1998). 51
- [176] Derancourt, J., Haiech, J. & Pechère, J. F. Binding of calcium by parvalbumin fragments. *Biochimica et Biophysica Acta* **532**, 373–375 (1978). 51



- [177] Coffee, C. J. & Solano, C. Preparation and properties of carp muscle parvalbumin fragments a (residues 1-75) and b (residues 76-108). *Biochimica et Biophysica Acta (BBA) - Protein Structure* **453**, 67 – 80 (1976). 51
- [178] Yancey, P., Clark, M., Hand, S., Bowlus, R. & Somero, G. Living with water stress: evolution of osmolyte systems. *Science* **217**, 1214–1222 (1982). 58
- [179] Reddy K., R. C., Lilie, H., Rudolph, R. & Lange, C. L-arginine increases the solubility of unfolded species of hen egg white lysozyme. *Protein Science* **14**, 929–935 (2005). 58
- [180] Taneja, S. & Ahmad, F. Increased thermal stability of proteins in the presence of amino acids. *Biochemical Journal* **303**, 147–153 (1994). 58
- [181] Lin, T.-Y. & Timasheff, S. N. On the role of surface tension in the stabilization of globular proteins. *Protein Science* **5**, 372–381 (1996).
- [182] Arakawa, T. & Tsumoto, K. The effects of arginine on refolding of aggregated proteins: not facilitate refolding, but suppress aggregation. *Biochemical and Biophysical Research Communications* **304**, 148 – 152 (2003). 58
- [183] Street, T. O., Bolen, D. W. & Rose, G. D. A molecular mechanism for osmolyte-induced protein stability. *Proceedings of the National Academy of Sciences of the United States of America* **103**, 13997–4002 (2006). 59
- [184] Tischer, A. *Effekte von Arginin und von ionischen Flüssigkeiten auf Proteine*. Ph.D. thesis, Martin-Luther-Universität Halle-Wittenberg (2011). 64
- [185] Liu, Y. & Bolen, D. W. The peptide backbone plays a dominant role in protein stabilization by naturally occurring osmolytes. *Biochemistry* **34**, 12884–12891 (1995). PMID: 7548045. 76
- [186] Müller-Späh, S. *et al.* Charge interactions can dominate the dimensions of intrinsically disordered proteins. *Proceedings of the National Academy of Sciences of the United States of America* **107**, 14609–14614 (2010). 92
- [187] Mashino, T. & Fridovich, I. Effects of urea and trimethylamine-N-oxide on enzyme activity and stability. *Archives of Biochemistry and Biophysics* **258**, 356 – 360 (1987). 100

- [188] Ansari, A., Jones, C. M., Henry, E. R., Hofrichter, J. & Eaton, W. A. The role of solvent viscosity in the dynamics of protein conformational changes. *Science* **256**, 1796–1798 (1992). 101, 103
- [189] Jas, G. S., Eaton, W. A. & Hofrichter, J. Effect of viscosity on the kinetics of alpha-helix and beta-hairpin formation. *The Journal of Physical Chemistry B* **105**, 261–272 (2001). 102
- [190] Cellmer, T., Henry, E. R., Hofrichter, J. & Eaton, W. A. Measuring internal friction of an ultrafast-folding protein. *Proceedings of the National Academy of Sciences* **105**, 18320–18325 (2008).
- [191] Wensley, B. G. *et al.* Experimental evidence for a frustrated energy landscape in a three-helix-bundle protein family. *Nature* **463**, 685–689 (2010). 101, 103
- [192] Schulz, J. C. F., Schmidt, L., Best, R. B., Dzubiella, J. & Netz, R. R. Peptide chain dynamics in light and heavy water: Zooming in on internal friction. *Journal of the American Chemical Society* **134**, 6273–6279 (2012). 102
- [193] Sasisekharan, V. *Collagen*, 39 (Interscience, New York, 1962). 104
- [194] Soranno, A. *et al.* Quantifying internal friction in unfolded and intrinsically disordered proteins with single-molecule spectroscopy. *Proceedings of the National Academy of Sciences* (2012). 106
- [195] Buck, M. Trifluoroethanol and colleagues: cosolvents come of age. Recent studies with peptides and proteins. *Quarterly Reviews of Biophysics* **31**, 297–355 (1998). 114
- [196] Main, E. R. G. & Jackson, S. E. Does trifluoroethanol affect folding pathways and can it be used as a probe of structure in transition states? *Nature Structural and Molecular Biology* **6**, 831–835 (1999). 114
- [197] Santoro, M. M. & Bolen, D. W. A test of the linear extrapolation of unfolding free energy changes over an extended denaturant concentration range. *Biochemistry* **31**, 4901–7 (1992). 123

# Acknowledgements

This work was carried out from January 2008 until August 2012 in the laboratory of Prof. Dr. Thomas Kiefhaber in the chair for Biophysical Chemistry at the Chemistry Department of the Technische Universität München.

At first, I want to express my gratitude to Prof. Dr. Thomas Kiefhaber for having the fantastic possibility to work in his group and for his excellent scientific supervision. I am very grateful for his support and independency I experienced during my thesis. Further, I would like to thank Prof. Dr. George Rose for fruitful discussions and his great insights he provided.

For partial funding, I would like to thank the International Graduate School of Materials Science of Complex Interfaces.

Many thanks to all current and former members of the group: Tobias Aumüller, Annett Bachmann, Matthias Berg, Christophe Bodenreider, María Dolores Crespo, Beat Fierz, Sarah Güthe, Judith Habazettl, Michael Hösl, Kerstin Hoffmann-Jacobsen, Kristine Steen Jensen, Peter Kämmerer, Richard Kil, Sabine Kullick, Stefan Langheld, Natalie Merk, Rita Müller, Alexander Ogrodnik, Sabine Rauch, Andreas Reiner, Manuela Schätzle, Ursula Seidel, Therese Schulthess, Tobias Schümmer, Karin Stecher, Matthias Stecher, Traudl Wenger, Joseph Wey, Daniel Winter and Ursula Zinth. Thanks for the good time and the nice atmosphere in the lab. In special, I would like to thank, Stephan for his introduction into the field of native chemical ligation, Andreas for his support and the BNAA, Prof. Dr. Matthew Auton and Dr. Alexander Tischer for the Arginine GTFE data and Joseph for the synthesis of xanthonic acid.

For proofreading parts of this thesis, I want to acknowledge Michael, Natalie, Kristine and Ursula.

Finally, I wish to express my gratitude to my family, my parents, my sister and Maria for their support not only during this thesis but through my whole life.



# Lebenslauf

## Persönliche Daten

Christian Nyffenegger

geboren am 21.07.1983 in Liestal

ledig

## Ausbildung und Tätigkeiten

8/1990 bis 6/1995	Primarschule Schulhaus Dorf, Füllinsdorf
8/1995 bis 6/1999	Sekundarschule Mühlacker, Frenkendorf
8/1999 bis 12/2002	Gymnasium Liestal, Liestal
12/2002	Maturität, Schwerpunkte Mathematik und Physik
2/2003 bis 5/2003	Wehrdienst
9/2003 bis 7/2005	Grundstudium der Biologie am Biozentrum der Universität Basel
8/2005 bis 7/2006	Bachelorstudium der Molekularbiologie am Biozentrum der Universität Basel
9/2006 bis 12/2007	Masterarbeit am Lehrstuhl für Biophysikalische Chemie von Prof. T. Kiefhaber: "Dynamics in Different States of Carp $\beta$ -Parvalbumin, Measured with Triplet-Triplet Energy Transfer".
1/2008 bis heute	Dissertationsstudium / Wissenschaftlicher Mitarbeiter am Lehrstuhl für Biophysikalische Chemie von Prof. T. Kiefhaber, Technische Universität München

