



Fakultät für Chemie

Lehrstuhl für Biophysikalische Chemie

End-to-End Distance Distribution and Intra-Chain Diffusion
in Unfolded Polypeptide Chains
Determined by Time-Resolved FRET Measurements

Ursula G. Zinth

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. Michael Groll

Prüfer der Dissertation:

1. Univ.-Prof. Dr. Thomas Kiefhaber
2. Univ.-Prof. Dr. Matthias Rief
3. Univ.-Prof. Dr. Claus Seidel
(Heinrich-Heine-Universität Düsseldorf)

Die Dissertation wurde am 11.11.2013 bei der Technischen Universität München
eingereicht und durch die Fakultät für Chemie
am 20.1.2014 angenommen.

Contents

1	Introduction	1
1.1	Proteins	1
1.1.1	Proteins in nature and their functions	1
1.1.2	Properties of polypeptide chains and the native structure	2
1.2	Protein folding and the native state	3
1.2.1	Protein stability	3
1.2.2	Protein folding	4
1.2.3	The energy landscape of protein folding	5
1.3	The unfolded state	6
1.3.1	The random coil model	7
1.3.2	Characteristics of the ensemble of unfolded conformations	9
1.3.3	The wormlike chain and the chain with excluded volume	11
1.4	Dynamics in the unfolded state	13
1.4.1	Methods to study dynamics in the unfolded state	13
1.4.2	Dynamics in unfolded peptides and proteins	16
1.4.3	Theory of diffusion with focus on intra-chain diffusion	17
1.5	Principles of Fluorescence Methods	20
1.5.1	Fluorescence	20
1.5.2	Fluorescence Resonance Energy Transfer (FRET)	21
2	Aims of research	27

3	Results	31
3.1	End-to-end distance distribution and intra-chain diffusion coefficient determined by FRET	31
3.1.1	FRET in unfolded peptide systems	31
3.1.2	Analyzing FRET experiments under consideration of intra-chain diffusion	33
3.1.3	Experimental design	34
3.1.4	Determination of the intra-chain diffusion coefficient by global analysis of different FRET experiments.	35
3.1.5	Generalization of global analysis method	39
3.1.6	Summary and discussions	42
3.2	Unfolded model peptides	45
3.2.1	Glycine-Serine repeat peptides	45
3.2.2	EF-loop from Parvalbumin.	46
3.2.3	End-to-end distance distribution and intra-chain diffusion coefficient	47
3.2.4	The chain with excluded volume as a model for unfolded peptide conformations	48
3.2.5	Monte Carlo modeling of unfolded peptide conformations	49
3.2.6	End-to-end distance and radius of gyration of unfolded polypeptides	51
3.2.7	Summary and discussions	52
3.3	Effect of Solvent friction and internal friction on intra-chain diffusion of unfolded polypeptides	55
3.3.1	Friction in the dynamics of polymers and unfolded polypeptides	55
3.3.2	Effect of glycerol on the intra-chain diffusion coefficient and the end-to-end distance distribution	57
3.3.3	The role of friction for the dynamics of unfolded polypeptides	61
3.3.4	Summary and discussions	63
3.4	Influence of denaturant and stabilizing osmolyte on the properties of unfolded polypeptides	67
3.4.1	Effect of denaturants and stabilizing osmolytes on protein stability	67

3.4.2	End-to-end distance distribution and intra-chain diffusion in the presence of GdmCl	69
3.4.3	End-to-end distance distribution and intra-chain diffusion in the presence of sarcosine	73
3.4.4	Monte Carlo simulations	76
3.4.5	Summary and discussions	82
3.5	Temperature dependence of intra-chain diffusion and roughness of energy landscape	85
3.5.1	Diffusion in a rough potential	85
3.5.2	End-to-end distance distribution and intra-chain diffusion coefficient at various temperatures	86
3.5.3	Summary and discussions	90
4	Summary	95
A	Materials and methods	99
A.1	Peptide synthesis and purification	99
A.1.1	Peptide synthesis	99
A.1.2	Labeling with chromophores	100
A.1.3	Purification	100
A.2	Fluorescence measurements	100
A.2.1	Sample preparation.	101
A.2.2	Determination of fluorescence quantum yields and Förster-distances	102
A.2.3	Fluorescence lifetime measurements	103
A.3	Global data analysis	104
A.3.1	Data input.	104
A.3.2	Solution of PDE	105
A.3.3	Re-convolution fit	105
A.3.4	Global fit	106
A.3.5	Support plane analysis.	106

Contents

List of Abbreviations	109
List of Figures	111
List of Tables	113
Bibliography	115

1. Introduction

1.1. Proteins

1.1.1. Proteins in nature and their functions

Proteins form the most versatile class of macromolecules in nature. They are made of 20 different α -L-amino acids as building blocks¹⁻³, which are pairwise covalently linked through amide bonds to form chains typically consisting of several hundred residues.⁴ In order to fulfill the required function, a polypeptide chain has to adopt a well defined 3-dimensional structure, the native state, which is stabilized by several non-covalent interactions. Physiological functions accomplished by proteins comprise the catalysis of reactions as enzymes, the transport and storage of metabolites or small molecules and the transmission of nerve impulses or other signals. Furthermore, proteins are building blocks for the cytoskeleton, act as molecular motors and provide immune protection.

The native, 3-dimensional structure of a protein is predetermined by its amino acid sequence,⁵ which in terms is encoded in the corresponding gene. A ribonucleotide transcript of this gene is used in protein synthesis on the ribosome. Upon leaving the ribosome through an exit channel, most of the proteins fold into a stable, well defined, native structure. However, around 40 % of all proteins lack a well defined structure or have unstructured regions and are considered to be intrinsically disordered.⁶⁻⁸ In some cases, folding can be assisted by folding helpers like chaperones⁹ or isomerases^{10,11} or influenced by interactions with the ribosome^{12,13}.

In competition to the folding process, mis-folding into aggregates or assembly into oligomers and fibrils can occur.¹⁴ Mis-folded proteins are involved in several severe diseases like Alzheimer's disease, Parkinson's disease, Bovine spongiform encephalopathy, AA-amyloidosis, etc.

A detailed characterization of proteins throughout the whole folding process is crucial for understanding the elementary principles of their biological functionality.

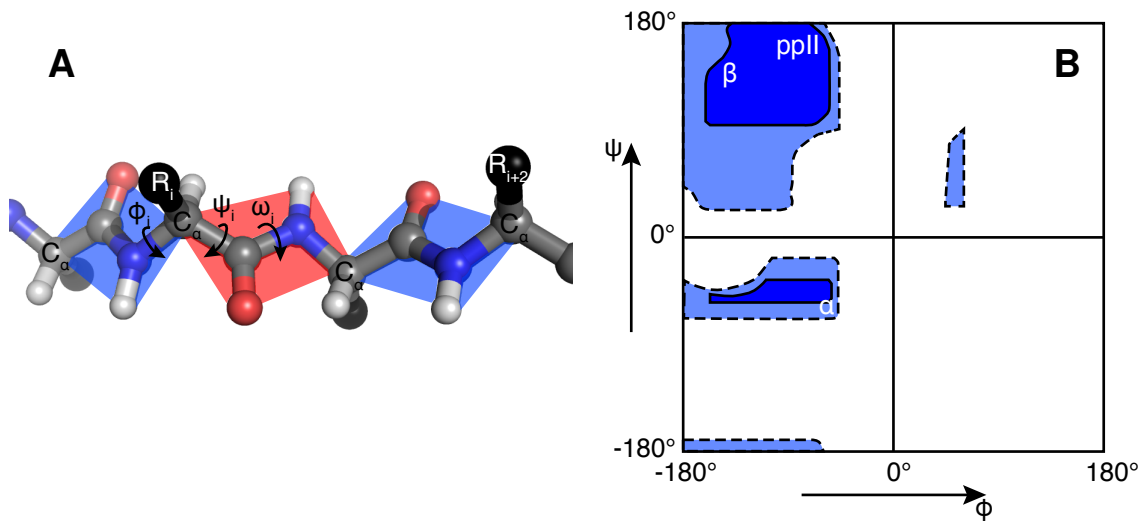


Figure 1.1.: A: Schematic representation of a polypeptide chain with the rotation angles ϕ , ψ and ω shown. Carbonates shown in gray, nitrogens in blue, oxygens red, hydrogens white. The positions of the side chains are marked in black. B: Allowed (dark blue) and outer limits (light blue) of ϕ - and ψ -angle combinations as proposed originally by Ramachandran¹⁵. The regions of the most important secondary structure elements are indicated by white labels (i.e., α -helix, β -strand and poly-proline II conformation).

1.1.2. Properties of polypeptide chains and the native structure

The native structure of proteins is strongly related to molecular properties of the polypeptide chain. The amino acids share the same backbone chemical structure, but differ in their side chains. The relative spatial arrangement of the atoms of an amino acid in a polypeptide can be depicted by the two backbone dihedral angles ϕ and ψ as well as by the angle ω of the peptide bond. The orientation of the side chain of a residue relative to the backbone is given by the χ_j -angles (see Figure 1.1 A). Considering steric repulsion between different atoms, the dihedral angles ϕ and ψ can only adopt certain combinations, which are indicated by the allowed regions in a Ramachandran-plot (see Figure 1.1 B).¹⁵ The partial double-bond character of the peptide bond between two consecutive amino acids prohibits fast rotation and allows only two stable conformation with $\omega \approx 0^\circ$ (*cis*) and $\omega \approx 180^\circ$ (*trans*), so that the backbone atoms C_α , C, O, N^{i+1} , H^{i+1} and C_α^{i+1} are (approximately) situated in a plane. In native structures, around 5–7 % of prolyl-prolyl bonds are in the *cis*-conformation¹⁶, while the population of the *cis*-conformation for secondary amide bonds is below 0.5 % in small unstructured peptides¹⁷ and below 0.1 % in native structures^{18,19}. In summary, steric effects within the polypeptide chain strongly constrict the conformational space available to a

polypeptide.

A variety of attractive, non-covalent interactions between different chemical moieties of the polypeptide may be formed. Those comprise dipole-dipole-interactions (hydrogen bonds, Van der Waals interactions), coulomb-interactions (salt bridges) and in aqueous solution also hydrophobic interactions. Together with steric repulsion, these interactions between backbone or side chain groups determine the native structure. The exact shape of the energy landscape in ϕ - ψ -space varies slightly with the nature of the side chains of two consecutive residues.²⁰ In addition, the energy landscape is influenced by interactions spanning the distance of several residues along a polypeptide chain.²¹

In 1951, Pauling observed that several structural patterns with a maximum number of satisfied backbone hydrogen bonds can be formed by the polypeptide chain. These *secondary structure* elements like α -helices and β -strands can be identified by their ϕ - ψ -angles in the Ramachandran-plot (see Figure 1.1 B). The secondary structure elements arrange in 3-dimensional super structure, the *tertiary structure*, stabilized by interactions between groups that may be far apart in sequence. Large proteins also possess a *quaternary structure* built up of several tertiary subunits. The first protein structure determined by X-ray diffraction was the one of Myoglobin in 1958.²² Since, structure determination at atomic resolution by X-ray crystallography²³ or by solution nuclear magnetic resonance (NMR)^{24,25} has become routine and a multitude of structures have been solved up to date.

1.2. Protein folding and the native state

One of the most remarkable properties of proteins is the fact, that denatured (unfolded) proteins fold into their unique native structure when transferred to folding conditions.⁵ This gives evidence that, despite the large number of possible interactions, the native structure is defined solely by the amino acid sequence.

1.2.1. Protein stability

Despite their functional and structural diversity based on structural sensitivity to sequential changes²⁶, the folded state of a protein needs to be stable against thermal fluctuation or external stress in a cellular environment. The stability of a protein is the result of a compensation between large stabilizing and large destabilizing energetic contribution from favorable or unfavorable interactions within the protein or between protein and solvent.^{27,28}

1. Introduction

This leads to a rather small overall difference and thus to stabilization energies between -10 and -60 kJ/mol^{29,30}. As a consequence, structure prediction from sequence, which is based on accurate calculation of stabilizing and destabilizing energies, remains challenging up to date.^{31,32}

The thermodynamic stability of the folded state N of a protein relative to an unfolded state ensemble U is given by the following equation:

$$\Delta G^0 = -RT \ln(K_{\text{eq}}), \quad \text{with } K_{\text{eq}} = \frac{[\text{N}]_{\text{eq}}}{[\text{U}]_{\text{eq}}} \quad (1.1)$$

and the change of the free energy difference is given by the Gibbs fundamental equation³³:

$$d\Delta G^0 = \Delta V^0 dp - \Delta S^0 dT + \sum_i \Delta \mu_i^0 dn_i \quad (1.2)$$

ΔV^0 is the difference in volume, ΔS^0 the difference in entropy and $\Delta \mu^0$ the difference in chemical potential each between U and N. The stability ΔG^0 can be changed by either varying the temperature T , the solvent composition n_i or the pressure p . The stability of a protein is generally determined from its thermal or denaturant induced unfolding transition.³⁴ In particular, protein stability is defined relative to the unfolded state. While intra-molecular interactions in the folded state can be easily assessed through the crystal- or NMR-structure, the exact contributions from interactions within an unfolded polypeptide chain or between the unfolded chain and solvent molecules is unknown. Hydrogen bonds are believed to play a crucial role in this context, as the number of hydrogen bonds per residue formed between an unfolded polypeptide and solvent molecules is about the same as the number of intra-chain hydrogen bonds in folded structures.³⁵ A better characterization of the unfolded state and interactions therein is therefore highly required.

1.2.2. Protein folding

In 1969 Levinthal stated his famous paradox for the folding time of a protein.³⁶ Based on the assumption of a random search of all possible conformations he estimated folding to take an astronomical amount of time until the correct folded state could be reached, but folding times of milliseconds to seconds can be observed in experiments^{29,37}. Obviously, specific pathways are followed to ensure efficient folding.

The existence of folding intermediates can optimize the search for native interactions and

speeds up folding.³⁸ It was also shown, that an energetic bias in the unfolded state towards the native conformation can reduce folding times down to observed values.³⁹ In order to solve the Levinthal's paradox, several theories have been proposed about the concrete folding mechanism. Sequential folding proceeding through intermediates^{40,41} or through molten globule states^{42,43}, diffusional collision of preformed micro domains small enough to rapidly search all possible conformations^{44,45}, or condensation of the chain around a preformed structural nucleus in an apparent two state manner (nucleation-condensation)^{46,47} have been proposed as mechanisms.

Even though intermediate states were observed for some proteins, the folding of many other, especially small proteins was often interpreted to follow a two-state mechanism. However, careful analysis of folding kinetics gave evidence for transiently populated intermediates for many of these apparent two-state proteins.^{48,49} Recently, even very complex pathways with (several) on- and off-pathway intermediates or parallel folding pathways were experimentally resolved through hidden Markov analysis.^{50,51}

However, a unifying description of protein folding is still under investigation and the exact mechanism of protein folding remains a challenging to conceive^{52,53}.

1.2.3. The energy landscape of protein folding

The folding of many small proteins can be described by a two-state mechanism between U and N with folding rate k_f and unfolding rate k_u .²⁹



Even though U consists of a multitude of different conformations it acts as one state if interconversion between different conformations is much faster than the folding or unfolding rate.⁵⁴ A relation between the kinetic rate constants and the height of the free energy barrier $\Delta G_{f/u}^{\ddagger}$ separating the two states is given from transition state theory.⁵⁵

$$k_{f/u} = k_0 \exp\left(-\frac{\Delta G_{f/u}^{\ddagger}}{RT}\right) \quad (1.4)$$

However, the pre-exponential factor k_0 is unknown for protein folding. Folding and unfolding rate can be determined from the re-equilibration kinetics after transferring denatured

1. Introduction

protein back to folding conditions or by transferring folded protein to unfolding conditions. The stabilities determined from kinetic measurements and from unfolding transitions should be the same if the protein is well described by a two-state mechanism.

To further characterize the free energy barrier and the transition state changes of the conditions x (i.e., temperature or pressure) can be applied. The relation between the change of the activation free energy ΔG_f^{\ddagger} to the change of the equilibrium free energy ΔG_f° can follow a linear relationship (rate equilibrium free energy relationship or REFER) with proportionality parameter α_x .⁵⁶

$$\alpha_x = \frac{\partial \Delta G_f^{\ddagger} / \partial x}{\partial \Delta G_f^{\circ} / \partial x}, \quad \text{with } 0 \leq \alpha_x \leq 1 \quad (1.5)$$

If α_x is close to 1, the transition state resembles the native state, whereas if α_x is close to 0, the transition state resembles the unfolded state in respect to the property probed by the disturbance (ΔV_f^{\ddagger} for changes in p , ΔS_f^{\ddagger} for changes in T , applying the Gibbs fundamental equation (1.2) to the transition state). In an analog approach, the presence of a side-chain interaction in the transition state can be tested by site-directed mutagenesis disturbing the interaction. The proportionality parameter determined in analogy to equation (1.5) is called ϕ_f -value.⁵⁷⁻⁶⁰ A ϕ_f -value of 1 indicates that the interaction is already formed in the transition state. However, as all changes are interpreted relative to the ground state (U), effects on this state can lead to misinterpretations. Especially, if the unfolded state has residual interactions the disturbance may also alter the unfolded state. A comprehensive analysis of kinetic data can reveal these effects, which were found to be present in quite a few proteins.⁶¹ This indicates for a complex nature of the unfolded state ensemble, with residual interactions and sensitivity to changes of temperature or denaturant.

1.3. The unfolded state

While the native structure can be studied by X-ray diffraction or solution NMR and has been major focus of research, the unfolded state is far less well characterized. However, the unfolded state plays a crucial role in the entire folding process. The unfolded state serves as the starting point the folding reaction. The energetic constitution in the unfolded state is the reference for the stability ΔG_f° of the native structure.⁶² Furthermore, properties of the unfolded state influence folding speed and can have implications for the mechanism of

folding.

In recent years, interest has also focused on proteins lacking the native structure under physiological conditions, the so called intrinsically disordered protein (IDP)s. Characterizing the properties of this disordered state may help to understand how these proteins function, i.e. the role of the unfolded state in the mechanism of binding targets⁶³ and the physiological benefit of disorder^{64,65}.

The unfolded state ensemble U of a protein comprises a large number of different conformations which rapidly interconvert through temperature driven Brownian motion. It is populated under physiological conditions for IDPs and unstructured model peptides. Moreover, it can be populated by shifting the equilibrium (equation (1.2)) towards the unfolded state for example by adding denaturants like GdmCl and urea to the protein solution^{66,67} or through an increase in temperature. The denaturant induced unfolded state is often referred to as denatured state. The unfolded state is intrinsically difficult to characterize in experiments under folding conditions. Therefore, unstructured sequences derived from IDPs or unfolded repeat sequences serve as model systems for experimental studies.

The unfolded state can easily be identified by its signature in the circular dichroism (CD) spectrum of the protein in solution, which shows a strong negative band near 200 nm and either a slightly positive or negative shoulder at 220 nm.⁶⁸⁻⁷¹

The unfolded ensemble is not identical under all these different conditions and presumably its properties also depend on the sequence of the protein. An adequate model description of the characteristics of the unfolded state, i.e. the conformations and dynamics, could improve the understanding of protein folding and function.

1.3.1. The random coil model

An approach to capture conformational properties of unfolded proteins is based on the application of polymer models. The *freely jointed* or *random flight* model, which is mathematically equivalent to the *Gaussian chain* model is the conceptually most simple model for polymers.^{72,73} It describes the properties of a chain with n infinitely small beads separated by bonds of length l . The joint at the beads are allowed to rotate freely with random distributions of the rotation angles and without memory of the rotation angles of the preceding joint. In the continuum approximation (for very large chains), the probability $p_{\text{eq}}(r)$ of finding the two ends of the chain at a distance r from each other can be described by the

1. Introduction

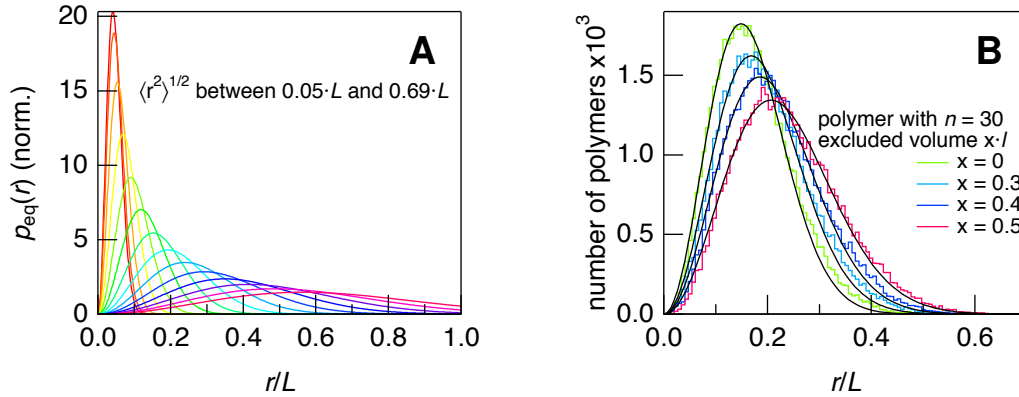


Figure 1.2.: A: End-to-end distance distribution of a 3-dimensional random coil polymer with total length $L = nl$; B: Simulation of a 3-dimensional random walk of a polymer of $n = 30$ segments (—) and of polymers with excluded volume modeled through solid spheres of diameter $x \cdot l$ around the joints between the segments (---, -.-, -.-). Fits with equation (1.6) (—). Random coil polymers of n segments with excluded volume show end-to-end distance distributions similar to those of longer polymers with $n'l' > nl$.

following probability distribution function⁷³:

$$p_{eq}(r) \propto 4\pi r^2 \exp\left(-\frac{3r^2}{2\langle r^2 \rangle}\right) \quad (1.6)$$

For the Gaussian chain model, the root mean square end-to-end distance is related to the number of segments:

$$\langle r^2 \rangle^{1/2} = l \cdot n^{1/2} \quad (1.7)$$

and the radius of gyration r_G of the Gaussian chain is related to the mean square end-to-end distance:^{72,74}

$$\langle r_G^2 \rangle = \langle r^2 \rangle / 6 \quad (1.8)$$

This relation is often used to compare results from experiments sensitive to $\langle r^2 \rangle$ to results of experiments sensitive to the $\langle r_G^2 \rangle$ or vice versa. The random coil model is widely used to model distance distributions in unfolded or intrinsically disordered proteins in experiments sensitive to intra-chain distances.^{75–95}

If the beads of the random coil polymer possess a finite size, the conformational space for placing the beads is restricted and the distance distribution $p_{eq}(r)$ is shifted towards larger r , but retains the shape described by equation (1.6) (see Figure 1.2 B). Therefore scaling

exponents ν different from the value $\nu = 1/2$ in equation (1.7) were introduced to describe the relation between the radius of gyration and the number of chain segments n in order to allow a comparison between theory and experiment for these cases, where an analytical solution is no longer available.⁹⁶

$$r_G = \rho_0 \cdot n^\nu \quad (1.9)$$

Values for the scaling exponent are indicative for the nature of the polymer. An expanded coil (with excluded volume) leads to a value of or close to $\nu = 3/5$ ^{96,97}, a random coil to a value of $1/2$ ^{96,98} and a compact globule to a value of $1/3$ ^{99,100}. The GdmCl denatured state of different proteins was shown to follow a scaling law with $\nu = 2/3$. The interpretation of this finding was, that GdmCl unfolded proteins are random coils and that secondary structure is completely removed by the denaturant.¹⁰¹

1.3.2. Characteristics of the ensemble of unfolded conformations

Radii of gyration measured by small angle X-ray scattering (SAXS)¹⁰² as well as hydrodynamic radii determined by NMR¹⁰³ were shown to obey the scaling law of equation (1.9). Model ensembles of unfolded protein conformations generated by Monte Carlo sampling of the dihedral angles also obey the scaling law of equation (1.9) supporting the interpretation of random coil behavior of unfolded proteins.¹⁰⁴ However, several observations strongly contradicted this interpretation. Acid or heat denatured proteins did not show random coil characteristics¹⁰⁵ and also some proteins were found to act like outliers from the scaling law equation (1.9)¹⁰⁶. Persisting structure^{107–109} or interactions¹¹⁰ were observed for many different unfolded or chemically denatured proteins. Also the observation, that persistent structure in the unfolded state may even show a topology similar to that of the native state^{111–119} has great implications not only for the lack of random coil behavior in the unfolded state, but also for folding speed^{118,119} as it gives evidence of an energetic bias towards the native state.

The paradigm of these contradicting results can be resolved in view of two key-publications from 2004. Fitzkee and Rose showed, that ensembles of conformations derived from a folded structure with only 8 % of flexible residues and 92 % of residues keeping their native secondary structure obey the scaling law of equation (1.9)¹²⁰ reinforcing early warnings about drawing conclusions about random coil behavior solely from scaling laws^{101,121}. Additionally, Jha et al. were able to construct theoretical model ensembles for several unfolded

1. Introduction

proteins simultaneously showing random coil scaling of the radius of gyration and the presence of significant amounts of local backbone structure as observed by NMR.¹²² In the whole, conformational properties of unfolded proteins, like for instance residual structure, are not in contradiction with the scaling law of equation (1.9).

Several experimental technics have been established to obtain structural information on unfolded or disordered proteins. Applicable NMR techniques comprise the analysis of chemical shifts (CSs), residual dipolar couplings (RDCs), paramagnetic relaxation enhancement (PRE) and Spin relaxation. CSs give useful information on the local environment of nuclear spins (of ^1H or ^{13}C) for example in IDPs when compared to CSs of reference peptides^{123,124} or of denatured proteins^{125–131}. RDCs between two nuclear spins (^1H or ^{15}N) can be detected if the proteins are oriented in a weak alignment medium (i.e., polyacrylamide gels or other) for regions of the protein lacking dynamic averaging. Therefore, RDCs report on local conformational order along the backbone.^{132–134} For the observation of transient contacts between parts of protein far apart in sequence, a spin label or extrinsic paramagnetic center can be brought into the polypeptide chain changing the relaxation of nuclear spins of the unfolded protein within a distance of 20–35 Å.^{135–137} SAXS can detect the radius of gyration r_G of particles in a mono-disperse solution through Guinier analysis.¹³⁸ In case of unfolded or (partly) disordered proteins the scattering intensity is a linear combination of the scattering intensities of all present radii of gyration¹³⁹ and can be obtained by help of atomistic models of the protein^{140–146}.

Fluorescence studies based on fluorescence resonance energy transfer (FRET) require the incorporation of two label molecules into the protein. The energy transfer rate is sensitive to distances between 10–100 Å dependent on the choice of labels and can therefore be used to determine distance distributions ($p_{\text{eq}}(r)$) or average distances ($\langle r^2 \rangle$) in unfolded proteins as will be discussed in detail in section 1.5.2.

Residue specific, high resolution, structural information can be obtained by combing these experimental methods with Monte Carlo based generation of ensembles of unfolded conformations and algorithms selecting sub-ensembles based on experimental parameters.^{147–149} This has been successfully applied to data from NMR^{136,150,151} and SAXS experiments^{143,145}, including combinations of both techniques^{134,152–155}. Sampling of dihedral angles can be performed based on dihedral angles of residues in coil conformations¹⁵⁶ not involved in secondary structure like α -helix and β -strand obtained from high resolution crystal structures.^{152,157}

Accordingly, the picture of the unfolded state ensemble drawn by these methods is highly

diverse. The unfolded ensemble is well described by conformations with ϕ - ψ -angles also found in folded structures, so they sample the same allowed regions of the Ramachandran plot, even though there are position dependent preferences of certain regions. For example, regions with preferential sampling of α/β regions or of polyproline II (ppII) regions are identified in IDPs.^{127,151} Furthermore, residue resolved, native-like contacts are found by a combination of RDCs, PREs and restricted molecular dynamics simulations in denatured ubiquitin giving indication for an energetic bias towards the folded state.¹³⁷ Despite the limitations mentioned above, scaling exponents still give useful information and can be used to classify unfolded/unstructured proteins as globular (folded), molten globule, pre-molten globule or native coil.¹⁵⁸ Single molecule (sm) FRET points out that the scaling exponent ν changes with the amount of denaturant and also varies with sequence, especially with the content of charged residues in IDPs.⁹¹

1.3.3. The wormlike chain and the chain with excluded volume

The findings summarized in the preceding section indicate, that the random coil model may be insufficient to describe overall properties of unfolded proteins due to the strict linkage between the width of the distribution and its center of mass as shown in Figure 1.2 A. In the following, two additional polymer models are discussed. The shape of the distance distribution for both models is determined by two independent parameters, which might allow a more accurate approximation of distances in real peptides.

The *wormlike chain model*¹⁵⁹ describes a rather stiff, almost rod-like polymer. The polymer has an overall length L , which is called contour length in this context, and a persistence length l_p , the length over which correlations in the direction of the tangent of a point along the chain are lost. The probability distribution of end-to-end distances $p_{\text{eq}}(r)$ for a worm-like chain can be written as¹⁶⁰

$$p_{\text{eq}}(r) \propto 4\pi r^2 \frac{1}{\left(1 - \frac{r^2}{L^2}\right)^{9/2}} \exp\left(-\frac{9L}{8l_p} \frac{1}{\left(1 - \frac{r^2}{L^2}\right)}\right). \quad (1.10)$$

If the persistence length l_p is increased, the (normalized) distance distribution changes from an almost symmetrical shape to a highly asymmetrical shape for large l_p (see Figure 1.3 A). The worm-like chain model has been applied to unfolded proteins in many occasions, i.e. to determine the persistence length of proteins unfolded by applying force to the ends^{161–163}, to model loop regions in proteins¹⁶⁴ and to compare properties like chain stiffness of different

1. Introduction

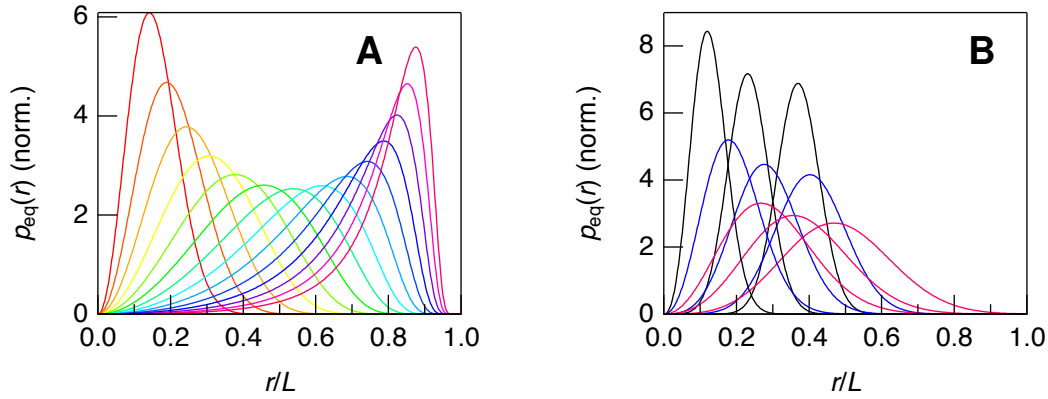


Figure 1.3.: A: Change of the distance distribution $p_{eq}(r)$ for the wormlike chain model. The ratio l_p/L is increased with colors from red, over green and blue to purple. B: Distance distribution $p_{eq}(r)$ for the Edwards model. An increase of parameter b leads to a shift of the distribution towards larger distances r (same color), while an increase of σ increases the width of the distribution (colors from black over blue to pink).

proteins ^{165,166}.

Another model describing a *polymer with excluded volume* was proposed by Edwards ¹⁶⁷. In this model a chain consisting of an infinite number of segments of length l expands outwards from the origin of the coordinate system. The probability of finding the n th segment along the chain at a distance r from the origin depends on the excluded volume v resulting from a repulsive potential between two segments of the polymer. ¹⁶⁷

$$p_{eq}(r) \propto 4\pi r^2 \exp \left[- \frac{\left(r - \left(\frac{5}{3}\right)^{\frac{3}{5}} \left(\frac{v}{3\pi l}\right)^{\frac{1}{5}} (nl)^{\frac{3}{5}} \right)^2}{\frac{20nl^2}{27}} \right] \quad (1.11)$$

Here, a simplified notation of this equation will be used.

$$p_{eq}(r) \propto 4\pi r^2 \exp \left[- \left(\frac{r - b}{\sigma} \right)^2 \right] \quad (1.12)$$

Both parameters b and σ are generally varied, as n and l may not be directly related to properties of the peptide backbone. This model has been successfully applied to model distance distributions of unfolded peptides or short protein fragments. ^{168–175}

Unfortunately, no analytical expression is known to relate the distance distribution and the

radius of gyration for both of these two models. Therefore, knowing the position of the ends of the polymer does not provide any knowledge about the position of the middle part of the polymer. The currently available experimental methods, which are sensitive to distance distributions, are insensitive to the radius of gyration, so only a combination of different methods can provide insight on the relation between distance distributions and the radius of gyration. Still, the use of polymer models provides good approximations for distance distributions and average distances in unfolded proteins.

1.4. Dynamics in the unfolded state

Dynamics in proteins or polypeptides in general describe changes between different conformations. On the one hand, dynamics can occur within an ensemble of energetically similar conformations, which can be separated by barriers of the order of $k_B T$ or smaller. This is the case for temperature driven Brownian motion in unfolded proteins and leads to sampling of conformational space through intra-chain diffusion. On the other hand, dynamics can also occur between discrete conformations or conformational ensembles separated by barriers in free energy much larger than $k_B T$. The methods used to determine dynamics or the underlying processes probing the dynamics are required to be faster than the dynamics themselves, otherwise only the dynamic average is accessible.

1.4.1. Methods to study dynamics in the unfolded state

Dynamics in proteins occur on many different time scales. They comprise different processes like local rotations of side-chains, rearrangements in the unfolded state or in folding intermediates as well as mediated function in enzymes. Nowadays, methods suitable for studying dynamics span the whole time range relevant for proteins (see Figure 1.4). However, not for all timescales comprehensive insight is provided, so that the entire picture of dynamical process remains incomplete up to date.

Dynamics in the unfolded state take place on the timescale from early nanoseconds (ns) to microseconds (μ s). Methods in this range typically determine one of the following parameters: rate constant for intra-chain loop formation k_c between two sites along the unfolded chain, the mean square displacement of one site relative to another site along the chain captured by the intra-chain diffusion coefficient D or the chain reconfiguration time τ_r . Also methods determining the regions of conformational space sampled by the protein give some

1. Introduction

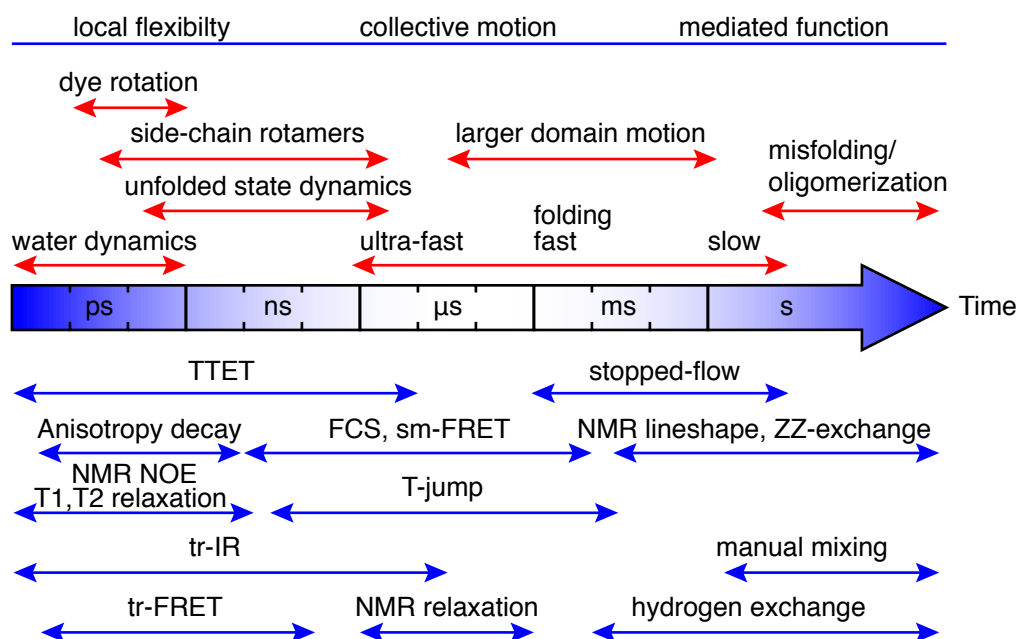


Figure 1.4.: Timescales of dynamics in proteins in the unfolded state and along the folding path (red). Methods used to study dynamics on different timescales (blue). Figure partly adopted from^{90,176,177}.

idea of dynamics and time-scales¹⁷⁸.

Rate constant for intra-chain loop formation can be directly determined by monitoring the depopulation of an excited state of a chromophore attached to the peptide or protein upon van-der-Waals contact with another photo-physically active compound, if the photo-physical process is diffusion controlled. Triplet-triplet energy transfer (TTET) between xanthone and naphthalene is suitable to determine k_c , as both, the excitation of xanthone-triplet through intersystem crossing, as well as the Dexter-type¹⁷⁹ energy transfer to naphthalene, are very fast (≈ 2 ps).^{180,181} The triplet state of xanthone shows a lifetime of 30 μs in degassed aqueous solution making it very useful to study peptide dynamics.¹⁸² Alternatively, also energy transfer between Zn-porphyrin and a Ru(NH₃)₅ in cytochrome c have been used to obtain k_c .¹⁸³ Furthermore, photo-physical processes like tryptophan triplet quenching^{93,165,166,184–188}, photon induced electron transfer (PET) between an oxazine derivative and tryptophan^{189,190} or FRET with the long lifetime donor Dbo^{191–194} have been also been applied. However, if the process is not diffusion controlled, further calculations and the assumption of models are

required to obtain k_c .

The intra-chain diffusion coefficient D describing the relative motion of two dye molecules attached to a peptide can be determined together with the distance distribution $p_{\text{eq}}(r)$ from analysis of time-resolved FRET experiments.^{170–175,195}

This approach was first proposed by Haas et al. for unfolded polypeptides.^{170,171} Applied to study polymers, the approach was advanced^{196–199} and successfully applied to study the dynamics in a variety of peptides^{170,172–174,200}. The influence of intra-chain diffusion D and $p_{\text{eq}}(r)$ can be described by a partial differential equation from which a fitting function for time resolved data can be obtained. However, the parameters describing $p_{\text{eq}}(r)$ and the intra-chain diffusion coefficient D can be strongly correlated in the fit.^{170,171} To resolve this correlation the assumption was made that $p_{\text{eq}}(r)$ remains unchanged under addition of a viscous co-solvent¹⁷⁰ or when changing the temperature²⁰¹, which may not be entirely justified. Another possibility is global analysis of measurements done with either two different fluorescence acceptors^{197,198} or two different donors^{172,199}. Approaches based on changing the properties of the fluorescence donor by adding quenching compounds to the solution¹⁹⁶ may not be suitable for resolving correlations²⁰². In some cases, $p_{\text{eq}}(r)$ and D can be directly obtained with relatively large error from one measurement.^{173,174,200} Alternatively, intra-chain diffusion coefficients D are calculated from quenching rate constants k_q in combination with numerically obtained distance distributions $p_{\text{eq}}^{\text{num}}(r)$ based on the assumption of validity of equation (1.19).²⁰³

Reconfiguration times can be obtained from ns fluorescence correlation spectroscopy (FCS) measurements through the determination of correlation times of photon emission of FRET chromophores.^{75,83,88} Further information from sm FRET experiments is contained in FRET efficiency histograms. In these histograms, photon shot noise from the dye fluorescence and noise from the detection system lead to broadening of the distribution of FRET efficiencies. Several methods have been developed to predict the width of the distribution dictated by the experimental settings.^{204–206} If the observed width of the histogram is wider than the width expected from shot noise, this can be interpreted as conformational dynamics occurring on the timescale slower than the binning of the experiment.²⁰⁷

Accessing dynamics by molecular dynamics (MD) simulations is still challenging nowadays, especially for proteins of around 50 residues or more, due to the limited time window for one simulated trajectory. In a work-around, dynamics can be achieved even from short trajectories by applying a Markov model, identifying the Markov states and the rate constants of interconversion connecting them.^{208–212} By linking some of the states to for

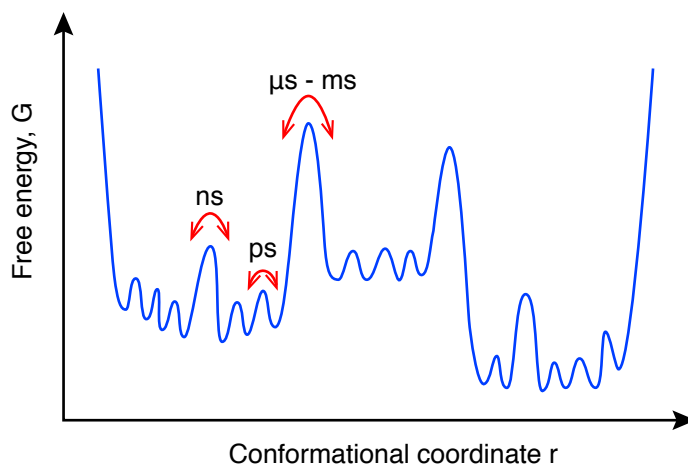


Figure 1.5.: One dimensional representation of the free energy landscape underlying the hierarchy of dynamical processes of a protein. The timescale of interconversion between different minima is determined by the height of the corresponding barrier. Chain diffusion comprises local structural changes for fast processes and larger rearrangements for slower processes. Figure based on references^{215,219,220}.

example photo-physical activity of a dye molecule, simulations can be compared to experiments.²¹³ In unfolded small peptides, dynamic parameters, like the rate constant for intra-chain loop formation between two sites, can be determined from the trajectories using a first passage time approach.²¹⁴

1.4.2. Dynamics in unfolded peptides and proteins

Parts of the dynamics of unfolded as well as folded proteins in solution are slaved to the dynamics of the solvent. Water fluctuations can be divided into two processes, one of which depends on viscosity, the second is identified by its Arrhenius like temperature dependence.²¹⁵ In folded proteins in solution, fluctuations are dependent on water dynamics and have their origin either in the bulk solvent or in the hydration-shell.^{216,217} In a water/glycerol mixture at ambient temperature, the viscosity dependent fluctuations occur on a timescale of 10 ns, while the temperature dependent fluctuations occur on a much faster timescale of 10 ps. Fluctuations can be assigned to hierarchically organized tiers (see Figure 1.5).^{215,218} Dynamics in small unfolded peptides studied by TTET measurements also show motions on different hierarchical level as indicated by different timescales of dynamics in the ps as well as in the ns range.²¹⁹ However, in unfolded systems, dynamics may not only be slaved

by water dynamics, but are strongly influenced by barriers imposed by the backbone itself through bond rotations or by interactions. On the ns timescale, loop formation of two sites of an unfolded peptides is well described by a single exponential kinetic in agreement with theory (equation (1.18)).¹⁸² The loop formation rate constants k_c of model repeat peptides like poly-(Gly-Ser) or poly-Ser peptides are length independent for short loops, and length dependent for longer loops of more than 10 residues obeying a scaling law similar to equation (1.9).^{182,221–223} The loop formation rate is influenced by the sequence of the peptide, for instance, the end-to-end loop formation is faster for poly-(Gly-Ser) peptides than for poly-Ser peptides.^{221,224} Loop formation is slowed down, if the N-terminal or C-terminal parts of the loop are extended by tail sequences.²²⁵ The temperature dependence of loop formation rate constants follows an Arrhenius law with activation energies in the 5–30 kJ/mol range.^{194,221,223,226} The viscosity η of the solvent strongly influences loop formation. End-to-end loop formation rate constants k_c are found to follow a $\eta^{-\beta}$ dependence in mixtures of water with viscous co-solutes.²²⁷ In this context, $\beta = 1$ indicates, that the full solvent viscosity η acts on the peptide, $\beta < 1$ indicates that loop formation is not entirely controlled by diffusion. This relation can also be used to correct for effects of changes of solvent viscosity. The viscosity corrected loop formation rate k_c^{corr} slows down with increasing denaturant concentration.²²⁷

Intra-chain diffusion coefficients D are not as systematically characterized as loop formation rate constants. Determined by the analysis of time resolved fluorescence measurements, intra-chain diffusion is found to increase with rising temperature for different fragments derived from α -Synuclein¹⁷⁴ and also shows a trend to increase with increasing peptide length¹⁷³. The intra-chain diffusion coefficients observed in experiments are on the order between 0–50 $\text{\AA}^2/\text{ns}$ in aqueous solution at ambient temperature.^{172–175} Intra-chain diffusion coefficients calculated from quenching rate constants k_q are found to be between 0.01 and ≈ 200 $\text{\AA}^2/\text{ns}$ for different proteins²²⁸ and to decrease with increasing temperature for different variants of α -Synuclein²⁰³.

1.4.3. Theory of diffusion with focus on intra-chain diffusion

To describe the behavior of particles subject to Brownian motion, Newton's equations are extended by a stochastic force $s\xi(t)$ leading to the Langevin equation^{229,230} or in presence

1. Introduction

of an external force ($\mathbf{F}(\mathbf{r})$) to the Smoluchowski equation²³¹.

$$m\ddot{\mathbf{r}} = -f\dot{\mathbf{r}} + s\xi(t) + \mathbf{F}(\mathbf{r}) \quad (1.13)$$

In this equation m is the mass of the particle, \mathbf{r} its position vector and f is the friction coefficient. In the limit of strong friction f , the corresponding Fokker-Planck equation for the probability p to find the particle at point \mathbf{r} at time t when it started from point \mathbf{r}_0 at time t_0 can be derived.²³²

$$\frac{\partial}{\partial t}p(\mathbf{r}, t|\mathbf{r}_0, t_0) = \left(\nabla^2 D - \nabla \frac{\mathbf{F}(\mathbf{r})}{f} \right) p(\mathbf{r}, t|\mathbf{r}_0, t_0) \quad (1.14)$$

The diffusion coefficient D is defined through $D = s^2/(2f^2)$. In most cases D is assumed to be independent of the position \mathbf{r} of the particle, but in general a position dependence of the diffusion coefficient $D(\mathbf{r})$ is possible. Combined with the Boltzmann-distribution, this leads to the *fluctuation-dissipation theorem*²³³,

$$\nabla D = \mathbf{F}(\mathbf{r}) \left(\frac{1}{f} - \frac{D}{k_B T} \right), \quad (1.15)$$

which states that the response of the system to an external force is the same as the response to an internal fluctuation, so the relaxation of the system is related to its statistical fluctuations. Including this, the Smoluchowski equation for a force $\mathbf{F}(\mathbf{r}) = -\nabla U(\mathbf{r})$ can be written in the following form:²³²

$$\frac{\partial}{\partial t}p(\mathbf{r}, t|\mathbf{r}_0, t_0) = \nabla \cdot D e^{-U(\mathbf{r})/k_B T} \nabla e^{U(\mathbf{r})/k_B T} p(\mathbf{r}, t|\mathbf{r}_0, t_0) \quad (1.16)$$

This equation describes diffusion in a potential $U(\mathbf{r})$.

Another important consequence of the fluctuation dissipation theorem is the *Einstein-Smoluchowski equation* for position independent diffusion in solution, relating the diffusion coefficient D to the friction coefficient f .^{234,235}

$$D = \frac{k_B T}{f} \quad (1.17)$$

The friction coefficient f depends on solvent viscosity η .

The above stated relations apply to free diffusion as well as to intra-chain diffusion between

two points in an unfolded protein or peptide, as both processes can be projected on one reaction coordinate and can therefore be seen as Brownian motion in one dimension. Thus, in the following, D will denote the intra-chain diffusion coefficient.

In a polypeptide subject to intra-chain diffusion, two sites along the chain can encounter with a certain frequency. The probability $p(t)$, that the two sites have not encountered each other for the first time, can be approximated to decay single exponentially over time.^{236–238}

$$p(t) \propto e^{-k_c t} \quad (1.18)$$

In this equation, k_c is the rate of loop formation. It is intuitively clear, that k_c depends on the intra-chain diffusion coefficient $D(r)$ and on the equilibrium distance distribution of the two points $p_{\text{eq}}(r)$. In the *mean first passage time* approach, the relation is given by:²³⁸

$$k_c = \left(\int_{r_{\min}}^{r_{\max}} dx [D(x)p_{\text{eq}}(x)]^{-1} \left[\int_x^{r_{\max}} dy p_{\text{eq}}(y) \right]^2 \right)^{-1}. \quad (1.19)$$

However, one of the major defects in this theory, is the implicit assumption, that the relaxation of the distribution towards equilibrium is governed by one single time constant k_c , i.e., one intra-chain diffusion coefficient $D(r)$.²³⁸

The existence of Brownian motion and intra-chain diffusion has other implication for protein folding. In order to cross a free energy barrier ΔG^{\ddagger} on the folding pathway, the polypeptide chain has to undergo intra-chain diffusion. The so called *Kramers's theory* describes a diffusional attempt for crossing a barrier ΔG^{\ddagger} in free energy. Modeled as a one-dimensional process in the high viscosity limit, the rate k for crossing the barrier is given.²³⁹

$$k \cong \frac{2\pi\omega\omega'}{\eta} \exp\left(-\frac{\Delta G^{\ddagger}}{k_B T}\right) \quad (1.20)$$

In this equation ω and ω' are the frequencies of the harmonic approximations of the potential-well preceding the barrier and of the potential at the barrier top. The influence of diffusion is not evident on first sight, but went into the derivation through the viscosity η , which is assumed to be the only source of friction and thus $\eta = f = k_B T/D$ (compare equation (1.17)).

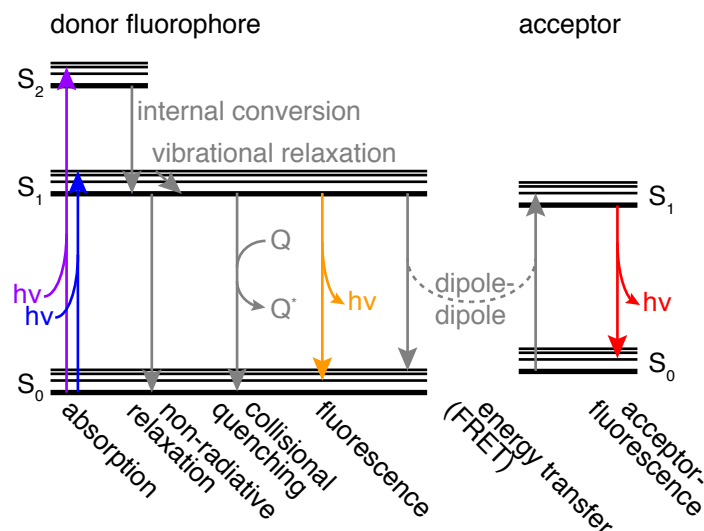


Figure 1.6.: Jablonski diagram illustrating the electronic states and the electronic transitions within a fluorophore or between donor fluorophore and acceptor. Processes involving the absorption or emission of photons are drawn in color with photon energy decreasing with colors from violet over blue and yellow to red. Processes involving the triplet state like for instance intersystem crossing or phosphorescence are not shown.

1.5. Principles of Fluorescence Methods

1.5.1. Fluorescence

The emission of photons from molecule which are in an electronically excited singlet state is called fluorescence. Typical fluorophores are aromatic molecules. Upon absorption of light, they are brought to an excited singlet state, which rapidly relaxes through internal conversion or by dissipating vibrational energy (within about 1 ps) to the lowest excited singlet state S_1 . From there, return to the ground state is possible either through emission of a photon, through non-radiative processes that dissipate energy to the surrounding or through quenching from other compounds in solution (compare Jablonski diagram in Figure 1.6). For most fluorophores, the depopulation of S_1 follows a single exponential with a decay time τ which lies in the range of up to 10 ns for most fluorophores. The fluorescence emission spectrum is shifted towards lower frequencies (higher wavelengths) due to the Stokes shift.²⁴⁰ The shape of fluorescence spectrum resembles a mirror image of the absorbance spectrum in many cases. The spectrum, as well as the decay time τ , is mostly independent of excitation wavelength due to fast internal conversion and vibrational relaxation to the

lowest excited singlet state S_1 (Kasha's rule).²⁴¹ The ratio between the number of photons emitted by a fluorophore to the number of photons absorbed is called fluorescence quantum yield Q . The natural lifetime τ_n of the excited state S_1 in the absence of non-radiative processes and quenching corresponds to the ratio between the observed lifetime τ and the quantum yield Q .

$$\tau_n = \frac{\tau}{Q} \quad (1.21)$$

1.5.2. Fluorescence Resonance Energy Transfer (FRET)

Another process depopulating the excited state of a fluorophore is *FRET* to a second chromophore. The basis for FRET is the assumption, that the electric field around a molecule in the excited singlet state S_1 can be approximated by the field generated by a classical harmonic electric oscillator. The oscillation of the first molecule (donor D) can induce oscillations and finally excitation of another molecule (acceptor A) under resonance conditions. The electrostatic interaction energy E between two electric dipoles μ_D and μ_A is

$$E \propto \frac{\mu_D \mu_A}{r^3} \quad (1.22)$$

with r the distance between the two dipoles. The energy transfer rate k_{ET} is proportional to the square of the interaction energy E and therefore proportional to $1/r^6$. Generally it is written in the following form:²⁴²

$$k_{ET} = \frac{1}{\tau_D} \left(\frac{R_0}{r} \right)^6 \quad (1.23)$$

with the so called Förster distance R_0 ,

$$R_0^6 = \frac{9000 \ln(10) \kappa^2 Q_D}{128 \pi^5 N_A n^4} \int_0^\infty F_D(\lambda) \varepsilon_A(\lambda) \lambda^4 d\lambda \quad (1.24)$$

The Förster distance R_0 depends on the spectral overlap between physical donor fluorescence spectrum F_D ²⁴³ and acceptor extinction spectrum ε_A , donor quantum yield Q_D , Avogadro's number N_A and the refractive index n of the medium separating donor and acceptor. κ^2 depicts a factor which accounts for the relative orientation of the dipoles of the two chromophores. For geometric reasons it is between 0 (perpendicular orientation) and 4 (head-to-tail parallel orientation). For dynamic averaging (fast rotation) of the two chromophores

1. Introduction

$\kappa^2 = 2/3$ and for random orientation it is 0.476²⁴⁴. Despite on-going discussions, $\kappa^2 = 2/3$ is used in most cases. This seems justified by recent findings, that the axial vibration of chromophores is sufficient to allow FRET even in a system where the dyes are oriented perpendicular to each other.²⁴⁵ Using equation (1.24), the Förster distance R_0 can be determined from the fluorescence and absorbance spectra of donor and acceptor dyes. Like the fluorescence lifetime τ and the quantum yield Q , R_0 can vary strongly with the experimental conditions like linker chemistry, solvent, pH and temperature.

Typical Förster distances lie between 10–60 Å. Accordingly, FRET is a powerful tool to determine distances on the length scale of protein dimensions and was often referred to serve as "spectroscopic ruler"^{246,247}. In the most simple case, donor and acceptor are kept at fixed distance r by a rigid linker, which only allows free rotation of the dyes. In this case, the distance r can be determined either through measurement of the fluorescence intensity, or measurements of the fluorescence lifetimes. The FRET-efficiency E , the relative amount of energy (in units of photons), transferred to the acceptor is defined as:

$$E = 1 - \frac{Q_{\text{DA}}}{Q_{\text{D}}} = 1 - \frac{F_{\text{DA}}}{F_{\text{D}}} \quad (1.25)$$

F_{DA} and F_{D} are the absolute fluorescence intensities of the donor in the presence and absence of the acceptor. Since intensity measurements depend on concentration, a more precise way is to determine the E from time resolved fluorescence measurements.

$$F = \int_0^{\infty} F_0 e^{-t/\tau} dt = F_0 \cdot \tau \quad (1.26)$$

The FRET-efficiency E is therefore related to the fluorescence lifetimes τ_{DA} and τ_{D} in presence and absence of the acceptor.

$$E = 1 - \frac{\tau_{\text{DA}}}{\tau_{\text{D}}} \quad (1.27)$$

The lifetimes τ_{DA} and τ_{D} can be determined from time-correlated single photon counting (TCSPC) measurements. τ_{DA} is the decay time of the excited state population $p^*(r, t)$ and depends on the rate of energy transfer k_{ET} (equation (1.23)):

$$\frac{dp^*(r, t)}{dt} = -\frac{1}{\tau_{\text{D}}} p^*(r, t) - k_{\text{ET}}(r) p^*(r, t) = -\frac{1}{\tau_{\text{DA}}} p^*(r, t) \quad (1.28)$$

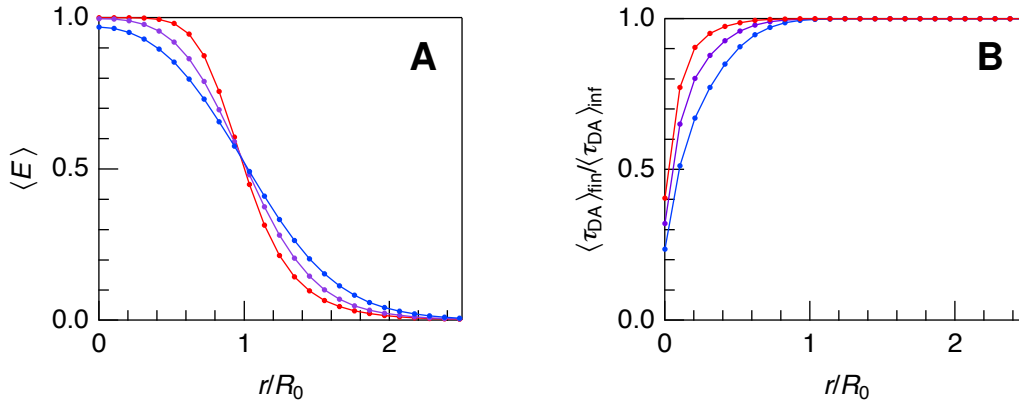


Figure 1.7.: A: Dependency of the FRET efficiency E on the distance r between the chromophores (\bullet). Dependence of the average FRET efficiency $\langle E \rangle$ on the midpoint r of a gaussian distribution of inter-dye distances with halfwidth $0.25 R_0$ (\bullet) and $0.4 R_0$ (\bullet). B: Difference between average lifetime $\langle \tau_{DA} \rangle_{fin}$ determined in an experiment with finite time resolution from the correct ensemble averaged lifetime $\langle \tau_{DA} \rangle_{inf}$ determined at infinite time resolution. The time resolutions are 1 % of the donor lifetime τ_D (\bullet), 2.5 % of τ_D (\bullet) and 5 % of τ_D (\bullet). The underlying distribution is gaussian with halfwidth $0.25 R_0$ and midpoint r .

Thus, E is related to r through

$$E = \frac{R_0^6}{R_0^6 + r^6} \quad (1.29)$$

In summary, if E and R_0 are determined from experiments, r can be calculated.

FRET is successfully used to determine distances between two chromophores in folded proteins which has been proved by comparison with known X-ray structures.^{248–251}

FRET becomes much more complex when unfolded proteins or peptides are studied. For unfolded proteins, not only one distance r between two points along the polypeptide chain carrying the chromophores is expected, but rather a distribution of distances $p_{eq}(r)$. The width and exact shape of the distance distribution influences the FRET-efficiency. The correct ensemble average plays a crucial role in such an experiment. Shorter distances lead to higher FRET-efficiencies and therefore contribute more to the ensemble averaged value $\langle E \rangle$, which is directly related to the ensemble averaged fluorescence intensity $\langle F_{DA} \rangle$:

$$\langle E \rangle = \int_{r_{min}}^{r_{max}} p_{eq}(r) \frac{R_0^6}{R_0^6 + r^6} dr = 1 - \frac{\langle F_{DA} \rangle}{F_D} \quad (1.30)$$

The inter-dye distance calculated from this is:

1. Introduction

$$r_{\langle E \rangle} = R_0 \left(\frac{1}{\langle E \rangle} - 1 \right)^{1/6} \quad (1.31)$$

In general, for broad distributions, $r_{\langle E \rangle}$ does not correspond to the correct ensemble averaged distance. Instead, the correct root mean square distance $\langle r^2 \rangle^{1/2}$ can be calculated in the following way:

$$\langle r^2 \rangle = \int_{r_{\min}}^{r_{\max}} p_{\text{eq}}(r) r^2 \mathrm{d}r \quad (1.32)$$

To obtain $\langle r^2 \rangle^{1/2}$, the shape of the distribution $p_{\text{eq}}(r)$ has to be known. In general, it is not possible to obtain $p_{\text{eq}}(r)$ from FRET efficiency, as the mathematical information content of the latter is too low. However, when making the assumption that $p_{\text{eq}}(r)$ obeys the random coil model (equation (1.6)) which is mathematically defined by only one parameter (compare Figure 1.2), $\langle r^2 \rangle^{1/2}$ of the Gaussian chain can be obtained.

The equations (1.30)–(1.32) are also crucial for interpreting FRET data of systems labeled with dyes with long linkers. These linkers possess conformational freedom and fast sampling of the conformational space given by the linkers superimposed to the properties of the protein.²⁵²

In case of a broad distance distribution, the time resolved fluorescence decay is not described by one single time constant, but by the continuous sum of different lifetimes $\tau_{\text{DA}}(r)$ weighted by the probability distribution $p_{\text{eq}}(r)$.

$$\langle F_{\text{DA}} \rangle(t) = \int_{r_{\min}}^{r_{\max}} p_{\text{eq}}(r) \cdot F_0 e^{-t/\tau_{\text{DA}}(r)} \mathrm{d}r \quad (1.33)$$

If intra-chain diffusion is completely negligible, this can serve as a fitting function for time resolved FRET data to determine $p_{\text{eq}}(r)$. In analogy to the more simple case of equation (1.26), the fluorescence intensity $\langle F_{\text{DA}} \rangle$ can be calculated from this by integrating over time.

$$\langle F_{\text{DA}} \rangle = \int_{r_{\min}}^{r_{\max}} \int_{t=0}^{\infty} p_{\text{eq}}(r) F_0 e^{-t/\tau_{\text{DA}}(r)} \mathrm{d}t \mathrm{d}r = F_0 \int_{r_{\min}}^{r_{\max}} p_{\text{eq}}(r) \tau_{\text{DA}}(r) \mathrm{d}r = F_0 \langle \tau_{\text{DA}} \rangle \quad (1.34)$$

In this way, the FRET efficiency $\langle E \rangle$ is related to the ensemble averaged lifetime $\langle \tau_{\text{DA}} \rangle$.

$$\langle E \rangle = 1 - \frac{\langle \tau_{\text{DA}} \rangle}{\tau_{\text{D}}} \quad (1.35)$$

A crucial step in the experimental determination of $\langle \tau_{\text{DA}} \rangle$ is the finite time resolution of TCSPC instruments. Thus, in reality, the lower boundary of the integration over time is $t = \tau_{\text{min}}$, which is the smallest time constant resolvable by the instruments. This removes the contribution of the smaller distances to the ensemble average, and thus the determined $\langle \tau_{\text{DA}} \rangle_{\text{fin}}$ for finite time resolution is smaller than the correct value $\langle \tau_{\text{DA}} \rangle_{\text{inf}}$ obtained for infinite time resolution as shown in Figure 1.7 B. A direct comparison between $\langle E \rangle$ obtained from intensity measurements and $\langle E \rangle$ obtained from lifetime measurements is in general only possible if the distribution $p_{\text{eq}}(r)$ is known, for example by applying equation (1.33) if intra-chain diffusion is negligible.

Unfolded proteins are in general not only characterized by a broad distance distribution $p_{\text{eq}}(r)$, but also by distinct intra-chain diffusion. These dynamics in the nanosecond range can have a strong influence on FRET as will be presented in section 3.1. In this sense, FRET measurements offer the possibility to determine the intra-chain diffusion coefficient and therefore to characterize the dynamics in unfolded polypeptides.

2. Aims of research

To conceive protein folding and function, an in-depth understanding of the intrinsic physical properties of polypeptide chains in all thermodynamically relevant states is indispensable. Among these states, the unfolded state is the conformationally most versatile and serves as the starting point of the folding process. Its characterization is of particular importance for the comprehensive understanding of folding, as dynamics in the unfolded state determine the speed of a conformational search for favorable interactions. The different conformations adopted by the unfolded polypeptide chain and the dynamical processes are a result of a number of different microscopic effects from favorable or unfavorable intra-chain and chain-solvent interactions activated by thermal energy. To identify the phenomenons, which limit the dynamics and determine the conformations in the unfolded state, is the aim of this study.

Determination of the end-to-end distance distribution and the intra-chain diffusion coefficient by time-resolved FRET measurements

The quantitative characterization of the conformational properties and the dynamics of unfolded proteins depends on the definition of measures for these properties. A conformational measure is the average end-to-end distance of an unfolded polypeptide or the distribution of end-to-end distances in the unfolded ensemble. Conformational rearrangements in the unfolded state can be described as intra-chain diffusion. It has been shown, that the end-to-end distance distribution and the intra-chain diffusion coefficient of an unfolded polypeptide can be simultaneously determined from time-resolved FRET measurements. However, this approach suffers from shortcomings mainly from a low accuracy of the determined parameters imposed by correlations between the different parameters. Therefore, the fundamental objective of this study is to advance this approach in order to determine the end-to-end distance distribution and the intra-chain diffusion coefficient of unfolded model peptides with high accuracy.

Do unfolded proteins behave like random coils?

Conformational properties of unfolded proteins have classically been described by the random coil model. However, experimental evidence for a more complex behavior of unfolded proteins containing significant amounts of secondary structure or native-like structural elements questions the validity of such assumptions.

Random coil behavior of unfolded proteins can be tested by determining the end-to-end distance distribution using time resolved FRET measurements of unfolded model peptides. Repeat peptides like for example (Gly-Ser)-peptides or native loop sequences can serve as models for the unfolded state. The distance distributions of these model peptides determined by time-resolved FRET measurements, allow for a critical comparison with the distance distributions of a random coil or an excluded volume chain. To evaluate how the intrinsic properties of the polypeptide chain determine the end-to-end distance distribution, the experimental results will be compared to unfolded peptide conformations obtained by Monte Carlo (MC) simulations using hard sphere models.

Does internal friction contribute to the dynamics in unfolded polypeptide chains?

In a macroscopical view, conformational changes in unfolded proteins are described by intra-chain diffusion in a continuous solvent limited by frictional forces. A major source of friction is the viscosity of the solvent, but other peptide intrinsic properties like steric effects or intra-chain interactions may be a source for additional internal friction.

The coefficients of friction can be directly calculated from the intra-chain diffusion coefficients determined by time resolved FRET measurements. The determination of the friction coefficients of unfolded model peptides at different solvent viscosities allows for a separation of internal friction from solvent friction through an extrapolation to zero solvent viscosity. Accordingly, the limits for chain dynamics in the presence and absence of solvent viscosity can be identified.

Are the conformations and dynamics of unfolded peptides affected by denaturants and stabilizing osmolytes?

Denaturants and stabilizing osmolytes are small co-solute molecules, that affect protein stability. Denaturants act by shifting the equilibrium from the native state N towards the

unfolded state U thus unfolding proteins. In contrast, stabilizing osmolytes shift the equilibrium towards N, thus stabilizing the native state or even forcing unfolded proteins to fold. Determination of the end-to-end distance distribution and the intra-chain diffusion coefficient of unfolded model peptides in presence of different concentrations of denaturants or stabilizing osmolytes gives information on the effect of these co-solutes on the unfolded state. Comparison of the resulting conformational changes with the properties of unfolded model ensembles derived from MC simulations based on hard sphere models can contribute to a better understanding of the principles underlying co-solute induced stability changes.

What are the properties of the free energy landscape of the unfolded state?

Many different conformations of the polypeptide chain coexist in the unfolded state ensemble. These conformations interconvert majorly unhindered at ambient temperature. The experimentally observed single exponential kinetics for intra-chain loop formation show, that these conformations are energetically similar and that barriers in free energy separating these different conformations are relatively small compared to thermal energy. Intra-chain diffusion is thus a process taking place in a rough energy landscape.

The roughness of the energy landscape can be probed by determining the temperature dependence of the intra-chain diffusion coefficient which yields the barrier heights. Additionally, the determination of the end-to-end distance distribution allows to test for temperature-induced conformational changes of the polypeptide chain.

3. Results

3.1. End-to-end distance distribution and intra-chain diffusion coefficient determined by FRET

3.1.1. FRET in unfolded peptide systems

As stated in section 1.5.2, FRET in unfolded peptides is influenced both by the broad equilibrium distribution of distances $p_{\text{eq}}(r)$ between the chromophores as well as by intra-chain diffusion caused by Brownian motion within the unfolded system. The influence of $p_{\text{eq}}(r)$ on time-resolved FRET is given by the solution of the differential equation (1.28). This equation can be extended by a diffusional term of the Fokker-Planck type (equation (1.16)). This leads to a partial differential equation (PDE), describing the time course of the relative population $\bar{p}^*(r, t)$ of excited donor molecules.¹⁷⁰

$$\frac{\partial \bar{p}^*(r, t)}{\partial t} = -\frac{1}{\tau} \cdot \bar{p}^*(r, t) - \frac{1}{\tau} \left(\frac{R_0}{r} \right)^6 \cdot \bar{p}^*(r, t) + \frac{1}{p_{\text{eq}}(r)} \frac{\partial}{\partial r} \left(p_{\text{eq}}(r) D \frac{\partial \bar{p}^*(r, t)}{\partial r} \right) \quad (3.1)$$

In this equation, r denotes the distance between donor- and acceptor-chromophore, τ is the lifetime of the donor in absence of an acceptor, R_0 is the Förster distance of the FRET-pair defined by equation (1.24) and D is the intra-chain diffusion coefficient. The relative population $\bar{p}^*(r, t)$ is defined as follows:

$$\bar{p}^*(r, t) = \frac{p^*(r, t)}{p_{\text{eq}}(r)}. \quad (3.2)$$

The first term on the right hand side of equation (3.1) accounts for radiationless relaxation of the excited chromophore to the ground state, the second term describes the distance dependent energy transfer to the acceptor and the third term accounts for reestablishment of the equilibrium distance distribution through intra-chain diffusion.

The time-resolved fluorescence decay of the donor can be calculated from the solution

3. Results

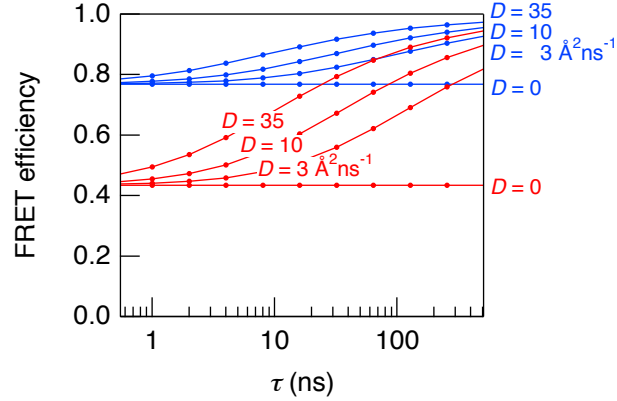


Figure 3.1.: The influence of intra-chain diffusion on FRET efficiencies increases with increasing donor lifetime τ for a FRET pair with Förster distance $R_0 = 20 \text{ \AA}$ (—) and a FRET pair with $R_0 = 30 \text{ \AA}$ (—).

$\bar{p}^*(r, t)$ of the PDE (equation (3.1)).

$$F(t) = \int_{r_{\min}}^{r_{\max}} \bar{p}^*(r, t) \cdot p_{\text{eq}}(r) dr \quad (3.3)$$

No analytical solution for the PDE (equation (3.1)) is known, so it has to be solved by numeric approximation of the solution.

Several important implications are based on this PDE. The FRET efficiency $\langle E \rangle$ (equation (1.25)) as well as the time-resolved fluorescence $F(t)$ can be significantly influenced by intra-chain diffusion. This influence is particularly strong for donor chromophores with a long lifetime τ as shown in Figure 3.1. Additionally taking into count the influence of a broad distance distribution as shown in section 1.5.2 Figure 1.7, the difficulties in directly interpreting FRET efficiencies for unfolded peptides becomes obvious. For instance, intra-chain diffusion always increases $\langle E \rangle$, while the width of the distance distribution can either decrease or increase $\langle E \rangle$ depending on the relative position of the distribution compared to the Förster distance R_0 . As $\langle E \rangle$ also depends on the position of the center of mass of the distribution, direct interpretation of FRET efficiency with respect to intra-chain diffusion is not straightforward. Instead, the time-resolved fluorescence decay will be considered.

The influence of both, the broad distance distribution and intra-chain diffusion on time-resolved fluorescence is depicted in Figure 3.2. The broad distance distribution leads to

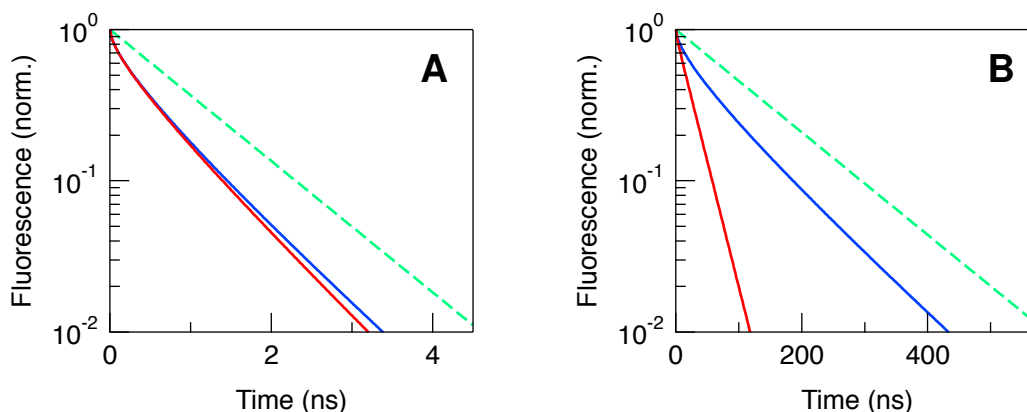


Figure 3.2.: Time-resolved fluorescence decay calculated from equations (3.1) and (3.3) with $\tau = 1$ ns (A) and $\tau = 128$ ns (B). Donor only decay ($-\cdot-$), donor fluorescence decay in presence of the acceptor ($R_0 = 20$ Å) for a broad distance distribution with $b = 12.5$ Å and $\sigma = 13.5$ Å calculated in the absence of intra-chain diffusion ($-$) and in the presence of intra-chain diffusion with $D = 10$ Å²ns⁻¹ ($-$). The influence of intra-chain diffusion on the signal is much stronger for the donor with long lifetime (B).

deviations from a single exponential fluorescence decay. The influence of intra-chain diffusion is weak for donor chromophores with relatively short lifetime of $\tau = 1$ ns, while the effect is very pronounced for donors with long lifetime of $\tau = 128$ ns (compare red and blue traces in Figure 3.2 A and B). The fluorescence decay is significantly faster for pronounced intra-chain diffusion than for negligible intra-chain diffusion. Moreover, the fluorescence decay of donors with long lifetime τ can be well approximated by a single exponential decay when intra-chain diffusion is pronounced.

3.1.2. Analyzing FRET experiments under consideration of intra-chain diffusion

The equations (3.1) and (3.3) can be used to numerically calculate the time-resolved fluorescence decay for a specific intra-chain diffusion coefficient D and a specific model for $p_{\text{eq}}(r)$. Different possibilities for $p_{\text{eq}}(r)$ were presented in section 1.3.1 and 1.3.3. The models differ in the shape of the distribution and in the number of free parameters mathematically defining the shape. The *random coil* or *Gaussian chain model* is defined by only one free parameter, the mean square distance $\langle r^2 \rangle$, while two free parameters are needed to determine the *worm like chain* and the *chain with excluded volume*. In view of the complex conformational properties of unfolded proteins discussed in section 1.3.2, it is certainly ap-

3. Results

appropriate to use one of the mathematically more complex models in order to obtain a less biased modeling of unfolded peptides. In the following, the focus will therefore be on the excluded volume model (see section 1.3.3).

Consequently, the solution of the PDE depends on three free parameters, the intra-chain diffusion coefficient D , as well as the parameters b and σ from the model for $p_{\text{eq}}(r)$ in equation (1.12). The time-resolved fluorescence of an unfolded peptide labeled with two different chromophores at specific positions can be experimentally measured by TCSPC. The donor lifetime τ and the Förster distance R_0 are characteristic for each FRET pair and can be determined in separated experiments. The experimental data can be analyzed by a fit with the function $F_{D,b,\sigma}(t)$ (equation (3.3)). In the fitting procedure, the average squared difference χ_R^2 between the i th datapoint $F_{\text{exp}}(t_i)$ and the function $F_{D,b,\sigma}(t_i)$ is minimized by varying the free parameters.

$$\chi_R^2 = \frac{1}{n-p} \sum_{i=1}^n \frac{(F_{\text{exp}}(t_i) - F_{D,b,\sigma}(t_i))^2}{(\Delta F_{\text{exp}}(t_i))^2} \quad (3.4)$$

Thus, the hyper surface $\chi_R^2(D, b, \sigma)$ in 3-dimensional space spanned by the 3 fitting parameters is searched for a minimum. In the ideal case, this hyper surface has one well defined global minimum at $(D_{\text{fit}}, b_{\text{fit}}, \sigma_{\text{fit}})$ determined from the converging optimization in the fitting process. If $F_{D,b,\sigma}^{\text{fit}}(t)$ gives a satisfying description of the experimental data, i.e., if χ_R^2 is close to 1 and there are no significant residuals, the solution is assumed to be close to the real values ($D_{\text{fit}} \approx D_{\text{real}}, b_{\text{fit}} \approx b_{\text{real}}, \sigma_{\text{fit}} \approx \sigma_{\text{real}}$) and to give a good approximation of reality.

3.1.3. Experimental design

In order to determine the intra-chain diffusion coefficient, D , and the end-to-end distance distribution $p_{\text{eq}}(r)$ of an unfolded polypeptide, an appropriate pair of chromophores (FRET pair) needs to be found. The FRET pair has to fulfill several requirements. Both chromophores should be compatible with peptide chemistry so that they can be attached to the peptide by linkers small in comparison with the peptide chain studied. Neither the linker, nor the peptide should interfere with the chromophores significantly changing their photo-physical properties. The disturbance of the unfolded peptide by the chromophores should be kept small, requiring the chromophores to be similar in size to side chains of natively occurring amino acids. The lifetime τ of the selected donor should be long enough to allow for an influence of intra-chain diffusion on the measured fluorescence signal. The Förster distance R_0 of the FRET pair should be in the range of the root means square distance between

the chromophores in the unfolded peptide to guarantee FRET to be most sensitive to differences in distances. The Förster distance R_0 is mostly determined by the spectral position of the overlap between donor emission and acceptor absorbance due to the λ^4 dependence in equation (1.24). Thus, values for R_0 typically increase with the emission wavelength of the donor. The absorption wavelength of an aromatic molecule is related to the size of the π -electron system. Thus the absorbance of small chromophores is in the ultraviolet (UV) range and corresponding Förster distances are in the range between 10 - 30 Å. This limits the length of the poly-peptide used in this study as the root mean square end-to-end distances of a polypeptide increases with chain length. Three different FRET pairs were characterized and tested for their applicability. As a model peptide, a glycine-serine sequence of 8 repeats, (GS)₈, was chosen. In Table 3.1 the FRET pairs consisting of 7-amino-4-methyl-3-coumarin (AMCA)¹ and 2,4-dinitrophenyl (DNP), naphthalene (Naph) and dansyl (Dans), as well as of pyrene (Pyr) and Dans are presented together with their chemical structure and their photo-physical characteristics when attached to a (GS)₈ peptide.

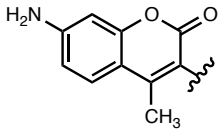
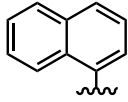
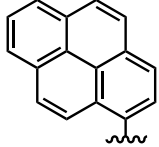
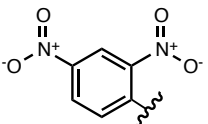
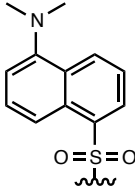
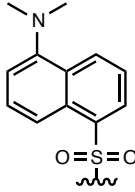
3.1.4. Determination of the intra-chain diffusion coefficient by global analysis of different FRET experiments

The time-resolved fluorescence of a (GS)₈ peptide labeled with the Pyr/Dans FRET pair at or near the termini dissolved in water was recorded by TCSPC. A fit of $F_{D,b,\sigma}(t)$ to the data resulted in $\chi_R^2(D_{\text{fit}}, b_{\text{fit}}, \sigma_{\text{fit}}) = 1.13$ and random distribution of the residuals indicating that the data is well described by the fitting function. In order to determine the errors of the obtained fitting parameters $D_{\text{fit}}, b_{\text{fit}}$ and σ_{fit} and to show, that the fit converged to a well defined minimum in the $\chi_R^2(D, b, \sigma)$ hyper surface, projections of this hyper surface on each fitting parameter were calculated. In case of the fitting parameter D for example, the projection is a function $\chi_R^2(D, b_{\text{fit}}(D), \sigma_{\text{fit}}(D))$. Every point in this function is calculated by varying b and σ while keeping D fixed. This method is often referred to as *support plane analysis*. The calculated projection for D together with the statistical confidence level is shown in Figure 3.3 A, the ones for the other two parameters in the Figures 3.3 B and C. In the case of Pyr, there is hardly any defined minimum observable for D and it can't be determined whether the parameter b has a finite value or tends towards zero. It is therefore not possible to determine the intra-chain diffusion coefficient D or the other parameters from this experiment,

¹In this study, this abbreviation is used for aminomethylcoumarin acetic acid linked to a polypeptide chain under reaction of the acid group.

3. Results

Table 3.1.: Overview of FRET pairs, the lifetimes τ of the donors and the Förster distances R_0 when attached to a (GS)₈ peptide dissolved in 10 mM KPh buffer at pH 7 and 22.5°C.

Donor			
	7-amino-4-methyl-3-coumarin (AMCA)	naphthalene (Naph)	pyrene (Pyr)
τ	5 ns	38 ns	150 ns
Acceptor			
	2,4-dinitrophenyl (DNP)	dansyl (Dans)	dansyl (Dans)
R_0	31 Å	19 Å	22 Å

even though D has a strong influence on the time-resolved fluorescence of the long lifetime donor Pyr.

Repeating the experiment and the fit with a (GS)₈ peptide labeled with the FRET pair AMCA/DNP, which comprises a rather short lifetime donor, gives a $\chi_R^2(D_{\text{fit}}, b_{\text{fit}}, \sigma_{\text{fit}}) = 0.95$ and randomly distributed residuals. Calculating the projection $\chi_R^2(D, b_{\text{fit}}(D), \sigma_{\text{fit}}(D))$ a minimum is indeed observed defining lower and upper limits for D_{fit} as shown in Figure 3.3 A. In the same sense, limits can be defined for the parameter σ as shown in Figure 3.3 C. In contrast, only an upper limit can be fixed for the parameter b . The better definition of the parameters in this fit compared to the fit of the Pyr/Dans peptide probably stems from the higher mathematical information content of the non-exponential fluorescence decay in this case. Despite these minor improvements, the minima are still flat and expand over a rather wide range of possible values so the definition of the parameters is still weak and a correlation between the parameters remains.

To remove this correlation between fitting parameters, data with a higher mathematical information content is required. Accordingly, the two datasets from the AMCA/DNP labeled

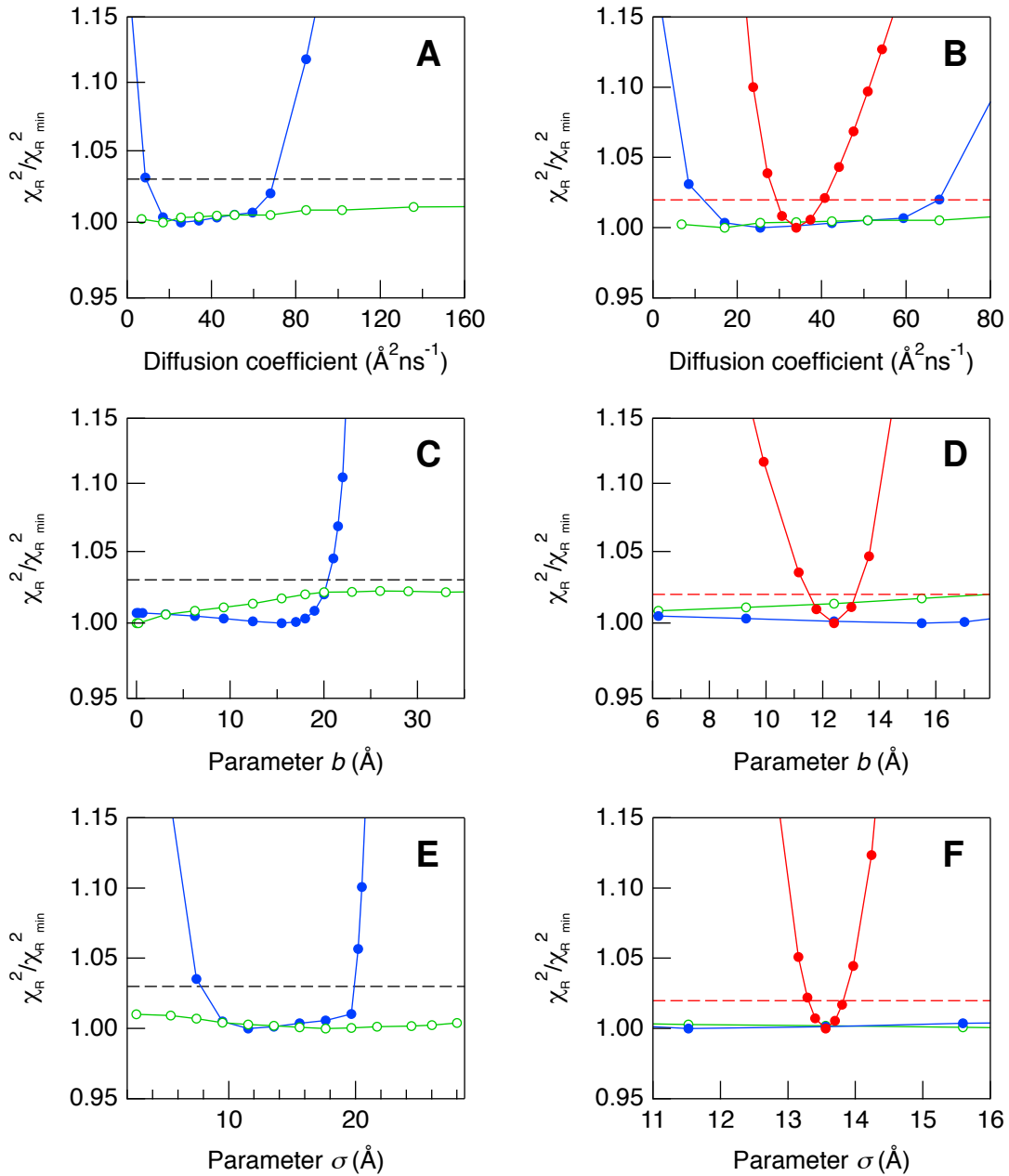


Figure 3.3.: A, C, E: Support plane analysis for the fits of $F_{D,b,\sigma}(t)$ (equation (3.3)) to the TCSPC data of one FRET pair. For the Pyr/Dans labeled (GS)₈ peptide (\circ), the calculated projections do not cross the 67 % confidence level within the range of the graphs indicating a very large error of the fitting parameters. For the AMCA/DNP labeled peptide (\bullet) limits for the fit parameters can be defined. B, D, F: Support plane analysis for the global fit of data of the Pyr/Dans and the AMCA/DNP FRET pair (\bullet). The calculated projections on the respective global fit parameters have parabolic shape and cross the 67 % confidence level indicating finite error ranges of the fit parameters.

3. Results

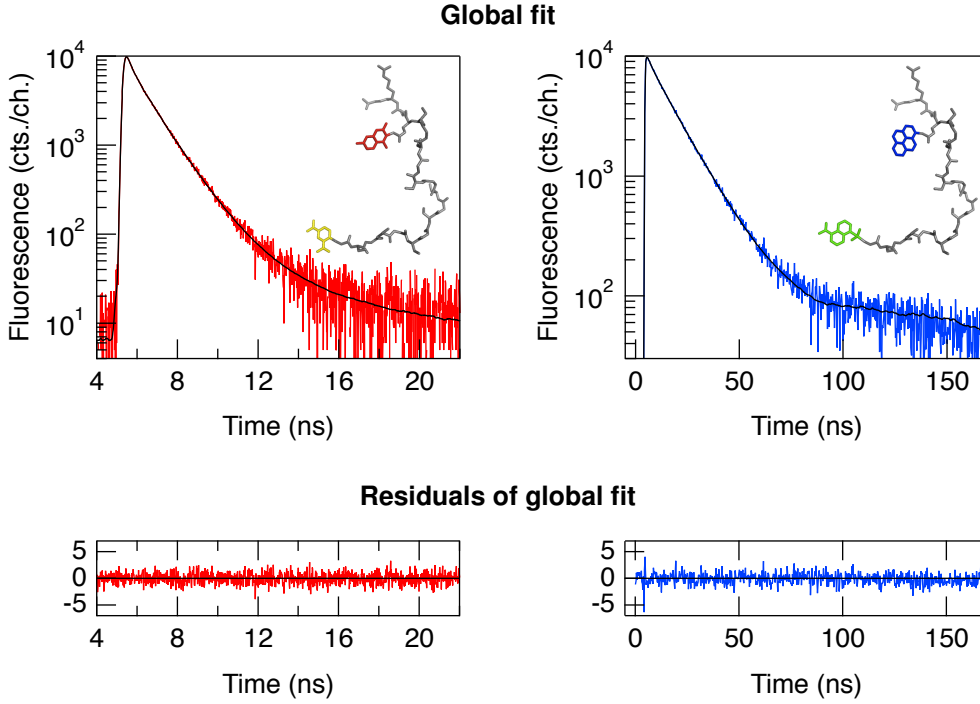


Figure 3.4.: The raw data of a $(GS)_8$ peptide labeled with the AMCA/DNP FRET pair (—) and labeled with the Pyr/Dans FRET pair (—). The global fit is shown (—), the residuals are of the global fit. The measurements were carried out in 10 mM potassium phosphate buffer at pH 7.0 and 22.5°C.

and from the Pyr/Dans labeled $(GS)_8$ are combined in a global analysis. To achieve this, the global hyper surface $\chi_{R, \text{global}}^2(D, b, \sigma)$ is searched for a minimum.

$$\chi_{R, \text{global}}^2(D, b, \sigma) = \frac{1}{2} (\chi_{R, \text{amc}}^2(D, b, \sigma) + \chi_{R, \text{pyr}}^2(D, b, \sigma)) \quad (3.5)$$

The global fit results in $\chi_{R, \text{global}}^2(D_{\text{fit}}, b_{\text{fit}}, \sigma_{\text{fit}}) = 1.04$. The fitted experimental data and the corresponding residuals of the global fit are shown in Figure 3.4. The nearly parabolic shape of the calculated projections of $\chi_{R, \text{global}}^2$ in the support plane analysis in Figure 3.3 B, D and F proofs, that the minimum $\chi_{R, \text{global}}^2(D_{\text{fit}}, b_{\text{fit}}, \sigma_{\text{fit}})$ is well defined and narrow. The resulting parameters $D_{\text{fit}}, b_{\text{fit}}$ and σ_{fit} are summarized in Table 3.3 together with their 67 % confidence intervals.

In conclusion, the intra-chain diffusion coefficient and the end-to-end distance distribution of a $(GS)_8$ peptide can be reliably determined from time-resolved FRET measurements by

Table 3.2.: Results from global analysis of time-resolved FRET data and the corresponding parameters for the chain with excluded volume model.

	results	confidence intervals		relative
	global fit	lower	upper	error
D ($\text{\AA}^2\text{ns}^{-1}$)	33.9	-4.7	+6.9	17 %
b (\AA)	12.4	-0.9	+0.8	7 %
σ (\AA)	13.6	-0.3	+0.2	2 %

global analysis of data from the two different FRET pairs AMCA/DNP and Pyr/Dans.

3.1.5. Generalization of global analysis method

In case of the $(\text{GS})_8$ peptide in water, the intra-chain diffusion coefficient and the end-to-end distance distribution $p_{\text{eq}}(r)$ could be successfully determined by global analysis of time-resolved FRET data of two different FRET pairs. However, this does not prove, that this method will be successful under different solvent conditions or for other peptides with unknown and maybe completely different values for D , b and σ . To test the applicability of the global analysis method to the general case, the theoretical time-resolved fluorescence $F_{D,b,\sigma}(t)$ was characterized for a wide range of parameters. At first, the theoretical fluorescence $F_{D_0,b_0,\sigma_0}(t)$ is calculated for a certain triple of parameters (D_0, b_0, σ_0) for one FRET pair. $F_{D_0,b_0,\sigma_0}(t)$ is compared the theoretical fluorescence signal $F_{D_i,b_i,\sigma_i}(t)$ of another triple (D_i, b_i, σ_i) . The two signals $F_{D_0,b_0,\sigma_0}(t)$ and $F_{D_i,b_i,\sigma_i}(t)$ are considered to be undistinguishable if their (time) averaged difference is smaller than the typical error in a time resolved FRET experiment. This error can be estimated based on the known uncertainty of the data points in a TCSPC experiment caused by the signal to noise ratio, the total number of data points, the time resolution and the dark counts. By testing a large enough number of triples (D_i, b_i, σ_i) for their indistinguishability from the first triple (D_0, b_0, σ_0) , it can be estimated, how large the uncertainty of the first triple (D_0, b_0, σ_0) would be in the fit.

Accordingly, calculations of $F_{D_i,b_i,\sigma_i}(t)$ were performed for 15,510 triples for D_i between $2.4 \text{ \AA}^2\text{ns}^{-1}$ and $50 \text{ \AA}^2\text{ns}^{-1}$, b_i between 6.8 \AA and 34.5 \AA and σ_i between 10 \AA and 20 \AA for all three FRET pairs displayed in Table 3.1. Physically meaningless combinations of b and σ were discarded. The results can be visualized in 3-dimensional graphs as shown in Figure 3.5 and Figure 3.6. In these figures, a triple (D_0, b_0, σ_0) is marked by a filled circle in the

3. Results

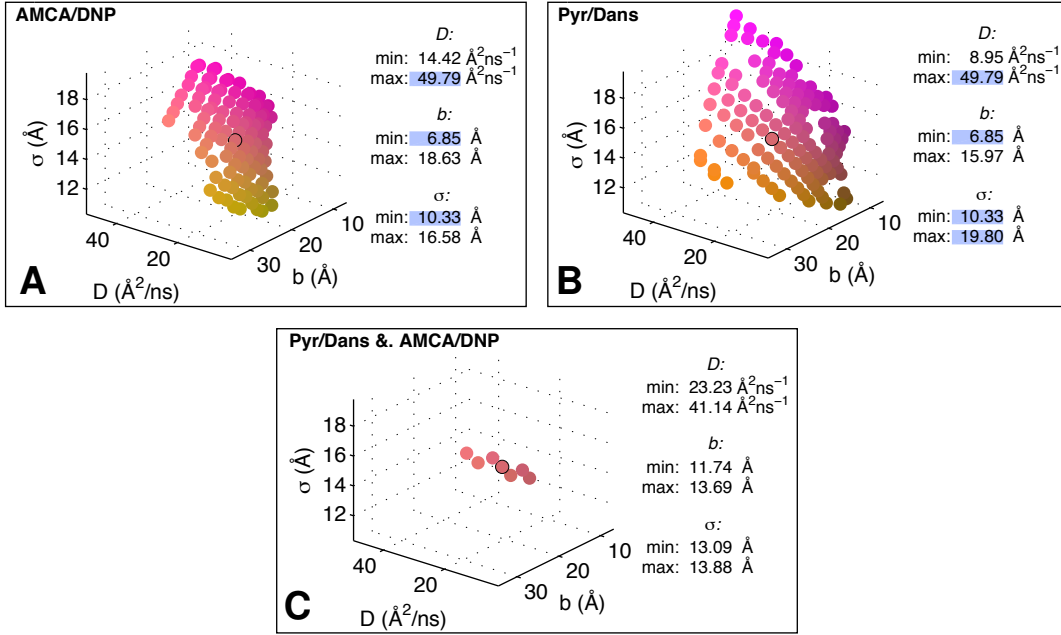


Figure 3.5.: Sensitivity of the calculated fluorescence signal $F_{D_0, b_0, \sigma_0}(t)$ towards changes in D , b and σ . The graphs shown are based on $D_0 = 30.91 \text{\AA}^2\text{ns}^{-1}$, $b_0 = 12.68 \text{\AA}$, $\sigma_0 = 13.48 \text{\AA}$ as marked (\circ). All parameter sets (D_i, b_i, σ_i) that would produce a $F_{D_i, b_i, \sigma_i}(t)$ which is within the typical experimental uncertainty of $F_{D_0, b_0, \sigma_0}(t)$ are marked by colored circles. The colors scale with the values of the parameters to highlight the depth effect, the size of the circles has an arbitrary value. Beneath the graphs, the lower and upper limits of the parameters are shown. The numbers are marked in blue, if the value corresponds to the limits of tested values for the parameter indicating that the actual error of the parameter is probably larger. A and B correspond to fits of data of one FRET pair, C corresponds to the global analysis.

3-dimensional space spanned by the positive D -, b - and σ -axis. All triples (D_i, b_i, σ_i) having a fluorescence signal $F_{D_i, b_i, \sigma_i}(t)$ within error of $F_{D_0, b_0, \sigma_0}(t)$ are equally marked in color by filled circles. The colored regions then have the shape of bent planes of finite thickness, which stretch out over the whole tested parameter range (see Figures 3.5 A and B). This indicates, that the error of the fit parameters is larger than the probed parameter range due to a strong correlation of the fit parameters.

To simulate a global fit, both theoretical fluorescence signals $F_{D_0, b_0, \sigma_0}^{\text{AMCA/DNP}}(t)$ and $F_{D_0, b_0, \sigma_0}^{\text{Pyr/Dans}}(t)$ as well as their typical experimental errors should be considered. Thus, for the global fit, (D_i, b_i, σ_i) is only within error range of (D_0, b_0, σ_0) if both theoretical fluorescence signals $F_{D_i, b_i, \sigma_i}^{\text{AMCA/DNP}}(t)$ and $F_{D_i, b_i, \sigma_i}^{\text{Pyr/Dans}}(t)$ are within the typical experimental error of $F_{D_0, b_0, \sigma_0}^{\text{AMCA/DNP}}(t)$ and $F_{D_0, b_0, \sigma_0}^{\text{Pyr/Dans}}(t)$. In the 3-dimensional graphs, this corresponds to the cross-section of the two

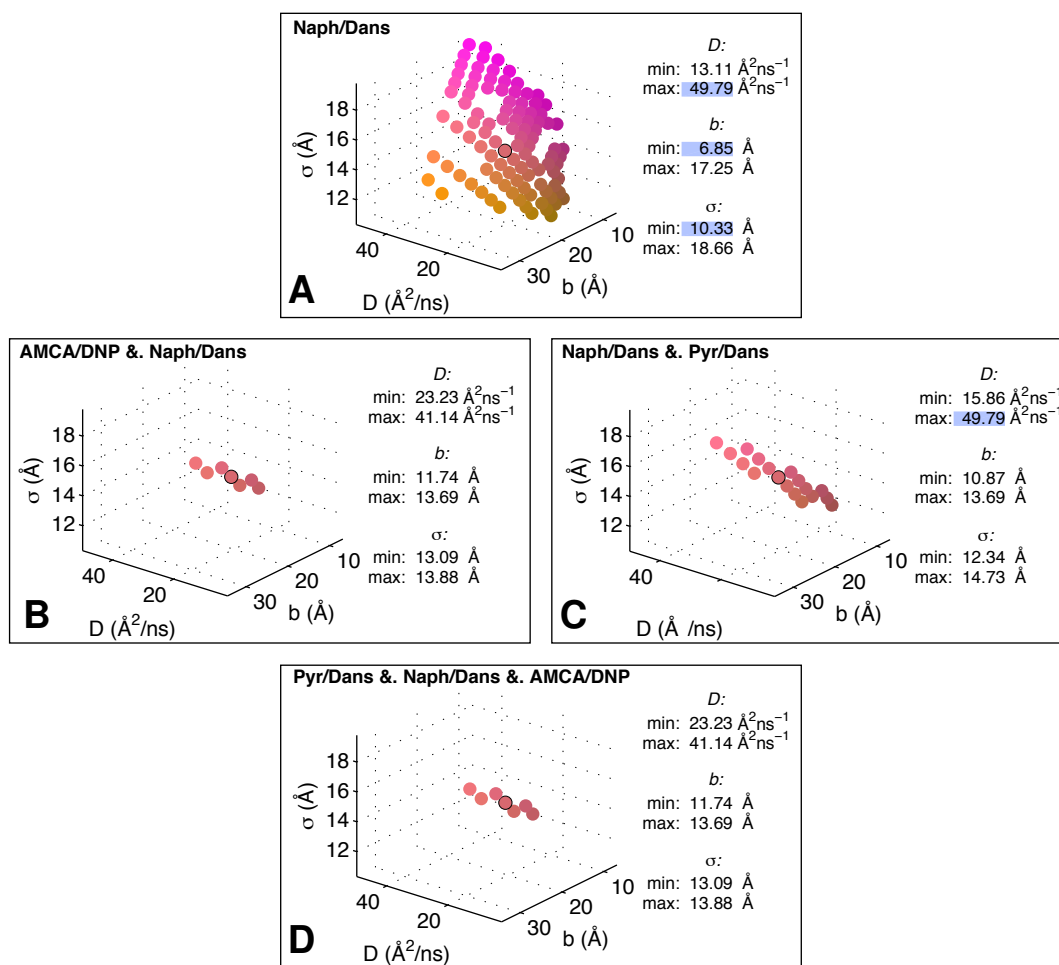


Figure 3.6.: Sensitivity of the calculated fluorescence towards changes in the parameters (see Figure 3.5 for details). A: fit of data of the Naph/Dans FRET pair. B, C: global analysis of the Naph/Dans data with one of the other FRET pairs. D: global analysis of all three FRET pairs.

colored regions in the graphs for AMCA/DNP and Pyr/Dans as shown in Figure 3.5 C.

By help of this procedure very different parameter sets (D_0 , b_0 , σ_0) can be tested for their correlation in the fit and especially for the possibility of removing the correlation by global analysis. This gives the possibility to design experiments for example for high solvent viscosities with weak intra-chain diffusion or high temperatures with strong intra-chain diffusion. Furthermore, this procedure allows to compare error bars for arbitrary parameters D_0 , b_0 and σ_0 in normal and global fits and for different FRET pairs.

Consequently, it is also possible to test another FRET pair, i.e., Naph/Dans, for the applicability to determine D and $p_{\text{eq}}(r)$. A fit of only the Naph/Dans data alone would give

3. Results

very similar results as the single fits of Pyr/Dans or AMCA/DNP as shown in Figure 3.6 A. The Naph/Dans pair was combined in a global analysis with AMCA/DNP (Figure 3.5 B) and with Pyr/Dans (Figure 3.5 C). The error of the fit parameters in global analysis of Naph/Dans and Pyr/Dans data would be quite large, while the error in global analysis of AMCA/DNP and Naph/Dans data is similar to the error for Pyr/Dans and AMCA/DNP. It was further tested if a global analysis of all three FRET pairs could narrow the error range of the fit parameters, but for most parameter sets (D_0, b_0, σ_0) no improvement was observed. Therefore, the supplementary expense of including another FRET pair in the experiments doesn't seem to be necessary.

3.1.6. Summary and discussions

The analysis showed, that the combination of time-resolved FRET measurements involving donors with long and short lifetimes with global data analysis is well-suited to determine the end-to-end distance distribution $p_{\text{eq}}(r)$ and the intra-chain diffusion coefficient D of unfolded polypeptides. This approach allows the accurate determination of all parameters with small errors as shown by the support plane analysis. While imprecise values for the parameters could be obtained using a single short lifetime donor, single long lifetimes are not applicable to define the parameters. Theoretical studies including up to 3 donor lifetimes showed, that combination of AMCA ($\tau \approx 5$ ns) and Pyr ($\tau \approx 150$ ns) is the best choice for investigating unfolded peptides. No significant improvement of precision of the parameters can be obtained by including Naph ($\tau \approx 40$ ns) as a third donor, or by combining Naph with one of the other two donors.

The global analysis approach is a significant improvement over the use of one single FRET pair as has been used in investigations of the proteins α -synuclein^{173,174}, p53¹⁷⁵ and RNase S-peptide²⁰⁰. The major drawback of using a single FRET pair is the large error originating in the strong correlation of the fit parameters.

Global analysis of FRET measurements with two different donor lifetimes allows to monitor changes of $p_{\text{eq}}(r)$ and D induced by changing solvent viscosity or temperature as all fit parameters can be determined independently. This is a drastic improvement over methods trying to reduce the correlation between the fit parameters by global analysis of data of one single FRET pair at varying viscosity¹⁷⁰ or temperature²⁰¹. In both these cases it has been assumed that $p_{\text{eq}}(r)$ is not influenced by the change in external conditions and therefore the diffusion coefficient for each condition (viscosity or temperature) could be determined.

However, if $p_{\text{eq}}(r)$ does indeed change, the resulting intra-chain diffusion coefficients are biased.

The choice of the chromophores in the global analysis critically determines the parameter range within which D can be accurately determined. The determination of D under conditions slowing down diffusional processes like high solvent viscosity or low temperature is most accurate if a long lifetime donor is combined with a short lifetime donor as shown by the theoretical studies. These studies further indicate, that this combination of long and short lifetime donors is an improvement over the use of two donors with long and intermediate lifetimes in the global analysis as has been used before¹⁷².

Global analysis methods involving two FRET pairs (either with different acceptors or different donors) have been applied to determine $p_{\text{eq}}(r)$ and D of alkyl chains.^{197–199} For alkyl chains, the inclusion of a long lifetime donor was not necessary due to the intrinsically strong intra-chain diffusion (around $110 \text{ \AA}^2\text{ns}^{-1}$).¹⁹⁸ As the intra-chain diffusion in peptides in aqueous solution with $D = 33.9 \text{ \AA}^2\text{ns}^{-1}$ is smaller by a factor of three, the use of a long lifetime donor is more important if the diffusion coefficient should be determined with high accuracy.

Intra-chain diffusion not only increases FRET in time-resolved or steady state ensemble based measurements, but also when FRET within a single molecule is studied. Even though the influence of diffusion on the FRET efficiency decreases with decreasing donor lifetime, it is not negligible at donor lifetimes of around 4 ns widely used in single molecule experiments as can be seen from Figure 3.1. The more intra-chain diffusion contributes to FRET, the shorter the intra-chain distances appear when intra-chain diffusion is discarded and consequently differ significantly from the real distances.

It has also been proposed to remove the correlation of parameters determined from an experiment of one single FRET pair by adding a quencher to reduce the donor lifetime τ and globally analyzing the data in presence and in absence of the quencher.²⁵³ This method has recently been questioned by the argument, that the relevant lifetime in the PDE (equation (3.1)) is the natural lifetime $\tau_n = \tau/Q$ (equation (1.21)) as the quantum yield Q enters the PDE through the Förster distance, $R_0^6 \propto Q$.²⁰² When adding a quencher, the natural lifetime should stay indeed unaffected. However, the approach of using different donor lifetimes as used here is based on a real difference between the respective donor lifetimes τ and also between the related natural lifetimes τ_n of the chromophores in use, even though this approach has also been questioned²⁰². The fluorescence decays in presence of the acceptor (see Figure 3.4) occur on very different absolute timescales showing average decay times of

3. Results

$\tau_{\text{av.}} = 0.8$ ns and 13 ns for the (GS)₈ peptide with AMCA/DNP and Pyr/Dans respectively. These aspects, together with the careful analysis of the support planes and the theoretical studies show, that with global analysis of short and long lifetime FRET measurements $p_{\text{eq}}(r)$ and D can be accurately determined.

3.2. Unfolded model peptides

The unfolded ensemble of a protein consists of many different conformations. Each conformer in the ensemble is characterized by the dihedral angles ϕ , ψ and ω of the residues as well as by overall properties of the polypeptide chain like the end-to-end distance r or the radius of gyration r_G . Consequently, the whole unfolded ensemble is characterized by the distribution of these parameters.

In the following, two model peptides, which are unfolded under physiological conditions, will be used to study the properties of the unfolded state. Firstly, the end-to-end distance distribution and the dynamics acting on end-to-end distance are determined by global analysis of time resolved FRET experiments. This approach allows a critical comparison between unfolded peptides and theoretical polymer models, i.e. the random coil model, the worm-like chain model and the chain with excluded volume model that were presented in section 1.3. Additionally, MC simulations of unfolded peptide conformations were carried out elucidating that the end-to-end distance distribution is imposed by the distribution of dihedral angles.

3.2.1. Glycine-Serine repeat peptides

Glycine-serine repeat peptides ($(GS)_x$ peptides) are widely used as models for the unfolded state and as model systems for studies of dynamics in the unfolded state.^{182,219,221,225,227,254} Due to the repeat sequence, they lack secondary structure¹⁷² and can serve as models for the unfolded state under physiological conditions. The high content of glycine results in a high flexibility of the peptide, as the lack of a side chain for glycine allows for more conformational freedom of the ϕ - and ψ -angles. No salt-bridges or hydrophobic interactions can be formed within a $(GS)_x$ peptides as no charged or non-polar side-chains are present. However, hydrogen bonds (H-bonds) can be formed within the polypeptide chain, or from the chain to solvent molecules. Thus, $(GS)_x$ peptides are good model systems to test for H-bonds in the unfolded state.

Distinct intra-chain diffusion is expected for $(GS)_x$ due its high flexibility. For instance, loop formation is faster in $(GS)_x$ peptides than in poly-serine peptides or native sequences of same length.^{221,224} For small chains, the loop formation rate constant of $(GS)_x$ peptides reaches a limiting value, while it scales with the number of residues for longer chains, $k_c \propto n^{-1.7}$.²²⁴ The (GS) -peptide used in the present study consists of 16 residues between

3. Results

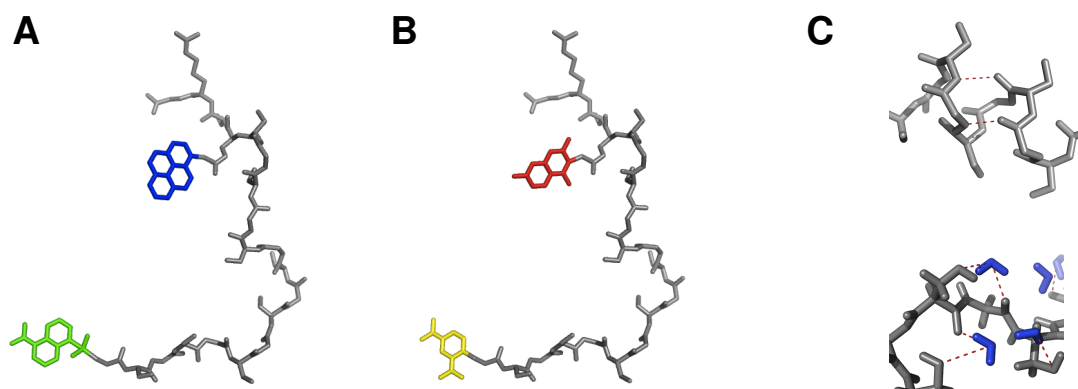


Figure 3.7.: (GS)₈ repeat peptides as used in the FRET measurements. The labels are pyrene (blue) and dansyl (green) (A) and aminomethylcoumarin (red) and dinitrophenyl (yellow) (B). Examples for intra-chain hydrogen bonds (C upper part) and hydrogen bonds to water molecules (C lower part).

the FRET labels (see Figure 3.7) and consequently the scaling law holds. Therefore, a (GS)₈ peptide is an interesting system to test whether the random coil model is applicable.

3.2.2. EF-loop from Parvalbumin

In order to compare the properties of the (GS)-repeat peptide to the properties of naturally occurring peptide sequences, a fragment from β -carp parvalbumin, a calcium binding protein, was chosen. The native structure of parvalbumin in the presence of calcium consists of several α -helices connected by loop regions. The peptide fragment PV(EF) used here (see Figure 3.8), links the two helices of the EF-hand motive responsible for binding one of the

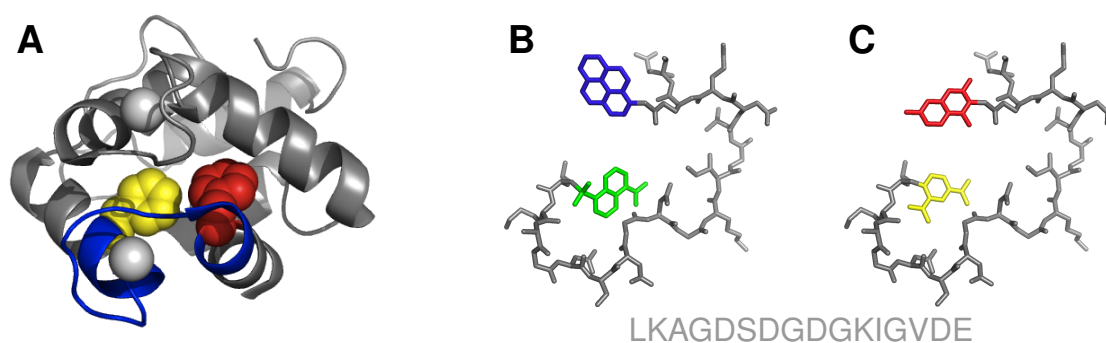


Figure 3.8.: (A) Crystal structure of β -carp Parvalbumin with calcium bound (pdb entry 4CPV). The EF-loop is shown in blue, the positions of the FRET-labels are highlighted in red (donor) and yellow (acceptor); isolated EF-loop sequence modified with FRET-labels pyrene (blue) and dansyl (green) (B) and modified with aminomethylcoumarin (red) and dinitrophenyl (yellow) (C)

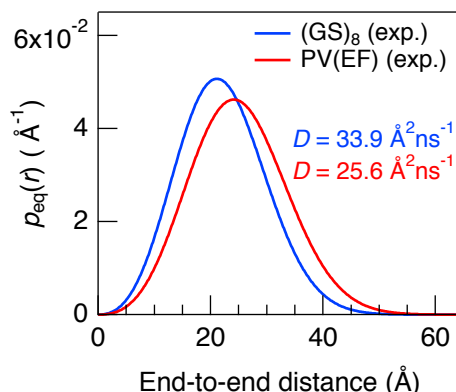


Figure 3.9.: End-to-end distance distributions for the two model peptides $(GS)_8$ and PV(EF) as obtained from global analysis of time resolved FRET experiments. The distance distributions clearly differ from the distance distribution of the random coil model.

calcium ions. In the isolated PV(EF), no secondary structure is formed in the presence and absence of calcium as indicated by the circular dichroism spectra.²⁵⁵ The EF-loop PV(EF) has 7 charged side-chains, several hydrophobic groups and 4 glycines. Measurements of the end-to-end loop formation by TTET have revealed a time constant of 53 ns,²²⁴ which is similar to the loop formation time constant of poly-serine of the same length and slightly faster compared to other natural sequences.²⁵⁶

3.2.3. End-to-end distance distribution and intra-chain diffusion coefficient

The two model peptides were chemically synthesized and modified with the corresponding FRET chromophores at the positions indicated in Figure 3.7 and 3.8. In both cases, the positions of the chromophores are separated by 16 amino acids. For the determination of the intra-chain diffusion coefficients and end-to-end distance distributions, the two different FRET-pairs AMCA/DNP and Pyr/Dans were used as described in section 3.1. Donor-only and acceptor-only labeled peptides were used for the determination of the corresponding donor lifetimes τ and the Förster distances R_0 (see chapter A). Time resolved fluorescence decays for the double labeled peptides in 10 mM potassium phosphate buffer (pH 7) were recorded using TCSPC. The data was analyzed by the global method described in section 3.1.4. Satisfactory fits could be obtained for both peptides using the end-to-end distance distribution $p_{eq}(r)$ of a chain with excluded volume (equation (1.12)) as shown in Figure 3.9. Intra-chain diffusion is faster for $(GS)_8$ with a diffusion coefficient of $33.9 \text{ \AA}^2 \text{ ns}^{-1}$ than

3. Results

for PV(EF) with a diffusion coefficient of $25.6 \text{ \AA}^2 \text{ ns}^{-1}$. The root mean square end-to-end distance is larger for PV(EF) with $\langle r^2 \rangle^{1/2} = 26.5 \text{ \AA}$ than for $(\text{GS})_8$ with $\langle r^2 \rangle^{1/2} = 23.5 \text{ \AA}$. For both peptides, the distance distribution differs from the one expected for a random coil. The difference is explicitly given through the parameter b of the excluded volume chain model. If b tends towards zero, the end-to-end distance distribution of the excluded volume model tends towards the distribution of the random coil model. Due to the accurate determination of the parameter b by global analysis, the random coil model can be excluded. The values for the parameters b and σ are summarized in Table 3.3.

3.2.4. The chain with excluded volume as a model for unfolded peptide conformations

The non-random coil like distribution of end-to-end distances is related to intrinsic properties of the poly-peptide chain. One possibility in this context is that unfolded peptides behave like polymers with excluded volume. The parameters b and σ are directly related to properties of the chain with excluded volume, i.e. σ is related to the number n of chain segments and to the length l of each segment, while b is related to n , l and the excluded volume v (see equation (1.11)). To simplify the calculations instead of $n \cdot l$ the contour length of the polymer L is written.

$$\sigma^2 = \frac{20}{27} L l \quad (3.6)$$

$$b = \left(\frac{5}{3}\right)^{\frac{3}{5}} \left(\frac{1}{3\pi}\right)^{\frac{1}{5}} \left(\frac{4}{3}\pi R^3\right)^{\frac{1}{5}} \left(\frac{1}{l}\right)^{\frac{1}{5}} L^{\frac{3}{5}} \quad (3.7)$$

In equation (3.7) the excluded volume $v = 4/3\pi R^3$ is used, with R the smallest possible distance between two arbitrary points along the chain resulting from a hard sphere repulsion. Thus, the volume v of a sphere of radius R around every point along the chain is devoid of other chain segments. By help of the equations (3.6) and (3.7), values for l and for the radius R of the excluded volume can be calculated for a given contour length L . Here we assume a contour length $L = 64.6 \text{ \AA}$ according to a distance of 17 residues of length 3.8 \AA separating the FRET labels. The resulting values for l and R are summarized in Table 3.3. For $(\text{GS})_8$ values of $l = 3.85 \text{ \AA}$ and of $R = 1.3 \text{ \AA}$ are found, while for PV(EF) $l = 4.45 \text{ \AA}$ and $R = 1.9 \text{ \AA}$. The calculated segment lengths l are close to the value 3.8 \AA expected for the length of one amino acid in a peptide. However, the values for the minimal allowed

Table 3.3.: Results from global analysis of time resolved FRET data and the corresponding parameters for the chain with excluded volume model.

	diffusion coefficient	parame- ter	parame- ter	contour length	segment length	radius (excluded volume)
	D ($\text{\AA}^2\text{ns}^{-1}$)	b (\AA)	σ (\AA)	L (\AA)	l (\AA)	R (\AA)
(GS) ₈	33.9	12.4	13.6	64.6	3.85	1.3
PV(EF)	25.6	15.3	14.6	64.6	4.45	1.9

distance of two chain segments appear to be very small. The van der Waals radius of carbon is 1.7 \AA and the van der Waals radius of nitrogen is 1.55 \AA ,²⁵⁷ resulting in a minimal possible distance between two carbon atoms of around 3.4 \AA . Thus the minimal possible distances R found here are too small by a factor of 2–3. The small values for R could be due to a potential of interaction between two segments along the polymer which differs from a hard sphere potential and comprises for example additional attractive interactions. It should also be noted, that the inclusion of steric repulsion into the random walk model does not lead to end-to-end distance distributions described by equation (1.12) but to distance distributions still well described by the random coil model as has been shown in Figure 1.2 B in section 1.3.1.

3.2.5. Monte Carlo modeling of unfolded peptide conformations

The shape of the distance distribution obtained from the experiments can originate in the properties of the poly-peptide chain rather than in properties of a polymer model. For instance, limitations of possible ϕ - ψ -angles imposed by steric repulsions could result in end-to-end distance distributions with a shape that can be well approximated by equation (1.12). To test this hypothesis, MC simulations of peptide conformations were carried out for both model peptides.

The MC sampling was done by the program *flexible meccano* (developed in the group of Martin Blackledge, Protein Dynamics and Flexibility by NMR, Institut de Biologie Structurale, Grenoble, France).^{134,258} The algorithm used by *flexible meccano* randomly selects ϕ - ψ -angle pairs from a database avoiding steric clashes through an amino acid specific excluded volume represented by spheres centered at the C_β atom or at the C_α atom for glycine.

3. Results

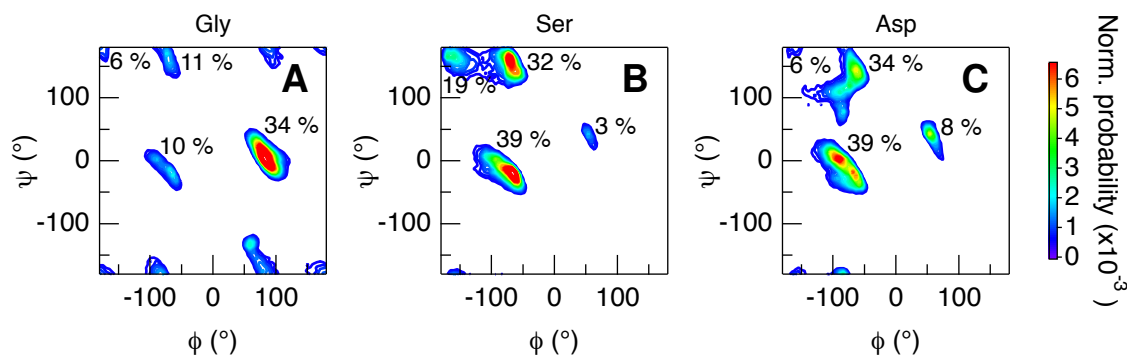


Figure 3.10.: Examples for amino acid specific Ramachandran plots obtained from the torsion angles of fragments of the PDB which are in coil conformation (see text for details). The Monte Carlo simulations for the two model peptides were carried out based on these ϕ - ψ -angles. The relative occupation of the major basins is displayed.

As the database provided by *flexible meccano* consists of only 500 ϕ - ψ -angle pairs per amino acid, a larger database obtained from the *The Protein Coil Library* (RoseLab, Jenkins Biophysics, John Hopkins University, USA)^{156,259} was implemented. This library lists the torsion angles of protein structure fragments from the Protein Data Bank (PDB), which are in the coil conformation and thus cannot be classified as α -helix or β -strand²⁶⁰. The torsion angles used for the database are extracted from crystal structures of resolution better than 2.0 Å. Examples for the Ramachandran plots used in the database are shown in Figure 3.10. 100,000 conformers of each peptide were calculated and further analyzed according to their conformational properties. The end-to-end distance, the distance between the N-terminal nitrogen and the C-terminal oxygen, of each member of the MC ensemble was calculated. The occurrence of end-to-end distances is shown in the histogram in Figure 3.11 A and B. The histograms were fit with the end-to-end distance distributions of the random coil model, of the wormlike chain model and of the chain with excluded volume model as shown in Figure 3.11 A and B. For (GS)₈, the best fit was obtained with the chain with excluded volume model, while for PV(EF) both, the wormlike chain and the chain with excluded volume model, produced acceptable fits. In both cases, the fits with the distance distribution of the random coil model did not result in adequate descriptions of the histograms. The parameters obtained from the fit of the MC histograms are compared with the experimentally obtained parameters. For both peptides, the parameters b obtained from the MC ensemble are slightly larger than the experimental values, while the parameters σ are smaller than the experimental values (compare Figure 3.11 and Table 3.3). This deviation can be due the

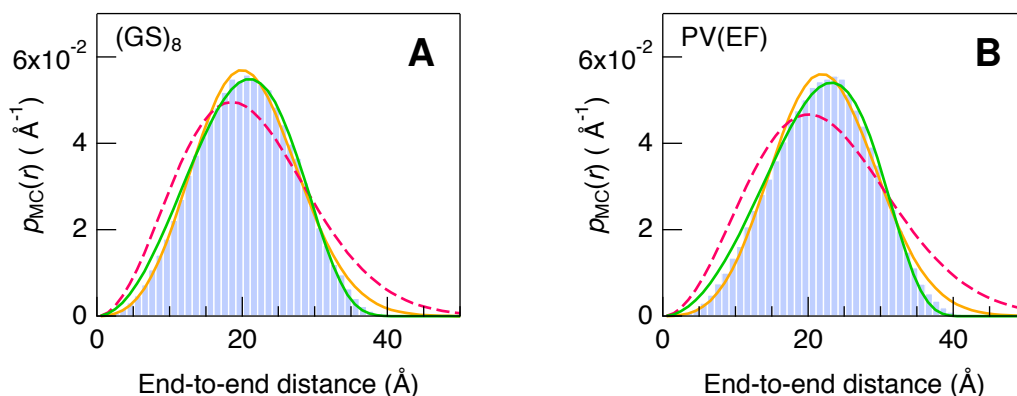


Figure 3.11.: Histograms obtained by calculating the end-to-end distances r of the Monte Carlo generated ensembles for $(GS)_8$ (A) and PV(EF) (B). Fits of the histograms with the chain with excluded volume model (—) result in the fit parameters $b = 12.9 \text{ \AA}$, $\sigma = 11.9 \text{ \AA}$ for $(GS)_8$ and $b = 15.6 \text{ \AA}$, $\sigma = 11.7 \text{ \AA}$ for PV(EF), fits with the wormlike chain model (—) result in $l_p = 7.6 \text{ \AA}$ and $L = 48.9 \text{ \AA}$ for $(GS)_8$ and $l_p = 9.1 \text{ \AA}$ and $L = 47.8 \text{ \AA}$ for PV(EF) and fits with the random coil model (---) result in $\langle r^2 \rangle^{1/2} = 22.7 \text{ \AA}$ for $(GS)_8$ and $\langle r^2 \rangle^{1/2} = 24.6 \text{ \AA}$ for PV(EF).

missing chromophores and their linkers in the MC simulation. Inclusion of chromophores and linkers into the simulations would most likely result in a broadened end-to-end distance distribution similar to the convolution of the actual distribution with a function modeling the freedom given by the linkers. However, the MC ensemble can serve as model for an unfolded ensemble and can help to understand properties of unfolded peptides when compared to experimental results.

It can be concluded, that the distribution of ϕ - ψ -angles imposed by the allowed regions of the Ramachandran plots for the different amino acids results in end-to-end distance distribution which can be well approximated by the formula for the end-to-end distance distribution of the excluded volume model (equation (1.12)).

3.2.6. End-to-end distance and radius of gyration of unfolded polypeptides

Through the MC generated ensemble, the relation between the end-to-end distance r and the radius of gyration r_G for unfolded peptides can be studied independent of the validity of a polymer model. The radius of gyration of each conformer was directly obtained from *flexible meccano* and correlated with the calculated end-to-end distance for both model peptides as displayed in Figure 3.12. Even though the radius of gyration and the end-to-end distance are correlated, the values of both parameters are broadly distributed over the en-

3. Results

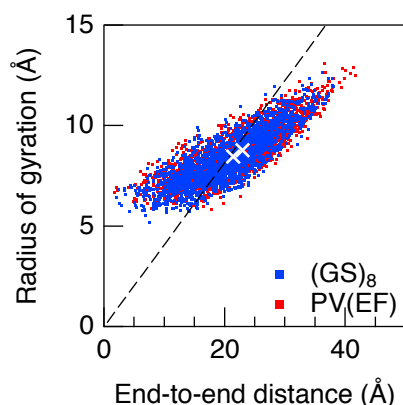


Figure 3.12.: Correlation between the radius of gyration r_G and the end-to-end distance of the members of the Monte Carlo generated ensembles for both model peptides. The relation $\langle r_G^2 \rangle = \langle r^2 \rangle / 6$ (equation (1.8)) of the random coil model is marked (--) and the average values $\langle r_G^2 \rangle^{1/2}$ versus $\langle r^2 \rangle^{1/2}$ for both peptides are indicated by white crosses.

semble. Surprisingly, the ensemble averaged values $\langle r_G^2 \rangle$ and $\langle r^2 \rangle$ show almost the same relation as would be true for a random coil polymer, $\langle r_G^2 \rangle = \langle r^2 \rangle / 6$ (equation (1.8)), as indicated in Figure 3.12. However, this is clearly not an indication for the validity of the entire random coil model, but rather for unfolded peptides showing some features that are well described by the random coil model. Moreover, this observation is based on the theoretical model ensemble and validity could not be checked for experimental results in this context.

3.2.7. Summary and discussions

The time resolved fluorescence data are well fit by the end-to-end distance distribution of the excluded volume chain model for both peptides studies. Due to the global analysis method used, the parameters were accurately determined, clearly excluding the random coil model for the description of the end-to-end distance distribution. As indicated by MC simulations of unfolded peptide conformations, the experimental distance distribution originates in the amino acid specific sampling of ϕ - ψ -angles.

The experimentally determined distance distributions clearly indicate, that unfolded peptides are not well described as random coil polymers, which had been assumed based on the validity of scaling laws (see section 1.3.1)¹⁰¹⁻¹⁰⁴. The interpretation of unfolded peptides behaving like random coil polymers based on scaling laws has already been put in doubt elsewhere.^{120,122}

Both the experimental findings and the MC ensembles show, that the end-to-end distance distribution of the random coil model is not a good approximation for unfolded ensembles of peptides and proteins, even though the random coil model is widely used as distance distributions⁷⁵⁻⁹⁵. Due to the accuracy of the parameters in the global analysis, it can be concluded, that the end-to-end distance distribution of the excluded volume chain model is a good approximation for the end-to-end distance distribution in an unfolded ensemble of peptides. This model has already been successfully used to fit time resolved fluorescence data of unfolded peptides¹⁶⁸⁻¹⁷⁴.

The cause for the observed shape of the distance distribution is most likely not due to the behavior of unfolded peptides as chains with excluded volume, but rather due to amino acid specific sampling of dihedral angles. End-to-end distance distributions of MC generated model ensembles can be well approximated by the excluded volume chain model or by the wormlike chain model. The qualitatively good agreement between the end-to-end distance distributions of the MC ensembles with the experimental distance distribution is in good agreement with findings, that parameters of intrinsically disordered proteins measured by NMR can be well described through similar MC generated model ensembles.^{134,148} The MC based model ensembles could, taking into count the chromophores and linkers, also serve as a starting ensemble from which sub-ensembles are selected to exactly fit the results from FRET experiments. This is an approach also used to gain comprehensive insight on unfolded proteins by combining different experimental techniques like NMR and SAXS.^{136,143,145,150}

3.3. Effect of Solvent friction and internal friction on intra-chain diffusion of unfolded polypeptides

3.3.1. Friction in the dynamics of polymers and unfolded polypeptides

An unfolded peptide in solution constantly changes conformation through temperature-driven Brownian motion. In the microscopic view, conformational changes occur via series of bond rotations impeded by local energy barriers, by favorable or unfavorable interactions or by the presence of solvent molecules. When passing over to the macroscopic view, a series of approximations has to be made, i.e. the solvent molecules are replaced by a continuum and conformational changes are described as intra-chain diffusion in the unfolded ensemble. In this view, intra-chain diffusion is hindered by frictional forces. A major source of friction is the viscosity of the solvent. However, other contributions caused by intrinsic properties of the chain, like steric effects or intra-molecular interactions may also limit intra-chain diffusion.²⁶¹ All these contributions are summed under the term *internal viscosity* or *internal friction*, while the contribution caused by the solvent is referred to as *solvent friction*.

The relation between the diffusion coefficient and friction is described by the Einstein-Smoluchowski equation^{234,235}:

$$D = \frac{k_{\text{B}}T}{f} \quad (3.8)$$

In this equation, $k_{\text{B}}T$ is the Boltzmann constant times the absolute temperature, f is the coefficient of friction and D stands for the intra-chain diffusion coefficient in the case discussed here. Thus, the coefficient of friction can be directly calculated from the intra-chain diffusion coefficient of an unfolded polypeptide determined by time-resolved FRET measurements as described in chapter 3.1.

The focus of the present study will be on the role of friction in the dynamics of unfolded peptides. To access in particular contributions from both, solvent friction and internal friction, the viscosity dependence of friction will be studied in detail. The determination of internal friction is particularly important because it sets an upper speed limit for the dynamics in the unfolded state.

In case of a free particle in solution, the coefficient of friction is directly proportional to the viscosity of the solvent according to the Stokes-Einstein law.^{234,262} In the case of intra-chain diffusion the relation of friction to solvent viscosity is less straight forward as the chain it-

3. Results

self most likely also affects the coefficient of friction.

The term *internal viscosity* has first been introduced to model the properties of threadlike polymers in solution^{263–266} and has been applied to describe the physical characteristics of dilute polymer solutions.^{267–269} The implementation of internal viscosity affects for example the diffusional attempt for crossing an activation barrier as in Kramers' theory leading to derivations from the linear viscosity dependence of friction at small viscosities^{270,271} or the relaxation modes in a Rouse polymer (a polymer consisting of infinitely small beads connected by springs)²⁷². Internal contributions to friction are best accessed in the absence of solvent viscosity, thus in the absence of solvent friction. However, this approach is experimentally impracticable. Accordingly, an analytical description is needed to allow for extrapolation to the zero viscosity limit and determination of the internal contributions.

Haas et al. were the first to apply the equations derived for linear polymers like polystyrene^{273,274} to unfolded polypeptides proposing an additive relation of contributions from the solvent and internal friction f_{int} to influence the dynamics of unfolded polypeptides.¹⁷⁰

$$f = f_{\text{solv}} \left(\frac{\eta}{\eta_0} \right) + f_{\text{int}} \quad (3.9)$$

In this equation, f_{solv} is the coefficient of solvent friction, η is the viscosity of the solvent and η_0 a reference viscosity. From this equation a linear dependence of the coefficient of friction on solvent viscosity with an intercept at vanishing viscosity, which corresponds to internal friction, is expected. This has recently been applied to describe the viscosity dependence of reconfiguration times in unfolded and intrinsically disordered proteins yielding a limiting internal reconfiguration time which vanishes with increasing denaturant concentration.⁷⁵ However, earlier studies on loop formation in model peptides suggested an internal viscosity equal to zero when applying equation 3.9.²²³

Friction in the dynamics of the unfolded state may also affect folding rates of proteins. The influence of intra-chain diffusion and thus friction on folding rate constants is described by Kramers' theory.²³⁹ Applying this theory, the influence of internal friction on folding time constants can be determined by extrapolating the viscosity dependence of folding times of proteins or peptides to zero viscosity in analogy to equation (3.9).^{275–284} Both, vanishing^{275,277,283} and non-vanishing^{279–282,284} internal friction have been reported.

An alternative expression with a scaling exponent rather than an additive contribution from internal friction has been proposed for the viscosity dependence of loop formation rate constants in unfolded model peptides determined by TTET.¹⁷⁶ In terms of the coefficient of

friction f , this expression reads:

$$f = f_{\text{solv}} \left(\frac{\eta}{\eta_0} \right)^\beta \quad (3.10)$$

For the loop formation in (GS)-peptides of different length, β has been found to be between 0.8 and 0.96.¹⁷⁶ The same expression has been tested on secondary structure formation rate constants resulting in $\beta = 0.64$ ²⁸³ and to simulations of the viscosity dependence of first passage times in different model peptides resulting in β values between 0.59 and 0.73.²¹⁴ In the present study, the applicability of equation (3.9) and (3.10) for the friction limiting intra-chain diffusion will be tested. These expressions can also be used to test for contributions from internal and solvent friction. Firstly, the viscosity dependence of the intra-chain diffusion coefficient and the end-to-end distance distribution is determined. From the intra-chain diffusion coefficients, the viscosity dependence of the coefficient of friction can be obtained. This allows to test the two different expressions equation (3.9) and (3.10) and to separate internal friction from solvent friction. Based on these experiments, a slightly non-ideal viscosity dependence is found and an upper limit for internal contributions can be set.

3.3.2. Effect of glycerol on the intra-chain diffusion coefficient and the end-to-end distance distribution

In order to increase solvent viscosity, a viscose co-solute is added to the solvent. Various substances like ethylene glycol, glycerol, sucrose, poly-ethylene glycols or ficoll are known to increase the macroscopic solvent viscosity. However, all these different co-solutes have been found to have different effects on the loop formation rate constants of unfolded model peptides determined by TTET.²⁵⁶ The larger co-solutes have a smaller effect on loop formation than small co-solutes, but below a certain size, all co-solutes have the same effect. Thus the small size co-solute glycerol was chosen for the present study. Glycerol has a hydrodynamic radius $r_H = 2.2 \pm 0.2 \text{ \AA}$ smaller than the length of one peptide unit of $\approx 3.8 \text{ \AA}$ and therefore has the same effect on loop formation in (GS)-peptides as ethylene glycol and sucrose as shown by TTET measurements.²⁵⁶

FRET-measurements were carried out in buffers with 10 mM potassium phosphate at pH 7 and 22.5°C containing different amounts of glycerol ranging from 0–56 % w/w. Donor lifetimes, quantum yields and Förster distances were determined for each condition using

3. Results

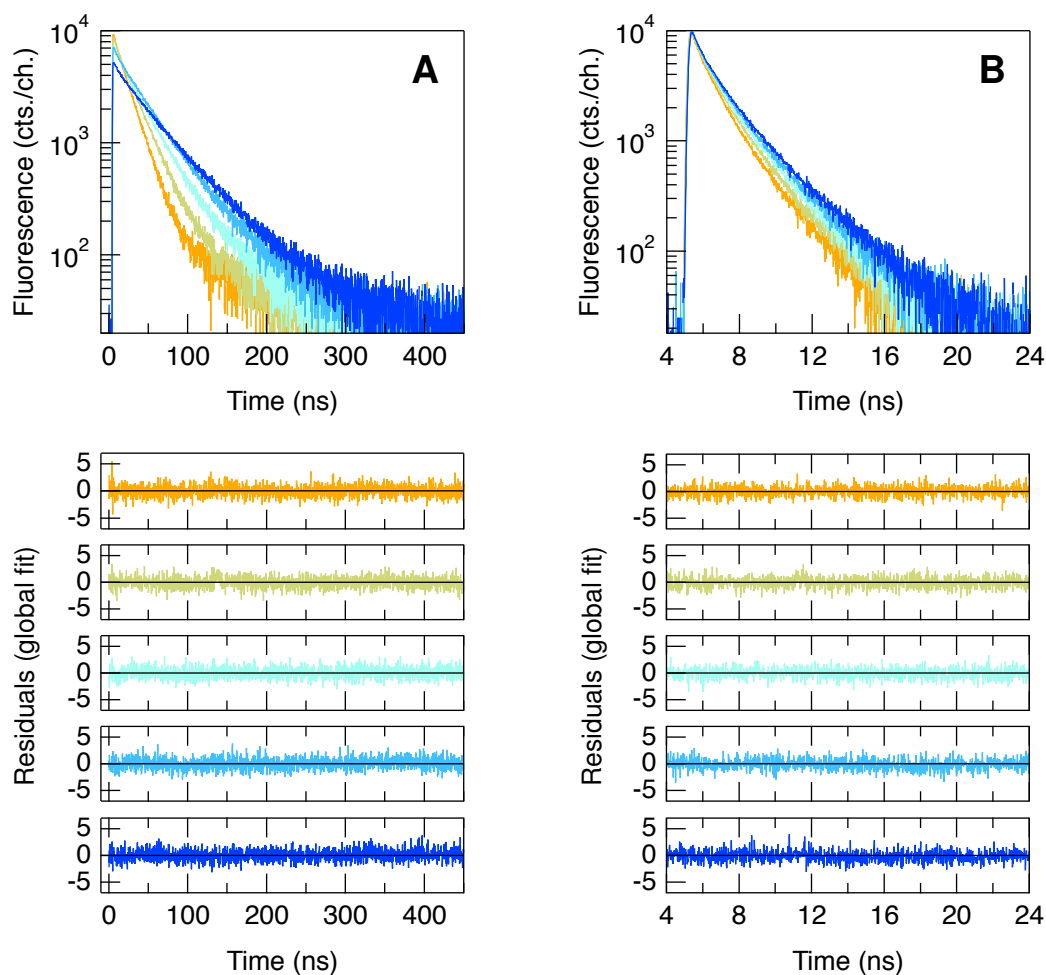


Figure 3.13.: Time resolved fluorescence decays for the (GS)₈ peptide labeled with the donor pyrene (Pyr) and the acceptor dansyl (Dans) near the termini (A) and with the donor aminomethylcoumarin (AMCA) and the acceptor dinitrophenyl (DNP) (B) in 10 mM potassium phosphate, pH 7 and glycerol contents of 16 %, 28 %, 40 %, 48 % and 56 % w/w (colors from yellow to blue). The residuals of the global analysis of the two different FRET pairs are shown; each color corresponds to one global fit, thus to one glycerol concentration. The decrease in the amplitude in A is due to an increase of background fluorescence of the buffer with the glycerol content which has been subtracted from the data.

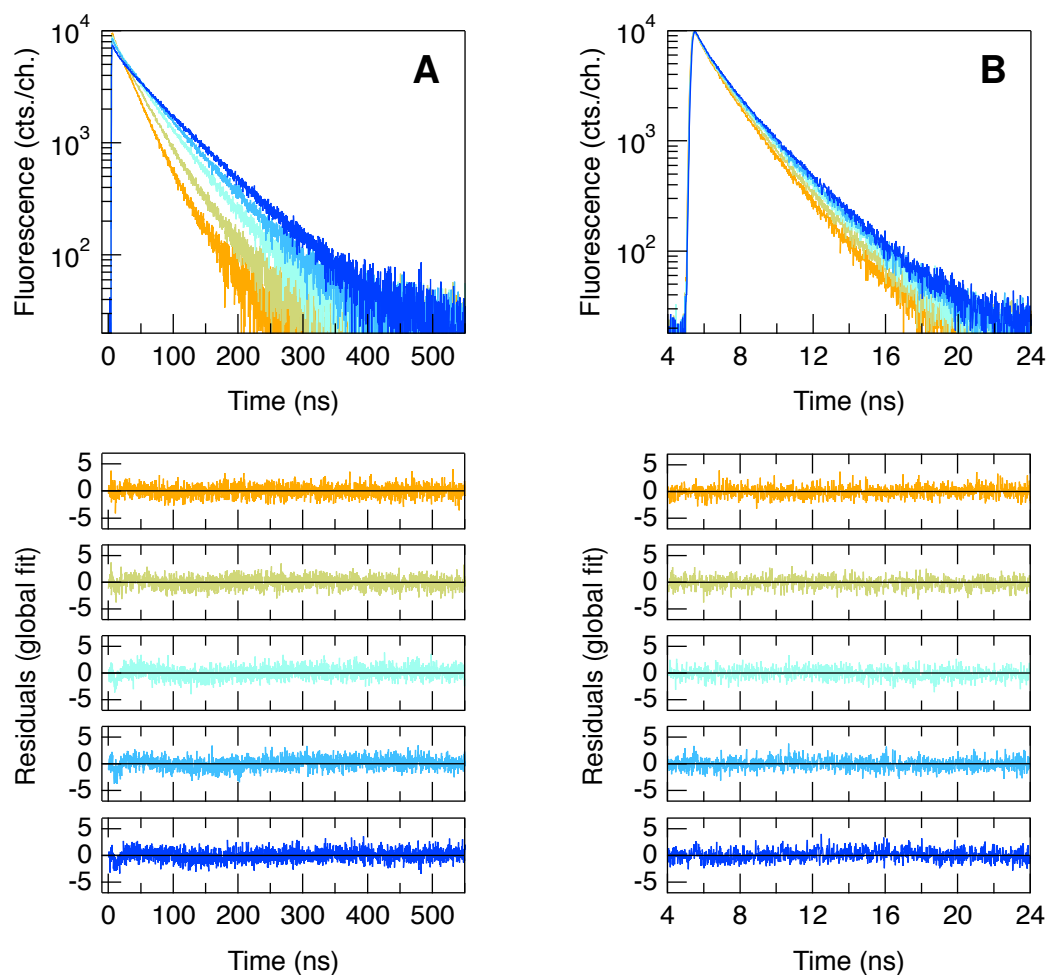


Figure 3.14.: Time resolved fluorescence decays for the PV(EF) peptide labeled with the donor pyrene (Pyr) and the acceptor dansyl (Dans) near the termini (A) and with the donor aminomethylcoumarin (AMCA) and the acceptor dinitrophenyl (DNP) (B) in 10 mM potassium phosphate, pH 7 and glycerol contents of 16 %, 28 %, 40 %, 48 % and 56 % w/w (colors from yellow to blue). The residuals of the global analysis of the two different FRET pairs are shown; each color corresponds to one global fit, thus to one glycerol concentration. The decrease in the amplitude in A is due to an increase of background fluorescence of the buffer with the glycerol content which has been subtracted from the data.

3. Results

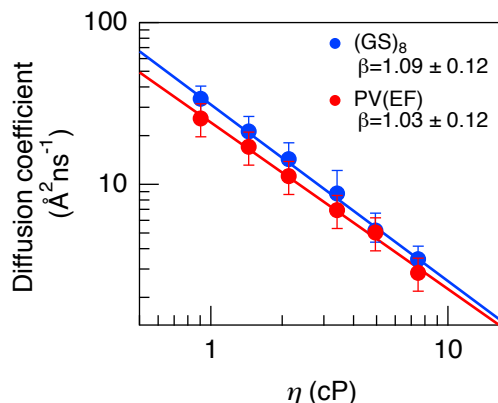


Figure 3.15.: Dependence of the intra-chain diffusion coefficient D on solvent viscosity η . The solid lines represent fits according to equation 3.11.

donor-only and acceptor-only labeled peptides. Time resolved fluorescence decays were recorded for each of the two FRET-pairs by TCSPC. These were globally analyzed as described in chapter 3.1 yielding the intra-chain diffusion coefficient D and the end-to-end distance distribution p_{eq} for the two different model peptides (GS)₈ and PV(EF) introduced in chapter 3.2.

Intra-chain diffusion is slowed down by a factor of 9.8 for (GS)₈ and of 9.0 for PV(EF) with a viscosity increase by a factor of 8.3 in 56 % glycerol compared to water. Figure 3.15 shows the decrease in D with increasing solvent viscosity in a double logarithmic graph. The dependence of D on the macroscopic solvent viscosity η can be described by the following empirical relation:

$$D = D_0 \cdot \left(\frac{\eta}{\eta_0} \right)^{-\beta} \quad (3.11)$$

In this equation η_0 is a reference viscosity and D_0 is the diffusion coefficient at that viscosity (here 10 mM potassium phosphate at 22.5°C is used as reference condition). The dimensionless scaling exponent β can account for deviations from the ideal η^{-1} dependence. Values of $\beta = 1.09 \pm 0.12$ and of $\beta = 1.03 \pm 0.13$ are found for (GS)₈ and PV(EF) respectively. A β above 1 indicates, that the increase in solvent viscosity slows down intra-chain diffusion more than would be expected for free diffusion for which β equals 1.

The end-to-end distance distributions of (GS)₈ and PV(EF) change slightly with glycerol

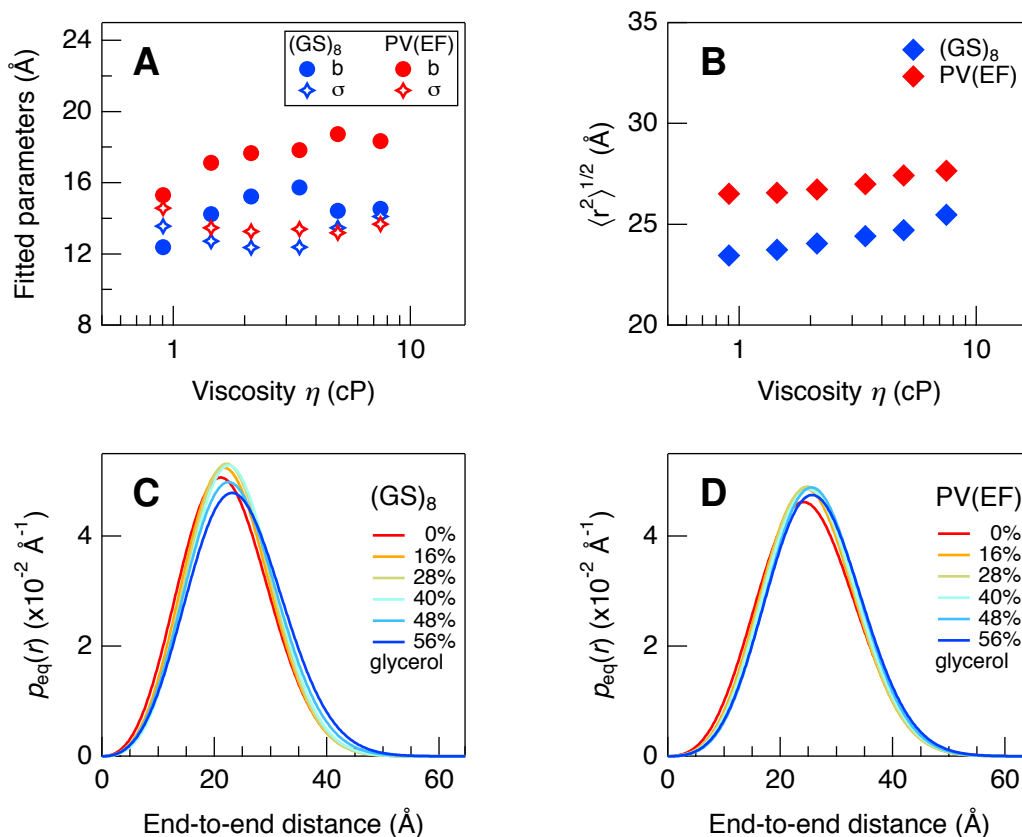


Figure 3.16.: Effect of a glycerol induced increase of solvent viscosity. A: Fitted parameters b and σ describing the end-to-end distance distribution according to equation 1.12. B: root mean square end-to-end distance. C and D: end-to-end distance distributions.

leading to an increase of the root mean square end-to-end distance of 9 % for $(GS)_8$ and of 4 % in 56 % glycerol compared to water (see Figure 3.16).

3.3.3. The role of friction for the dynamics of unfolded polypeptides

The total coefficients of friction f were calculated from the intra-chain diffusion coefficients using equation (3.8). As shown in Figure 3.17, a slight upward curvature is observed for the viscosity dependence of friction for $(GS)_8$. The viscosity dependence of friction was analyzed by help of the following equation:

$$f = f_{\text{solv}} \left(\frac{\eta}{\eta_0} \right)^\beta + f_{\text{int}} \quad (3.12)$$

3. Results

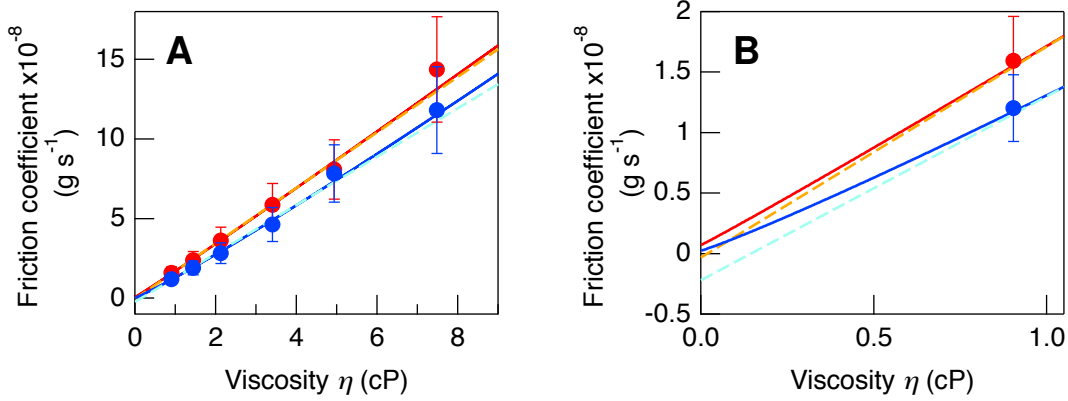


Figure 3.17.: Viscosity dependence of the coefficient of friction for $(GS)_8$ (●) and PV(EF) (●). Solid lines are fits of equation (3.9), dashed lines are fits of equation (3.12) with β fixed at the values determined in Figure 3.15. The resulting fit parameters are $f_{\text{solv}} = (1.15 \pm 0.18) \cdot 10^{-8} \text{ gs}^{-1}$, $f_{\text{int}} = (0.03 \pm 0.4) \cdot 10^{-8} \text{ gs}^{-1}$, $\beta = 1.09$ (—) and $f_{\text{solv}} = (1.48 \pm 0.25) \cdot 10^{-8} \text{ gs}^{-1}$, $f_{\text{int}} = (0.07 \pm 0.5) \cdot 10^{-8} \text{ gs}^{-1}$, $\beta = 1.03$ (—) for fits with according to equation (3.12) as well as $f_{\text{solv}} = (1.37 \pm 0.22) \cdot 10^{-8} \text{ gs}^{-1}$, $f_{\text{int}} = (-0.2 \pm 0.4) \cdot 10^{-8} \text{ gs}^{-1}$ (---) and $f_{\text{solv}} = (1.57 \pm 0.26) \cdot 10^{-8} \text{ gs}^{-1}$, $f_{\text{int}} = (-0.03 \pm 0.5) \cdot 10^{-8} \text{ gs}^{-1}$ (---) for fits according to equation (3.9).

For fits of the viscosity dependence of the coefficient of friction, β was fixed at the values obtained earlier (see Figure 3.15). The resulting fit describes the curvature of the data for $(GS)_8$. The intercept at zero solvent viscosity gives information about the internal friction. Extrapolation to zero solvent viscosity results in values close to zero, $f_{\text{int}} = (0.03 \pm 0.36) \cdot 10^{-8} \text{ gs}^{-1}$ for $(GS)_8$ and $f_{\text{int}} = (0.07 \pm 0.5) \cdot 10^{-8} \text{ gs}^{-1}$ for PV(EF) as shown in Figure 3.17 A and B. This result is however not surprising, as equation (3.12) and equation (3.11) are only equivalent for zero internal friction.

To investigate, which equation yields the better description of the data, fits with β set to 1 were performed. For $\beta = 1$ equation (3.12) reduces to equation (3.9). Intercepts of $(-0.2 \pm 0.4) \cdot 10^{-8} \text{ gs}^{-1}$ and $(-0.03 \pm 0.54) \cdot 10^{-8} \text{ gs}^{-1}$ are found for $(GS)_8$ and PV(EF) respectively (see Figure 3.17 B). Negative values for f_{int} are clearly physically meaningless, therefore only upper limits for internal friction can be obtained from these fits. The upper limits are $0.2 \cdot 10^{-8} \text{ gs}^{-1}$ for $(GS)_8$ and $0.5 \cdot 10^{-8} \text{ gs}^{-1}$ for PV(EF). However, these fits fail to describe the curvature in the data.

Internal friction seems to be absent in the dynamics of the two unfolded model-peptides, at least as an additive contribution. Consequently, all processes relevant for chain motion scale with solvent viscosity and are governed either by the term f_{solv} or by the exponent β . Solvent

friction is found to be higher for PV(EF) than for (GS)₈ adopting values of $f_{\text{solv}} = 1.48 \pm 0.25 \text{ gs}^{-1}$ for PV(EF) and $f_{\text{solv}} = 1.15 \pm 0.18 \text{ gs}^{-1}$ for (GS)₈ if the experimental β -values are used. Both peptides have the same number of residues between the labels, excluding chain length as possible cause for the difference. The difference in solvent friction thus results from the different amino acid compositions of the peptides.

3.3.4. Summary and discussions

Solvent viscosity strongly affects intra-chain diffusion. The intra-chain diffusion coefficient decreases with increasing glycerol content, thus with increasing solvent viscosity following a slightly non-ideal dependence with $D \propto \eta^{-\beta}$. The coefficient of friction f for intra-chain diffusion is found to be dominated by solvent friction as additive contributions from internal friction are not detectable within error. The end-to-end distance distributions are only slightly affected by glycerol.

Intra-chain diffusion is strongly affected by the increase of solvent viscosity following a non-ideal dependence with β -values slightly above one. The $\beta \neq 1$ indicate, that processes besides solvent viscosity influence chain dynamics. One explanation for $\beta > 1$ could be the slight change of the root mean square end-to-end distance, which increases by 9 % for (GS)₈ ($\beta = 1.09$) and by 4 % for PV(EF) ($\beta = 1.03$). Up to date, it has however not been studied systematically whether the intra-chain diffusion coefficient is directly related to the root mean square end-to-end distance. Alternatively, an upward curvature of the coefficient of friction deviating from the η^{-1} dependence has also been proposed for diffusional attempt for crossing a free energy barrier.²⁷⁰ However, for the peptides used here, no indication for an energy barrier close to the average end-to-end distance is given. It is more probable, that the deviations of β from 1 result from a rough energy landscape with minor barriers hindering motion (see results in chapter 3.5). Solvent friction is the dominant term in friction limiting peptide intra-chain diffusion. This is in good agreement with the folding rate constant of a small β -hairpin peptide which is likewise dominated by solvent friction.²⁸⁵ Solvent friction is larger for the native sequence PV(EF) than for the (GS)₈ peptide, which differ in their amino acid composition. This can either be due to the larger conformational freedom of glycine and thus the higher flexibility of glycine rich peptides or to supplementary interactions formed by the charged and hydrophobic side chains of PV(EF).

It has been reported that intrinsic viscosity is a frequency dependent term in macromolecules.²⁶⁹ Here an additive term from intrinsic viscosity is absent. However, solvent viscos-

3. Results

ity affects different modes of chain dynamics differently. For modes covering rearrangements not far from the average equilibrium end-to-end distance like intra-chain diffusion, the observed β -values are 1.09 ± 0.12 for (GS)₈ and 1.03 ± 0.12 for PV(EF). The β -values describing the viscosity dependence of end-to-end loop formation determined by TTET are 0.97 ± 0.01 for (GS)₈²²⁷ and 0.92 ± 0.01 ²⁸⁶ for PV(EF) respectively and thus deviate from the ones describing intra-chain diffusion. It could be argued, that end-to-end loop formation and end-to-end intra-chain diffusion are coupled processes.²³⁸ However this is only true if reconfiguration within the chain is governed by one single diffusion coefficient.²³⁸

Internal friction for intra-chain diffusion is zero within error. This is consistent with results from TTET experiments using different viscous co-solutes.²⁵⁶ Internal friction is a peptide intrinsic parameter and thus independent of the co-solute used in the viscosity dependence. For different co-solutes, different β -values are found in TTET, but under consideration of these values, all viscosity dependences of loop formation time constants extrapolate to a time constant equal zero in the absence of solvent viscosity.²⁵⁶

In summary, an additive term for internal friction is neither present in the viscosity dependence of intra-chain diffusion nor in the viscosity dependence of loop formation time constants^{256,286}.

Internal friction has also been reported to equal zero in loop formation kinetics following a linear viscosity dependence with $\beta = 1$.²²³ On the other hand, Soranno et al report a non-zero internal friction for the reconfiguration times of unfolded proteins at zero viscosity, which were calculated from single molecule FRET efficiencies and nanosecond fluorescence correlation spectroscopy.⁷⁵ However, these findings are based on the physically inconsistent assumption of the simultaneous validity of different polymer models like the random coil model for the distance distribution and the Rouse model for the reconfiguration times and discard curvatures in the viscosity dependences.

In contrast to often made assumption, the end-to-end distance distribution is slightly affected by glycerol. Glycerol has been found to interact favorable with charged side-chains but unfavorably with serine and the backbone.²⁸⁷ This can lead to conformational changes in the unfolded ensemble which are found to be slightly more pronounced for (GS)₈ than for the highly charged PV(EF).

In conclusion, the viscosity dependence of the coefficient of friction for intra-chain diffusion differs from the viscosity dependence expected for a Rouse polymer with an additive contribution from internal friction. Instead, the viscosity dependence of the coefficient of friction is well described by a sequence dependent coefficient of solvent friction and a scal-

3.3. Effect of Solvent friction and internal friction ...

ing exponent β accounting for deviations from the viscosity dependence of free diffusion. No additive term for internal friction is needed to describe the data.

3.4. Influence of denaturant and stabilizing osmolyte on the properties of unfolded polypeptides

3.4.1. Effect of denaturants and stabilizing osmolytes on protein stability

Co-solutes, such as low-molecular substances or ions, can influence protein stability by interacting with the polypeptide chain. They can act on the unfolded ensemble thus changing its properties or interact with the folded structure.

The influence of co-solutes especially on the unfolded state can be tested using unfolded model peptides lacking a well defined native structure (i.e., the peptides introduced in section 3.2.1 and section 3.2.2). The influence of certain co-solutes on conformational properties and on the dynamics in the unfolded state can be studied by determining the end-to-end distance distribution and the intra-chain diffusion coefficient from time-resolved FRET measurements. Characterizing the unfolded state in presence of different co-solvents can provide insights on the mechanism of co-solutes induced changes in protein stability.

Various co-solvent are present in cells, helping organisms to survive under relatively harsh environmental stress caused by high salt concentrations, dehydration or high temperatures. Some of these co-solutes assure correct hydration of the cell through osmosis or functionality of the cellular proteins.^{288,289} Therefore, these compounds are called natural *osmolytes*. The stability of proteins can be dramatically influenced by co-solutes in both directions. Denaturants like urea or GdmCl unfold proteins, while protecting osmolytes like Trimethylamine N-oxide (TMAO) or sarcosine shift the equilibrium toward the native state thus preventing unfolding²⁹⁰ or counteracting denaturation²⁹¹. Protecting osmolytes can even force unfolded proteins like intrinsically disordered proteins (IDPs) to adopt the folded state.^{292,293} The influence of co-solutes on protein stability can be explained by assuming some interaction between the co-solute molecule and the polypeptide chain. In the *weak binding model* proposed by Schellman,^{66,67,294,295} the co-solute molecules can interact with binding sites along the polypeptide chain. These binding sites are identical and interaction is only possible if the binding site is solvent exposed. To account for the typically high concentrations of co-solute (typically in the range of several molar) and the weak interaction between co-solutes and proteins, a site exchange formalism is introduced.^{296,297} In this formalism, the free energy of exchanging a water molecule for a co-solute molecule with equilibrium constant K_{ex} is taken into count rather than the binding free energy. The stability ΔG° of the protein in presence of a mole fraction $X_{\text{cos.}}$ of co-solvent in the solution depends on the

3. Results

difference in number of binding sites Δn exposed in U and N.²⁹⁶

$$\Delta G^\circ = \Delta G^\circ(\text{H}_2\text{O}) + \Delta n \cdot RT \ln(1 + (K_{\text{ex.}} - 1)X_{\text{cos.}}) \quad (3.13)$$

Δn between U and N was empirically found to be proportional to the change in solvent accessible surface area (SASA) of the protein upon folding.²⁹⁸

A slightly different theory was proposed in Tanford's *transfer energy model*.^{299,300} Applying the concept of summing the change of free energy over all chemical groups i times their solvent accessible surface α_i , the total free energy of transfer $\Delta G^{\circ \text{tr}}$ of the protein (in either N or U) from water to a co-solute solution can be obtained.³⁰¹

$$\Delta G_{\text{U}}^{\circ \text{tr}} = \sum_i \alpha_i^{\text{U}} \Delta g_i^{\circ \text{tr}} ; \quad \Delta G_{\text{N}}^{\circ \text{tr}} = \sum_i \alpha_i^{\text{N}} \Delta g_i^{\circ \text{tr}} \quad (3.14)$$

The transfer free energies $\Delta g_i^{\circ \text{tr}}$ of the individual groups when transferred from water to a co-solute solution are determined by the help of solubility studies³⁰² of small model compounds in the presence and absence of the corresponding co-solute.³⁰³ This procedure was adapted by Bolen and co-workers to a variety of different osmolytes separating the influence of backbone and side chain contributions to $\Delta g_i^{\circ \text{tr}}$.³⁰⁴⁻³⁰⁸ The transfer free energies can be placed in a thermodynamic cycle together with the stabilities of protein ΔG° in presence and absence of co-solute (see Figure 3.18).³⁰⁰ Through this circle, comparison of the difference in transfer free energy between U and N to experimentally determined values is possible²⁸⁷. In order to calculate the transfer free energy of the unfolded state in presence and absence of co-solvent, the exact solvent accessible surface areas need to be known. Bolen et al use models for the unfolded state. However, different models for the unfolded state may be required in the presence and absence of the co-solvent. Denaturants for example, are believed to favor poly-proline II conformation of the backbone dihedral angles³⁰⁹⁻³¹¹ resulting in different GdmCl than for other dihedral angles.³¹² On the other hand, stabilizing osmolytes increase the free energy of the unfolded state,³⁰⁷ which is believed to lead to compaction³¹³ and strengthening of intra-chain H-bonds³¹⁴.

The end-to-end distance distribution of an ensemble of unfolded peptides together with the intra-chain diffusion coefficient determined by time resolved FRET measurements provide insights on conformational changes in the unfolded state. In this study, it is found that the denaturant GdmCl increases the end-to-end distance indicating an expansion of the unfolded state, while the stabilizing osmolyte sarcosine has only little influence on the end-to-end

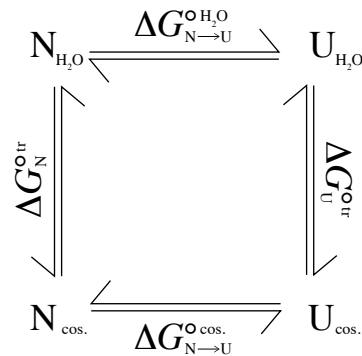


Figure 3.18.: Thermodynamic cycle used in the transfer model. Unfolding free energies $\Delta G_{N \rightarrow U}^{o}$ in the presence and absence of co-solute and the transfer free energies $\Delta G^{o, tr}$ of the native state N or the unfolded state U from water to co-solute solution. (See for example reference³⁰⁷)

distance. For both co-solutes, the viscosity corrected intra-chain diffusion coefficient is not affected.

3.4.2. End-to-end distance distribution and intra-chain diffusion in the presence of GdmCl

For the two model peptides, the $(GS)_8$ -peptide and the unfolded loop sequence PV(EF), fluorescence measurements were carried out under 9 different denaturant concentrations ranging from 0–8 M GdmCl. For each peptide, time resolved fluorescence decays were recorded for the two different FRET pairs as described in chapter 3.1. For each peptide and FRET pair, the donor lifetime τ and the Förster distance R_0 were determined using donor-only and acceptor-only labeled peptides in the corresponding GdmCl buffer. For each buffer condition, global analysis of the time resolved FRET measurements of the two different FRET pairs (Pyr/Dans and AMCA/DNP) were performed based on equation (3.1) as described. The raw data together with the residuals of the global fits are shown in Figure 3.19 and Figure 3.20.

The apparent intra-chain diffusion coefficient determined from global analysis is found to decrease by a factor of 2.4 for $(GS)_8$ and a factor of 2.1 for PV(EF) from water to the highest GdmCl concentration (see Figure 3.21 A). The viscosity of 8 M GdmCl solution is 2.3 fold higher than the viscosity of water at 22.5°C. The intra-chain diffusion coefficients were thus

3. Results

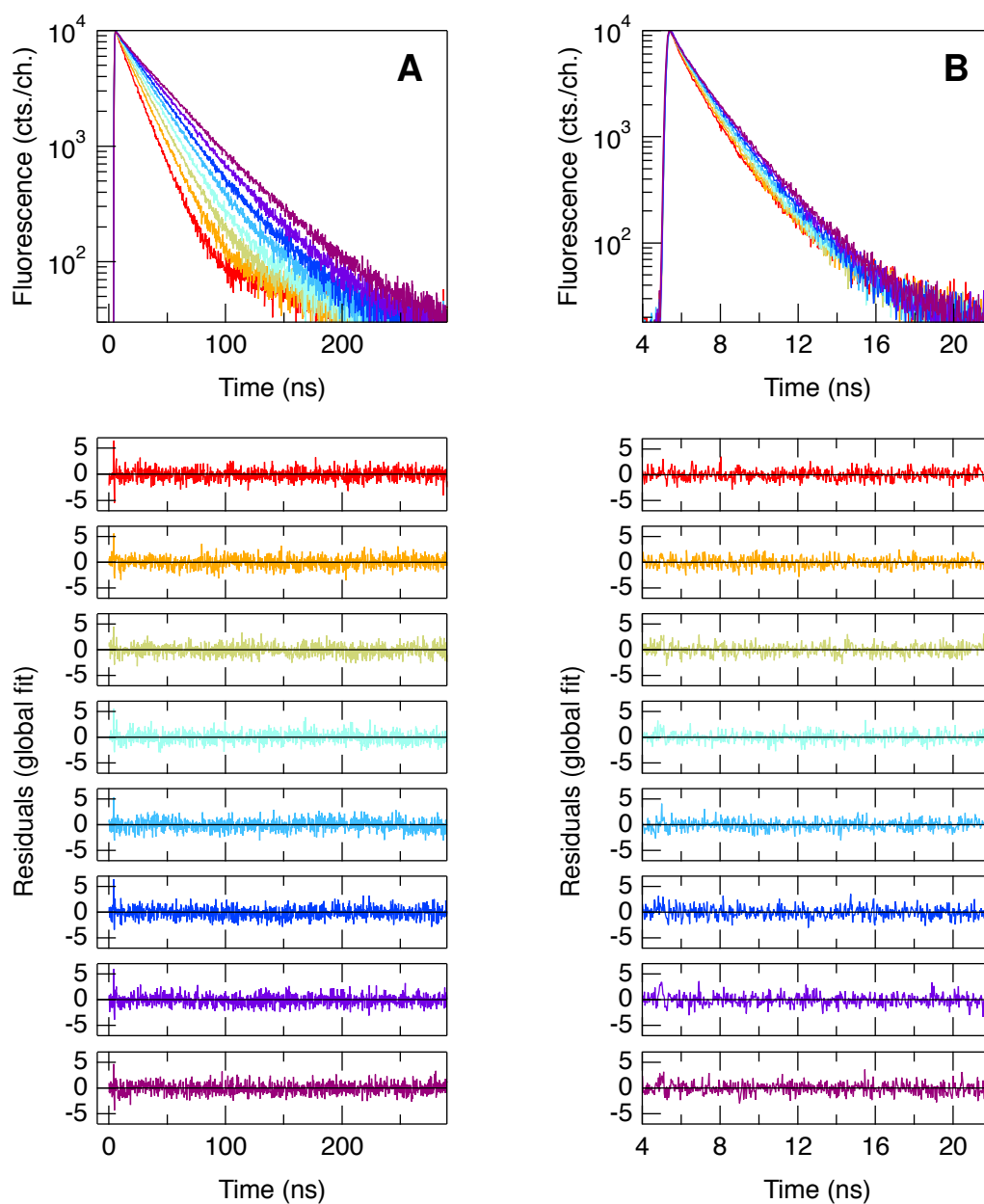


Figure 3.19.: Time resolved fluorescence decays for a (GS)₈ peptide labeled with the donor pyrene (Pyr) and the acceptor dansyl (Dans) near the termini (A) and with the donor aminomethylcoumarin (AMCA) and the acceptor dinitrophenyl (DNP) (B) in buffers containing 1–8 M GdmCl (colors from red to purple). The residuals of the global analysis of the two different FRET pairs under identical buffer conditions are shown; each color corresponds to one global fit.

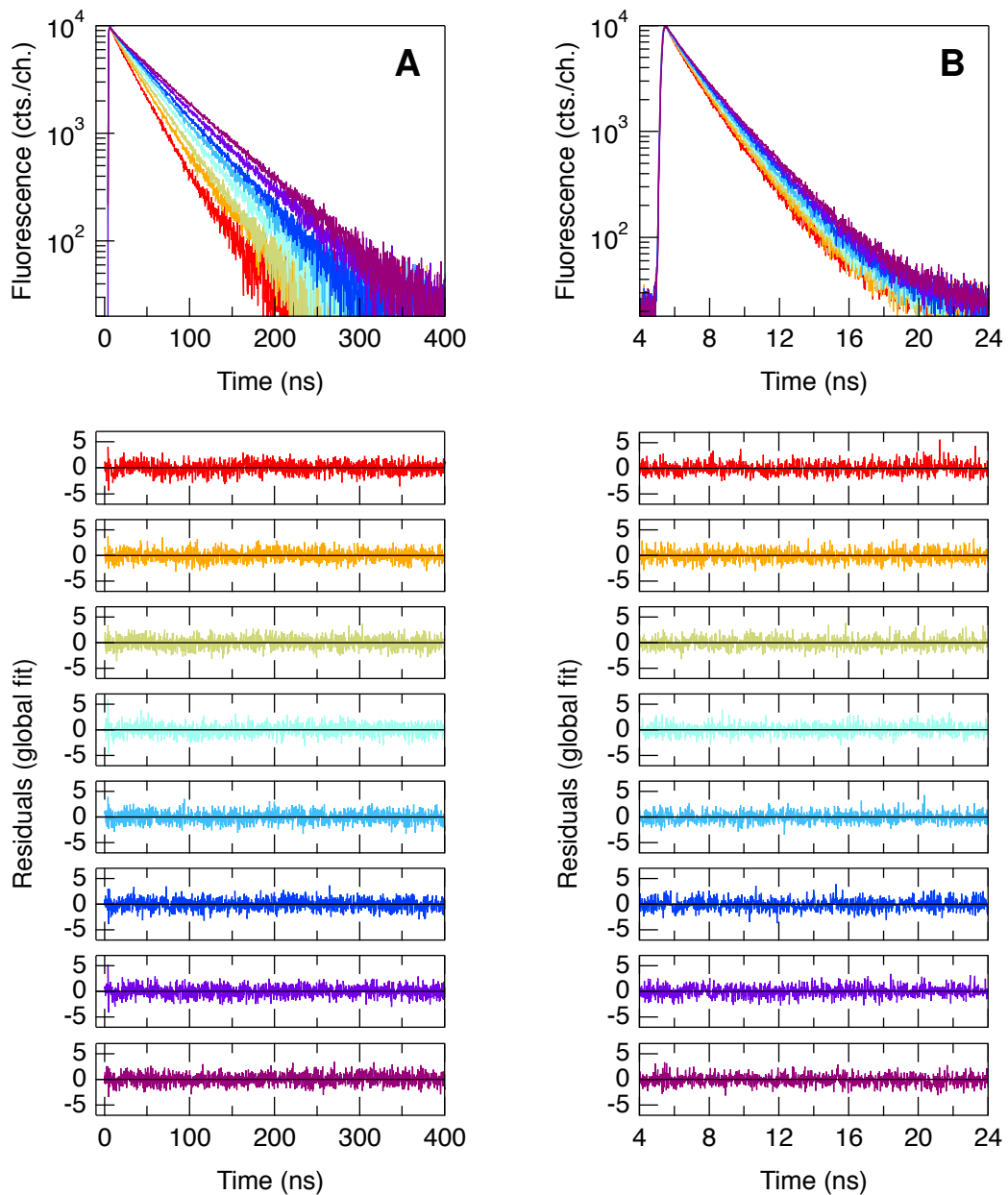


Figure 3.20.: Time resolved fluorescence decays for the PV(EF) peptide labeled with the donor pyrene (Pyr) and the acceptor dansyl (Dans) near the termini (A) and with the donor amino-methylcoumarin (AMCA) and the acceptor dinitrophenyl (DNP) (B) in buffers containing 1–8M GdmCl (colors from red to purple). The residuals of the global analysis of the two different FRET pairs under identical buffer conditions are shown; each color corresponds to one global fit.

3. Results

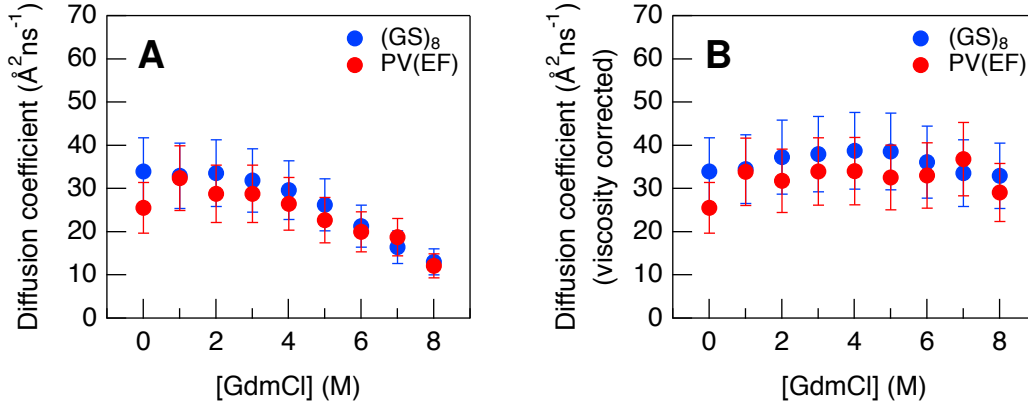


Figure 3.21.: Intra-chain diffusion coefficients of the two model peptides as obtained from global analysis (A) and corrected for the effect the changing of solvent viscosity (B).

corrected for effects of solvent viscosity according to the solvent viscosity dependence of intra-chain diffusion presented in chapter 3.3.

$$D_{\text{corr.}} = D \cdot \left(\frac{\eta}{\eta_0} \right)^\beta \quad (3.15)$$

In this equation, D is the experimental intra-chain diffusion coefficient, η the viscosity of the solvent, η_0 the viscosity of the reference solution (here 10 mM potassium phosphate, pH 7, 22.5°C) and β is the scaling exponent determined in section 3.3.3 (1.09 for (GS)₈ and 1.03 for PV(EF)). Within the given error, the viscosity corrected intra-chain diffusion coefficient remains almost unchanged when the denaturant concentration of the buffer is varied as shown in Figure 3.21 B. Similar to water, the intra-chain diffusion coefficient is on average slightly higher for the (GS)₈ peptide than for the native sequence PV(EF) at all GdmCl concentrations.

The root mean square end-to-end distance increases strongly by 21 % for GS₈ and 20 % for PV(EF) from 0–8 M GdmCl (see Figure 3.22 B). This effect is mostly due to a strong increase of the parameter b , which shifts the whole distribution towards larger end-to-end distances. The change of the parameter σ , which determines the width of the distribution, is small (Figure 3.22 A).

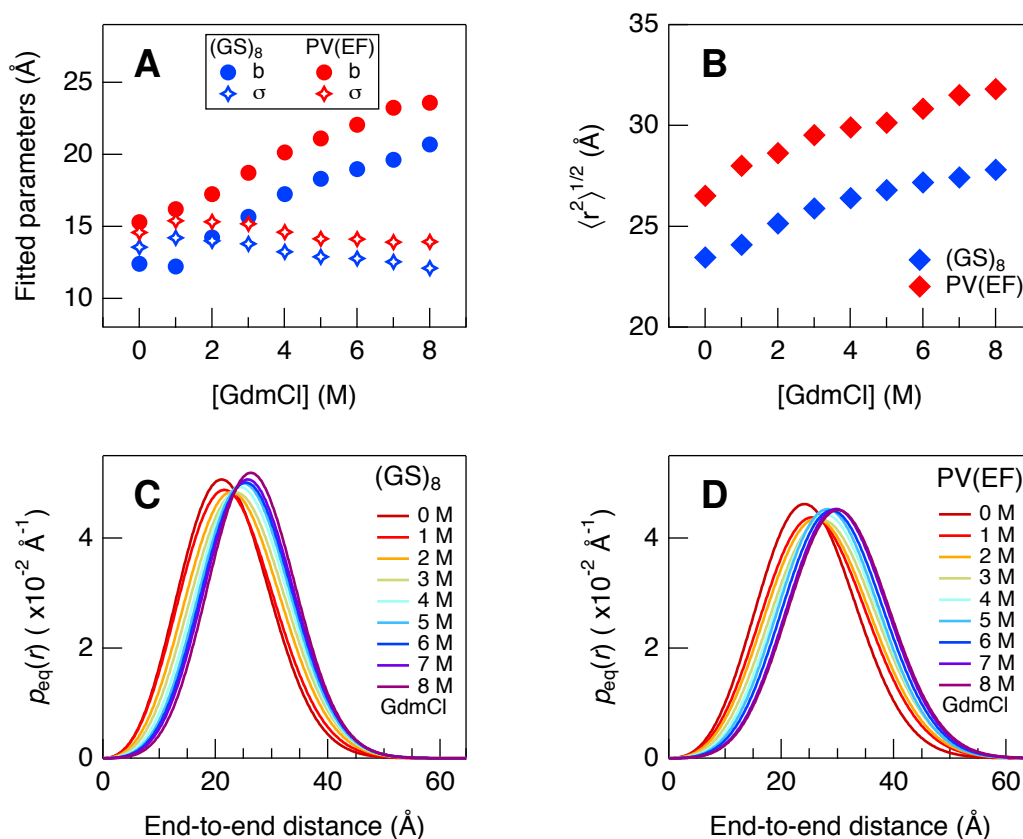


Figure 3.22.: A: The parameters b and σ obtained from the global fit determine the shape of the end-to-end distance distribution. The parameter b increases strongly with increasing denaturant concentration. B: The root mean square end-to-end distance increases strongly with denaturant. C and D: The influence of GdmCl on the end-to-end distance distributions of the $(GS)_8$ peptide and on the native loop sequence PV(EF) is very similar.

3.4.3. End-to-end distance distribution and intra-chain diffusion in the presence of sarcosine

The co-solute sarcosine, which occurs naturally for example in rays, is used to test for the influence of stabilizing osmolytes on intra-chain diffusion and to gain insight on the influence of conformational properties through the end-to-end distance distribution. Alike the procedure described in the previous section for GdmCl, time resolved FRET measurements were carried out in different sarcosine buffers ranging from 0–5 M sarcosine for the two model peptides $(GS)_8$ and PV(EF). The donor lifetimes τ and the Förster distances R_0 were determined for each buffer condition using donor-only and acceptor-only labeled peptides.

3. Results

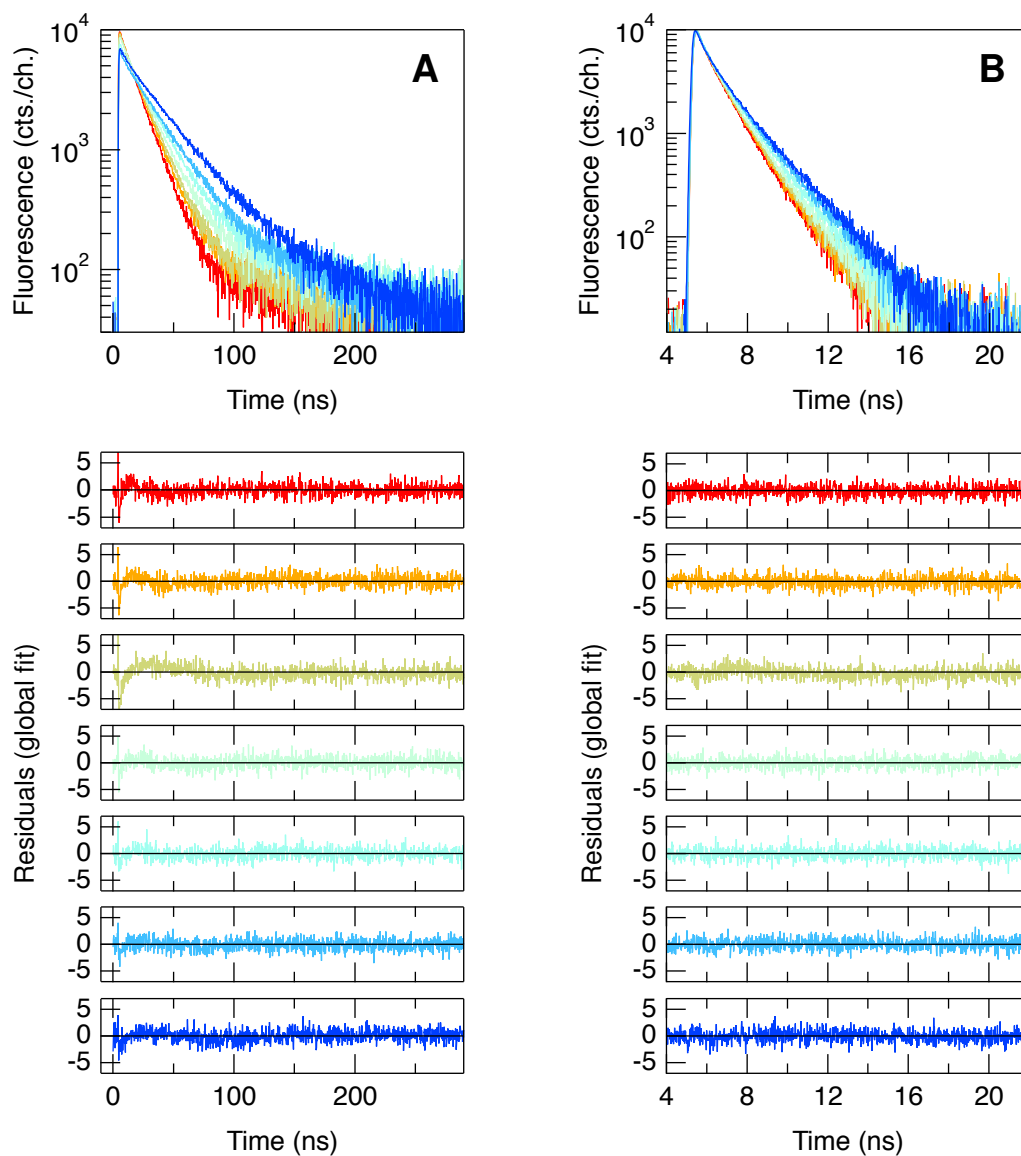


Figure 3.23.: Time resolved fluorescence decays for a (GS)₈ peptide labeled with the donor pyrene (Pyr) and the acceptor dansyl (Dans) near the termini (A) and with the donor aminomethylcoumarin (AMCA) and the acceptor dinitrophenyl (DNP) (B) in buffers containing 0.5 M, 1.0 M, 1.5 M, 2.0 M, 3.0 M, 4.0 M and 5.0 M sarcosine (colors from red to blue). The residuals of the global analysis of the two different FRET pairs under identical buffer conditions are shown; each color corresponds to one global fit. The decrease in the amplitude in A is due to an increase of background fluorescence of the buffer with the sarcosine concentration which has been subtracted from the data.

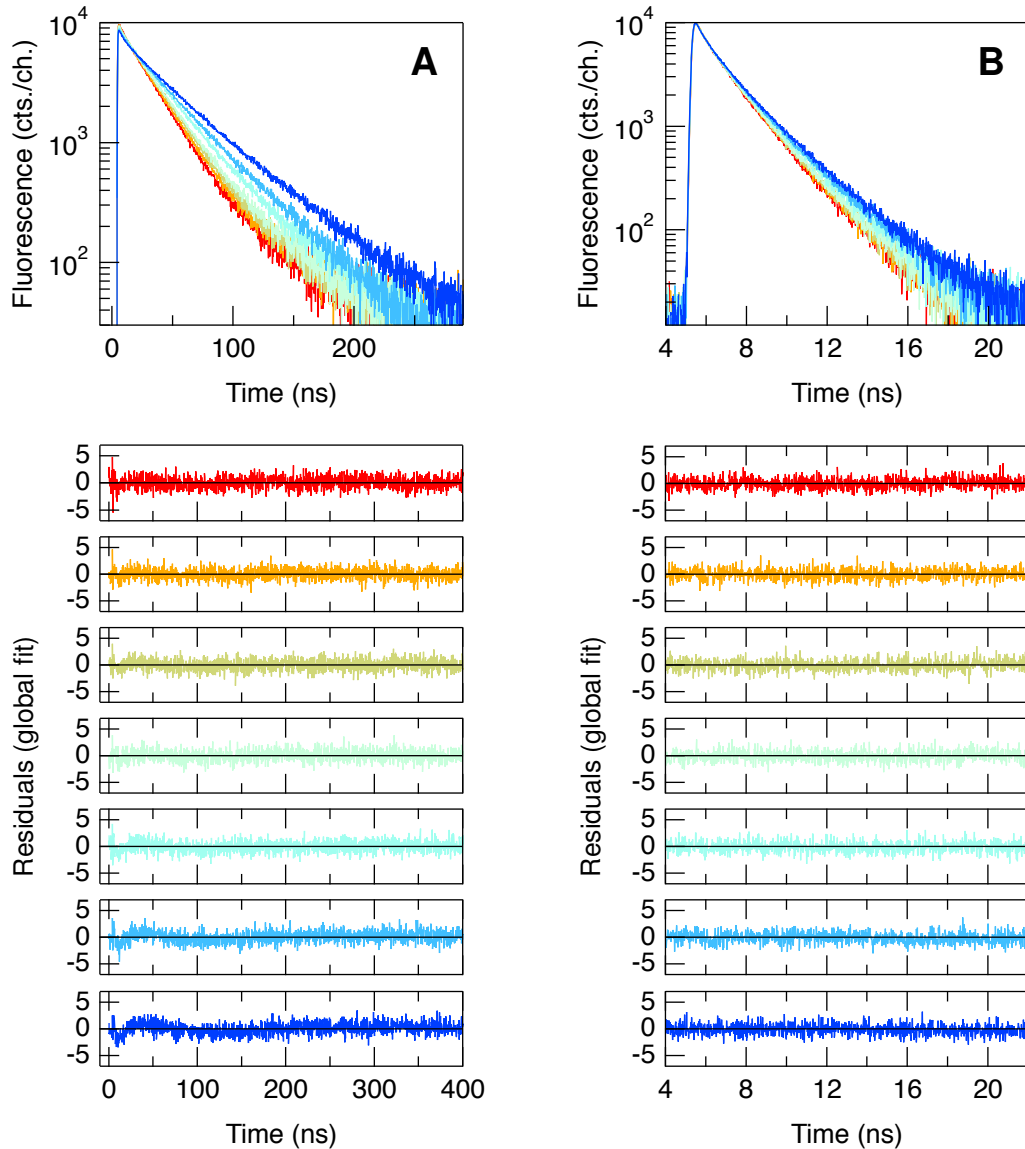


Figure 3.24.: Time resolved fluorescence decays for a PV(EF) peptide labeled with the donor pyrene (Pyr) and the acceptor dansyl (Dans) near the termini (A) and with the donor aminomethylcoumarin (AMCA) and the acceptor dinitrophenyl (DNP) (B) in buffers containing 0.5 M, 1.0 M, 1.5 M, 2.0 M, 3.0 M, 4.0 M and 5.0 M sarcosine (colors from red to blue). The residuals of the global analysis of the two different FRET pairs under identical buffer conditions are shown; each color corresponds to one global fit.

3. Results

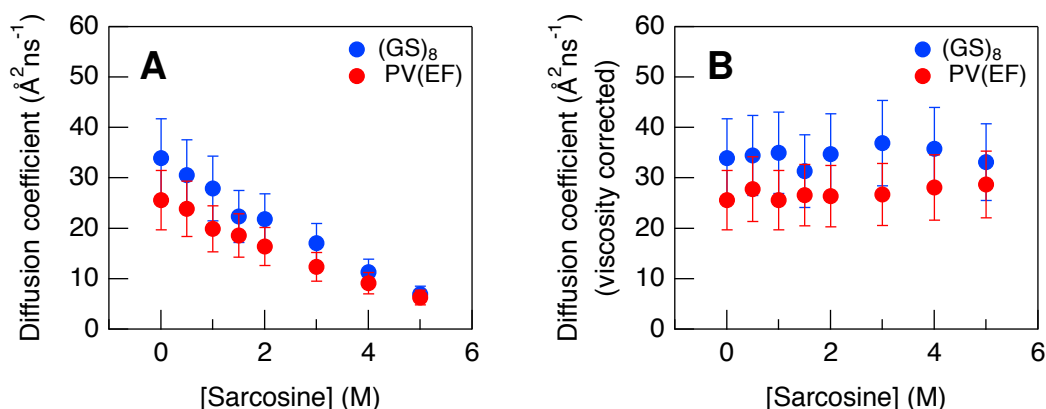


Figure 3.25.: Intra-chain diffusion (A) and viscosity corrected intra-chain diffusion (B) in the presence of the stabilizing osmolyte sarcosine.

The intra-chain diffusion coefficient and the equilibrium end-to-end distance distribution were obtained from global analysis of the two different FRET pairs as described.

The intra-chain diffusion coefficient D decreases with increasing sarcosine concentration. However, when correcting for the effects of altered solvent viscosity according to equation (3.15), intra-chain diffusion is found to be independent of the osmolyte concentration (see Figure 3.25 A and B).

Sarcosine seems to have an effect on both fitting parameters b and σ which determine the shape of the end-to-end distance distribution $p_{\text{eq}}(r)$ (Figure 3.26 A). However, the shape of the end-to-end distance distribution changes only very little (Figure 3.26 C and D). The same is true for the root-mean square end-to-end distance, which remains almost constant with increasing sarcosine concentration (Figure 3.26 B).

3.4.4. Monte Carlo simulations

To gain insight how the end-to-end distance distribution is related to the conformations of the backbone, i. e. the probability distribution of the dihedral angles, the MC simulations presented in chapter 3.2 are considered. This can help to better understand why denaturants and stabilizing osmolytes act on the end-to-end distance distribution as observed in the previous sections.

From the MC generated model ensembles, end-to-end distance dependent Ramachandran plots were calculated. Examples for these end-to-end distance specific Ramachandran plots

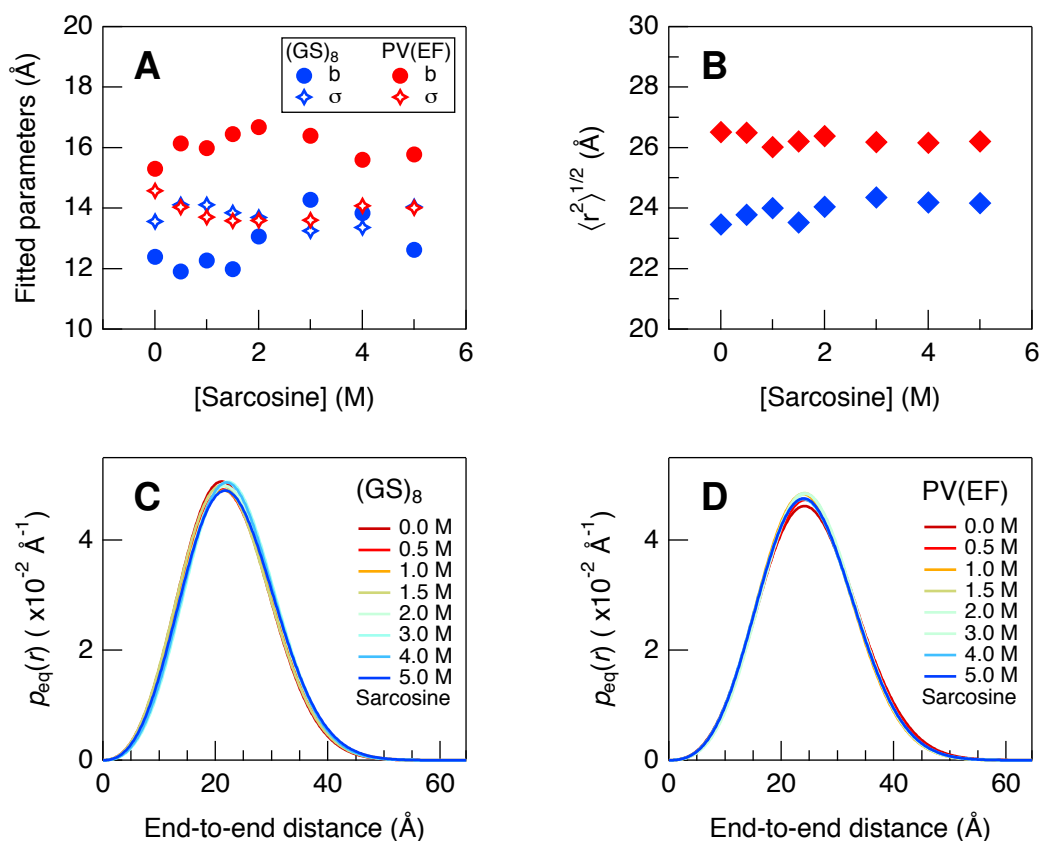


Figure 3.26.: Effect of sarcosine on the parameters b and σ of the end-to-end distance distribution (A), on the root mean square end-to-end distance (B) and on the end-to-end distance distributions (C and D).

are shown in the Figures 3.27 and 3.28 B, C and D. The obtained Ramachandran-plots differ significantly in the population of the major basins from the ones used as database for the MC sampling (compare Figure 3.10 in section 3.2.5). For each Ramachandran plot, the relative occurrences of ϕ - ψ -angles assigned to one of the major basins are calculated and are shown in Figure 3.27 A and 3.28 A. The native sequence PV(EF) has a higher probability of ϕ - ψ -angles assigned to poly-proline II and α -helical conformation than the $(GS)_8$ peptide. For both peptides, the probability of finding ϕ - ψ -angles corresponding to α -helical or left-handed helical conformations decreases with increasing end-to-end distance, while the probability of ϕ - ψ -angles corresponding to poly-proline II and β -strand conformations increases. However, at small end-to-end distances the population of all four major basins stays constant.

3. Results

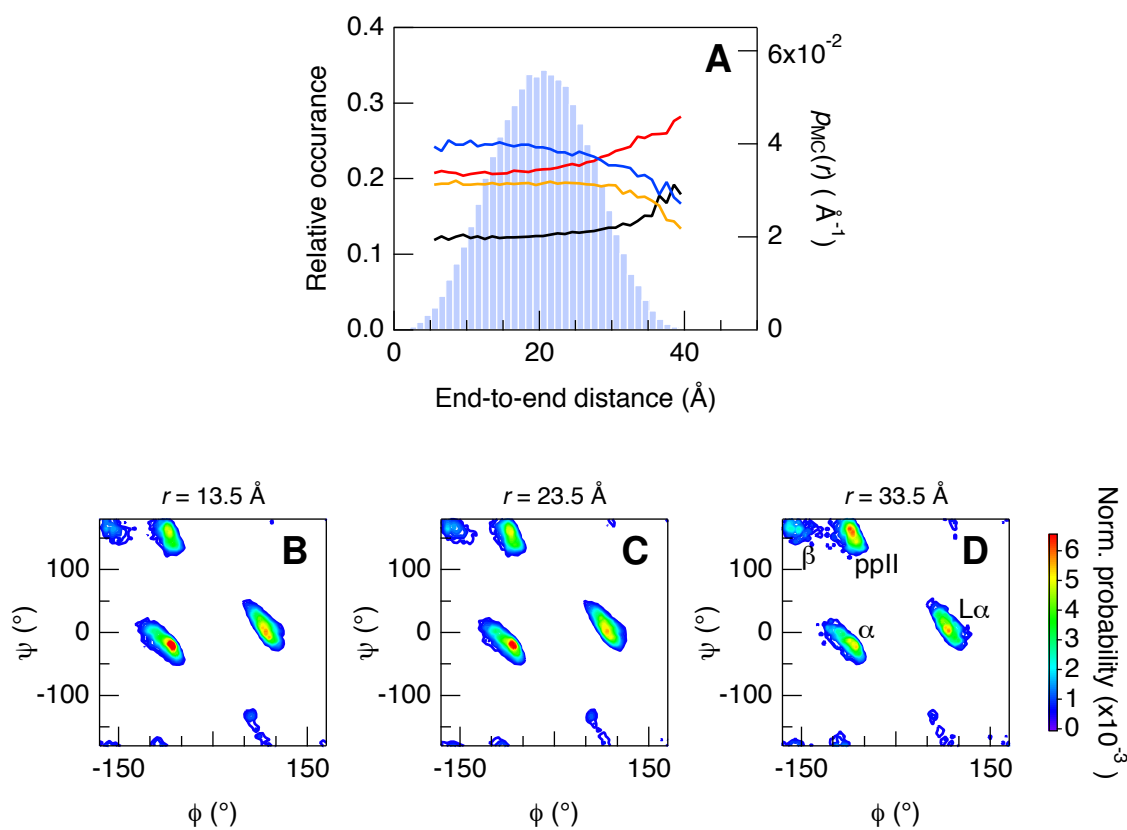


Figure 3.27.: Histogram of the probability to find a certain end-to-end distances in the Monte Carlo generated ensemble of conformations for the (GS)₈ peptide (A). For each end-to-end distance a Ramachandran plot is generated (for example B, C, D) and the relative occurrence of ϕ - ψ -angles which can be assigned to the four basins indicated in D is calculated. With increasing end-to-end distance, the relative occurrence of α -helical (—) and left-handed helical (—) ϕ - ψ -angles decreases, while the relative population of ϕ - ψ -angles indicating β -strand (—) and poly-proline II conformation (—) increases (A).

The changing probability of finding certain ϕ - ψ -angles along the end-to-end distance can also lead to differences in intra-chain interactions in the polypeptide. The numbers of intra-chain backbone H-bonds for each conformer can be calculated from the MC ensembles, by calculating the distances between backbone oxygen atoms and nitrogen bound hydrogens. A hydrogen bond was assumed to be formed, if the hydrogen-oxygen distance is below 2.76 Å while the angle formed by the oxygen-, hydrogen- and nitrogen atoms is larger than 120°. Accordingly, the average number of H-bonds per conformer for each end-to-end distance was calculated and is shown in Figure 3.29. The average number of H-bonds decreases with increasing end-to-end distance for both peptides. However, even though the (GS)₈ ensemble

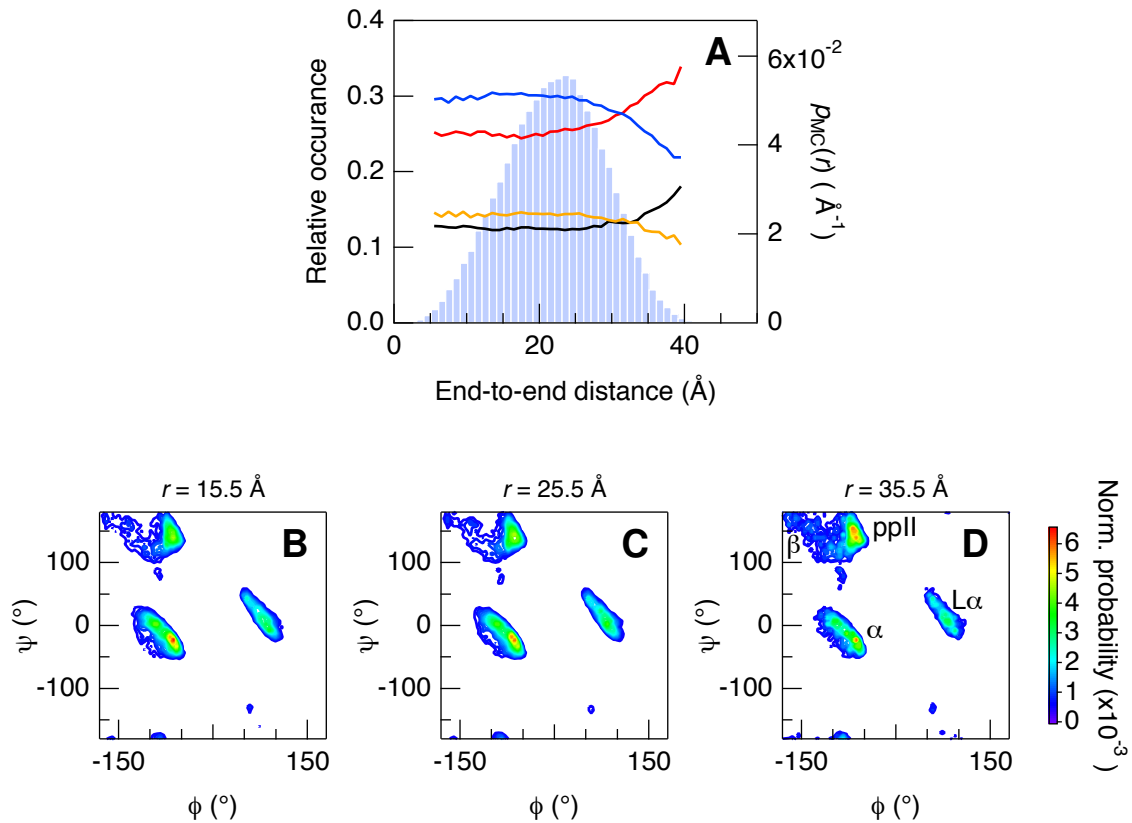


Figure 3.28.: Histogram of the probability to find a certain end-to-end distances in the Monte Carlo generated ensemble of conformations for the PV(EF) peptide (A). For each end-to-end distance a Ramachandran plot is generated (for example B, C, D) and the relative occurrence of ϕ - ψ -angles which can be assigned to the four basins indicated in D is calculated. With increasing end-to-end distance, the relative occurrence of α -helical (—) and left-handed helical (—) ϕ - ψ -angles decreases, while the relative population of ϕ - ψ -angles indicating β -strand (—) and poly-proline II conformation (—) increases (A).

has on average smaller end-to-end distances, the average number of H-bonds is lower than for the native sequence PV(EF).

These observations indicate, that an increasing end-to-end distance is related to an increase of the content of dihedral angles in ppII- and β -conformation and to a decrease in the number of backbone H-bonds.

Another critical question concerning the end-to-end distance distribution is, whether the latter is always affected by a change of the sampling of ϕ - ψ -angles. It has been shown, that an increase in the end-to-end distance is related to an increase in ppII and β sampling and to a decrease in the sampling of right- and left-handed helical conformations of ϕ - ψ -angles in

3. Results

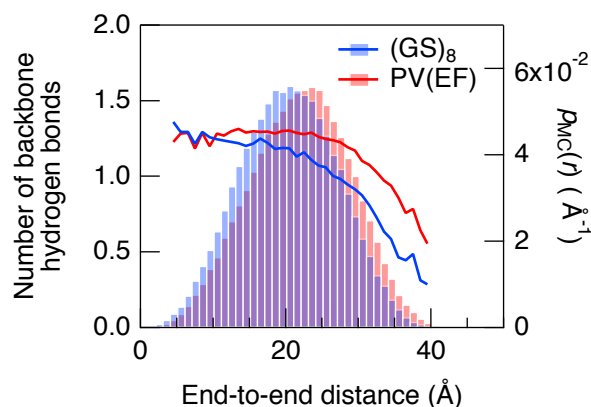


Figure 3.29.: The average number of backbone hydrogen bonds per conformer with a certain end-to-end distance. As backbone hydrogen bond all backbone oxygen atoms that showed a distance to a nitrogen bound hydrogen atom smaller than 2.76 Å and with a bonding angle above 120° were counted. For both peptides, the number of hydrogen bonds decreases with increasing end-to-end distance.

the unfolded ensemble. However, what would be the effect of an increase of helical conformations? In order to answer this question, further MC simulations were performed.

The starting Ramachandran plots for the MC simulation of glycine and serine were modified by systematically reducing the numbers of ϕ - ψ -angle pairs in the β -, ppII- and left-handed α basin thus increasing the probability of sampling conformations from the α basin. Based on these Ramachandran plots 100,000 (GS)₈ conformers were calculated and analyzed as described above. As expected, the probability of finding ϕ - ψ -angles belonging to the right-handed α -helical basin increased, but surprisingly also the probability of left-handed helical conformations increased compared to the original simulation (see Figure 3.30 compared to Figure 3.27 A). However, the end-to-end distance distribution of the ensemble as approximated by a histogram is remarkably similar to the original one (see Figure 3.31 A). In the same sense is the root mean square radius of gyration $\langle r_G^2 \rangle^{1/2}$ almost the same as before (Figure 3.31 B). The only significant difference between the two MC ensembles is the average number of backbone H-bonds formed within the chains.

In conclusion, it is possible to have a significant difference both in the sampling of α helical dihedral angles and the number of backbone H-bonds without significantly altering the end-to-end distance distribution and the root mean square radius of gyration.

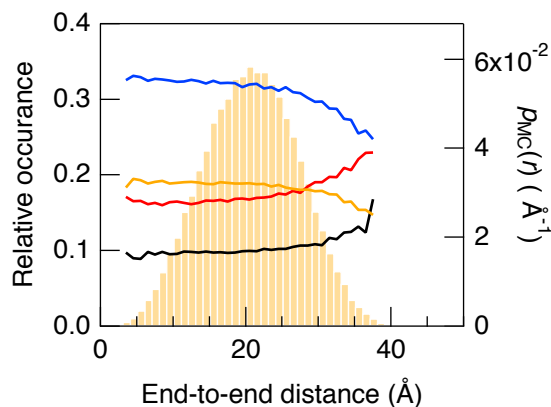


Figure 3.30.: Monte Carlo ensemble of $(GS)_8$ with increased probability of sampling α -helical conformations. Probability of ϕ - ψ -angles corresponding to one of the major basins of the Ramachandran plot: α -helical (---), left-handed helical (---), β -strand (---) and poly-proline II conformation (---).

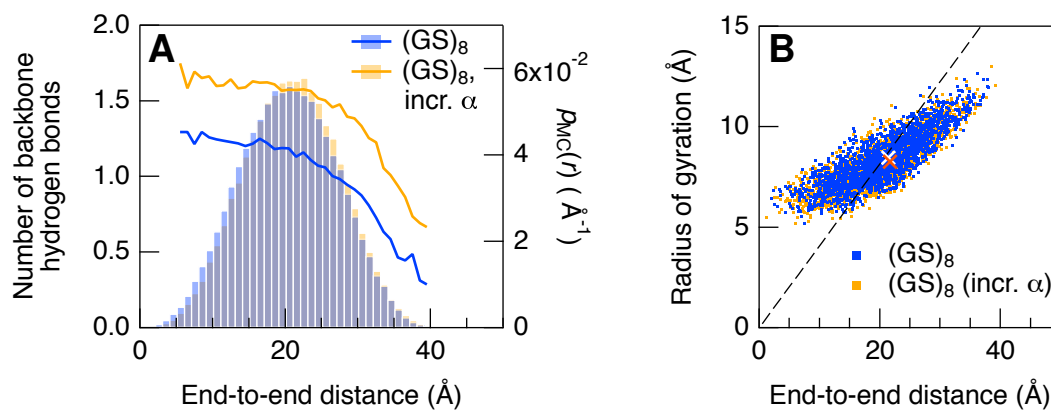


Figure 3.31.: A: Comparison of the end-to-end distance distributions of the Monte Carlo ensembles of $(GS)_8$. The ensemble with increases probability of sampling α -helical conformations has an increased number of backbone hydrogen bonds. B: Correlation between the radius of gyration r_G and the end-to-end distance. The relation $\langle r_G^2 \rangle = \langle r^2 \rangle / 6$ (equation (1.8)) of the random coil model is marked (---) and the average values $\langle r_G^2 \rangle^{1/2}$ versus $\langle r^2 \rangle^{1/2}$ for the original and the α increased simulations are indicated by white and orange crosses.

3. Results

3.4.5. Summary and discussions

Denaturants (GdmCl) and stabilizing osmolytes (sarcosine) have different effects on the unfolded state. GdmCl strongly increases the root mean squared end-to-end distance, while sarcosine has little effect on the latter. Neither of the two compounds changes the viscosity corrected intra-chain diffusion coefficient of the model peptides significantly. MC simulations of unfolded peptide conformations indicate, that an increasing end-to-end distance is related to preferred sampling of ppII and β -strand dihedral angle and to a decrease in the number of intra-chain backbone H-bonds. The simulations further point out the possibility of increasing the probability of sampling right- and left-handed helical ϕ - ψ -angles and the number of backbone H-bonds without changing the end-to-end distance distribution significantly.

The increase of the end-to-end distance with increasing GdmCl can be seen as a swelling of peptide dimensions as seen similar in many FRET studies.^{76,78,86,172} Even though some of these studies assume a random coil distribution of end-to-end distances and neglect the influence of intra-chain diffusion on FRET when deriving the distance distribution,^{76,78,86} this swelling of the end-to-end distance or the related radius of gyration of the random coil model (see equation (1.8)) is observed. The root mean squared end-to-end distance of unfolded protein L is found to increase by 20 % (2–7 M GdmCl)⁸⁶ or by 47 % (1–7 M GdmCl)⁷⁶, while for CspTm it is found to increase by either 38 % (1.2–7 M GdmCl)⁸⁶ or 80 % (0–7.2 M)⁷⁸. For the highly charged IDPs IN and ProT α , a collapse at low GdmCl concentrations is followed by an increase of the root mean square end-to-end distance of 14 % from 0.5–7 M GdmCl.⁷⁸ The neglect of intra-chain diffusion in these studies leads to an underestimation of end-to-end distances, especially when intra-chain diffusion is high as is the case at low GdmCl concentrations according to the results found here. Therefore, the expansion of end-to-end distances with GdmCl would be less pronounced than the given values, thus closer to the expansion of 20 % observed here. It has been reported, that the swelling of the denatured state was not observed in SAXS measurements,³¹⁵ even though in other studies an increase of the radius of gyration with denaturant is indeed observed³¹⁵.

The expansion of the peptides with increasing GdmCl observed in the present study indicates a higher content of ϕ - ψ -angles of ppII or β -strand conformation and a lower content of α -helical ϕ - ψ -angles in the unfolded ensemble as shown by the MC simulations. GdmCl has the same effect on both model peptides, indicating that the backbone is the major interaction partner for the denaturant, while side chains play a minor role. The solvation free

energy of peptides in water depends on the conformation of the backbone dihedral angles due to a difference in the hydrophobic solvent accessible surface area when H-bonds to the water molecules are taken into count in solvation.^{309,312} A similar concept may explain the favoring of extended conformations like ppII in the GdmCl unfolded state compared to the unfolded state in water. The favoring of ppII in the denatured state has been observed on many occasions^{310,311,316,317} in consistency with the observations in the present study. An increase in the end-to-end distance distribution further leads to a decrease in the number of backbone H-bonds as seen from the MC simulation in Figure 3.29. In conclusion, the denaturant GdmCl not only breaks native state interactions, but also disfavors the formation of backbone intra-chain H-bonds in the unfolded state.

Sarcosine, on the other hand, has been reported to acts on protein stability mainly by increasing the free energy of the unfolded state.³⁰⁷ The favorable interaction with polar and charged side-chains and the unfavorable interaction with the backbone³¹⁴ lead to preferential exclusion of stabilizing osmolytes from the proximity of the backbone³¹⁸. Denatured proteins or IDPs can react to an increase of the free energy of the unfolded state through a shift of the equilibrium, thus populating the native state. However, the two model peptides used here lack a well defined native structure and act as models systems to observe changes in the unfolded state.

Stabilizing osmolytes like sarcosine are believed to promote the formation of intra-chain H-bonds and favor compact conformations to decrease the SASA.^{313,314} However, no indication for a compaction of the model peptides with sarcosine is seen from the end-to-end distance distribution. Two explanations are possible for this observation - either, sarcosine has no effect on the unfolded state conformations, or, the osmolyte induced conformational changes have small or compensating effects on the end-to-end distance distribution.

Fitzkee et al showed, that the radii of gyration of proteins with 92 % of the amino acids in there native conformation and only 8 % flexible residues obey the same statistics as completely unfolded proteins.¹²⁰ This indicates, that the radius of gyration of an unfolded protein may not be affected by small changes in the secondary structure content, which could similarly also apply for the end-to-end distance distribution.

Holthauzen et al observed the CD signature at 228 nm to get more negative for the denatured state of Nank4-7* protein in the presence of sarcosine (sarcosine strongly absorbs at 220 nm usually used for these considerations).³¹⁴ Based on the theory of Uversy and Fink, a such change in the CD signature indicates a decrease of the hydrodynamic radius of the protein and thus a compaction.^{314,319} The CD signal at 228 nm is mostly due to α -helical

3. Results

conformations. Furthermore, stabilizing osmolyte was observed to induce helix formation in MD simulations of small peptides.³¹⁸

None of the two model peptides is very likely to entirely fold into helical secondary structure as the helix content of both model peptides is close to zero as predicted by the program agadir^{320–322}. However, sarcosine could increase the probability of the unfolded peptide to sample ϕ - ψ -angles from the α -basin thus leading to an increase in the number of backbone H-bonds without changing the end-to-end distance distribution significantly as suggested by the MC simulations and the experimental findings. The formation of backbone H-bonds leads to preferential exclusion of sarcosine from the proximity of the backbone and a shielding of backbone polar groups from the solvent. The loss of solvent interactions increases the free energy of the unfolded state and dominates over the decrease of free energy when backbone H-bonds are formed. Thus the free energy of the unfolded state increases as proposed³⁰⁷. These effects lead to the favoring of folding in the presence of stabilizing osmolytes.

The dynamics are for both, GdmCl and sarcosine, largely unaffected when the change of solvent viscosity is taken into account. Thus, the dynamics don't seem to be correlated with the number of intra-chain backbone hydrogen bonds, as neither a breakage nor a preferential formation of hydrogen bonds has any effect. This further indicates, that H-bonds are not the major cause of the friction f limiting intra-chain diffusion (equation 3.8 in chapter 3.3).

Up to date little is known about the effect of denaturants and stabilizing osmolytes on intra-chain diffusion. In one instance, the intra-chain diffusion coefficient in the presence of GdmCl was calculated from distance distributions determined by single molecule FRET⁷⁶ using Szabo-Schulten-Schulten-theory²³⁸ (see equation (1.19)).⁸¹ A value of $D = 4.2 \pm 0.4 \text{ \AA}^2\text{ns}^{-1}$ is found which is essentially insensitive to GdmCl concentrations ranging from 3–7 M. The viscosity corrected D_{corr} would thus increase with GdmCl. This value is much smaller than the value found in the present study and also the changes with GdmCl are not alike. This is probably due to the neglect of the effect of intra-chain on the FRET signals when determining the distance distribution in the single molecule experiment. Not only is the distance distribution then biased, but also are the afterwards obtained diffusion coefficients biased.

3.5. Temperature dependence of intra-chain diffusion and roughness of energy landscape

3.5.1. Diffusion in a rough potential

During folding, a protein has to undergo conformational rearrangements in the unfolded state on the search for favorable interactions. These conformational dynamics can macroscopically be seen as intra-chain diffusion taking place in the free energy landscape of the unfolded state. In the present study, the focus will be on the unfolded state aiming to answer the following questions:

- 1) What does the energy landscape of unfolded proteins look like?
- 2) What limits the dynamics in the unfolded state?

Many different conformations of the protein coexist in the unfolded state. Interconversion between these conformation constantly takes place at ambient temperature. A simultaneous population of many different conformations suggests that these conformations are energetically very similar resulting in a highly frustrated energy landscape. Interconversion between the different conformations occurs majorly unhindered. Consequently, energy barriers between the conformations are small compared to thermal energy and can easily be overcome. This results in an energy landscape of the unfolded state which can be well approximated by a rough potential, with minor barriers separating the many energetically similar sub-states. Diffusional processes in a such rough energy landscape have been proposed to depend on the root mean squared height of minor barriers ε and to show a characteristic temperature dependence.³²³

$$D = D^* \exp\left(-\frac{\varepsilon^2}{(k_B T)^2}\right) \quad (3.16)$$

In this equation, D is the intra-chain diffusion coefficient, D^* is the diffusion coefficient in absence of minor barriers and $k_B T$ is the Boltzmann constant times the absolute temperature. This theory has first been reported by Zwanzig³²³ and a year later by Bryngelson and Wolynes³²⁴. An equivalent expression has been obtained in terms of first passage time approach.³²⁵

The concept of a rough potential has been widely used to explain experimental results on many different protein^{85,88,89,326–330} and has also been observed in simulations^{331–334}. However, whenever equation 3.16 has been applied to intra-chain diffusion in unfolded proteins aiming to determine ε , relatively vague assumptions had to be made for the value of D^* as

3. Results

measurements have been carried out at one single temperature.^{85,88} In order to follow a more systematical approach in determining the roughness of the potential and D^* , the temperature dependence of the intra-chain diffusion coefficient in an unfolded polypeptide has to be determined. Using time-resolved FRET measurements of unfolded model peptides combined with global analysis of data obtained from different FRET pairs, the intra-chain diffusion coefficient is determined together with the end-to-end distance distribution as presented in chapter 3.1. The temperature dependence of the intra-chain diffusion coefficient of the two model peptides presented in chapter 3.2 is well described by equation 3.16 and can be used to determine the roughness of the potential ε and the intra-chain diffusion coefficient in absence of barriers D^* . This gives not only informations on the shape of the energy landscape in the unfolded state, but also allows for a better understanding of the processes limiting chain dynamics in the unfolded state.

3.5.2. End-to-end distance distribution and intra-chain diffusion coefficient at various temperatures

Time resolved FRET measurements were carried out in order to determine the intra-chain diffusion coefficient D and the end-to-end distance distribution $p_{\text{eq}}(r)$ for the two unfolded model peptides GS₈ and PV(EF) (section 3.2.1 and 3.2.2) at temperatures ranging from 2.5°C to 37.5°C. In order to accurately determine D and $p_{\text{eq}}(r)$ two different FRET-pairs were used for each peptide and the datasets were globally analyzed as described in chapter 3.1. Donor lifetimes τ , donor quantum yields and Förster distances R_0 were determined for each temperature using donor-only and acceptor-only labeled peptides.

The intra-chain diffusion coefficient is found to strongly increase with temperature for both model peptides as shown in Figure 3.34 A. According to the Einstein-Smoluchowski equation (1.17), D is proportional to the absolute temperature T . Therefore, D is expected to increase with temperature. Additionally, the friction coefficient f is expected to decrease due to the decrease in solvent viscosity with temperature, leading to an even stronger increase in D . Based on the determination of the viscosity dependence presented in chapter 3.3, a correction for the change of solvent viscosity η can be applied according to the following equation:

$$D_{\text{corr.}} = D \cdot \left(\frac{\eta}{\eta_0} \right)^\beta \quad (3.17)$$

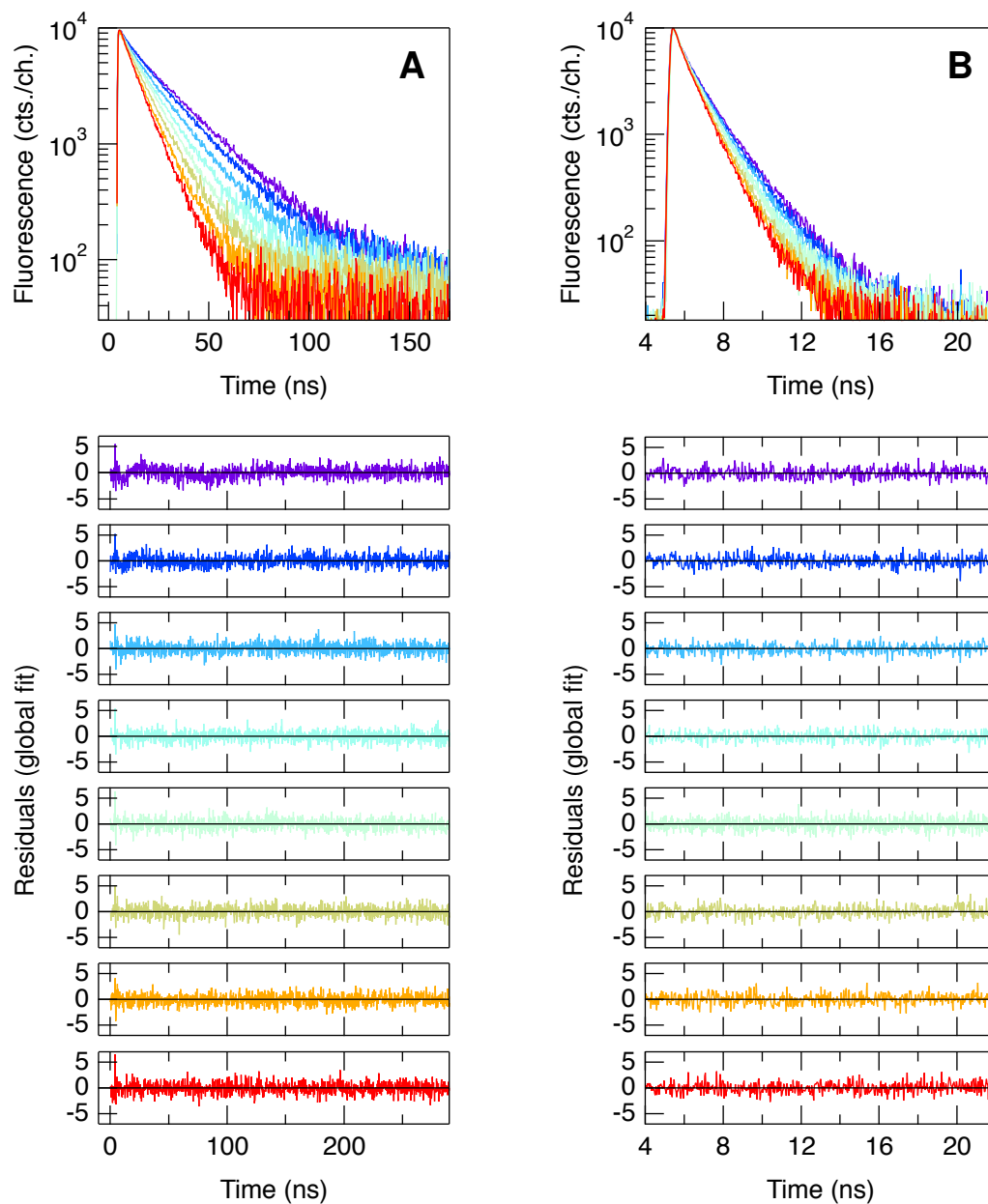


Figure 3.32.: Time resolved fluorescence decays for the (GS)₈ peptide labeled with the donor pyrene (Pyr) and the acceptor dansyl (Dans) near the termini (A) and with the donor amino-methylcoumarin (AMCA) and the acceptor dinitrophenyl (DNP) (B) in 10 mM potassium phosphate, pH 7 at temperatures ranging from 2.5°C to 37.5°C (colors from purple to red). The residuals of the global analysis of the two different FRET pairs are shown; each color corresponds to one global fit, thus to one temperature.

3. Results

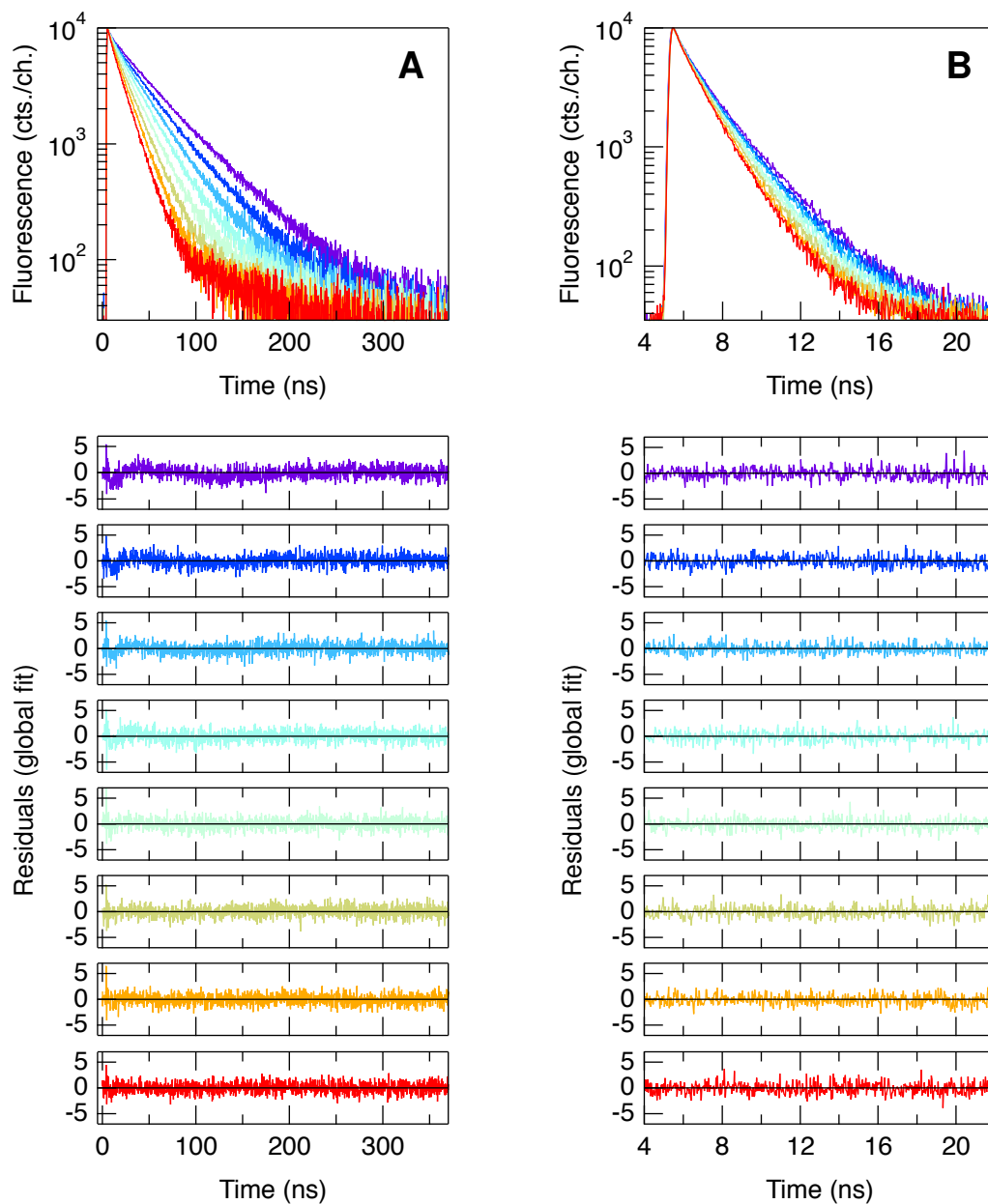


Figure 3.33.: Time resolved fluorescence decays for the PV(EF) peptide labeled with the donor pyrene (Pyr) and the acceptor dansyl (Dans) near the termini (A) and with the donor aminomethylcoumarin (AMCA) and the acceptor dinitrophenyl (DNP) (B) in 10 mM potassium phosphate, pH 7 at temperatures ranging from 2.5°C to 37.5°C (colors from purple to red). The residuals of the global analysis of the two different FRET pairs are shown; each color corresponds to one global fit, thus to one temperature.

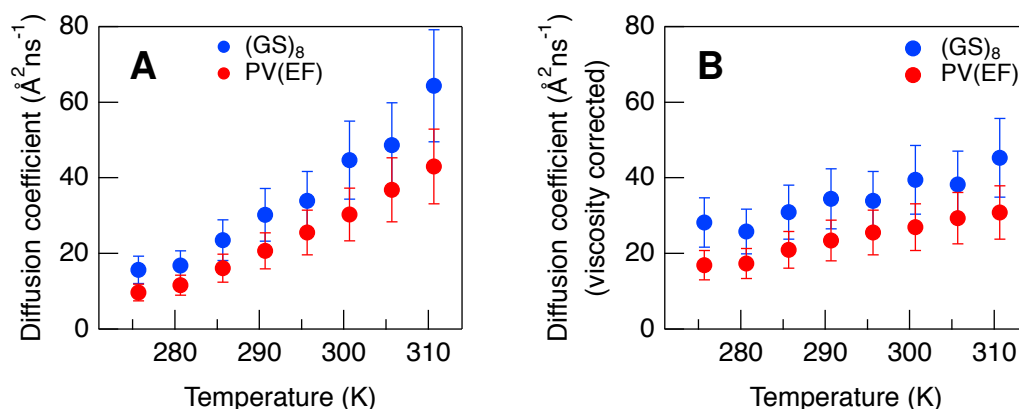


Figure 3.34.: Temperature dependence of the intra-chain diffusion coefficient (A) and the viscosity corrected intra-chain diffusion coefficient (B).

In this equation, D is the experimental intra-chain diffusion coefficient, η the viscosity of the solvent, η_0 the viscosity of the reference solution (here 10 mM potassium phosphate, pH 7, 22.5°C) and β is the scaling exponent determined in section 3.3.3 (1.09 for (GS)₈ and 1.03 for PV(EF)). The viscosity corrected intra-chain diffusion coefficients still show an increase with temperature, so the increase is not entirely due to the decrease in viscosity as can be seen from Figure 3.34 B.

In order to test the applicability of equation (3.16) to intra-chain diffusion in unfolded polypeptides, a semi-logarithmic plot of the viscosity corrected intra-chain diffusion coefficient versus the square of the inverse temperature is made (see Figure 3.35). As can be seen from the linearity in the plot, that the Zwanzig model indeed applies, indicating that intra-chain diffusion can be well described as one-dimensional diffusion in a rough potential with minor energy barriers. Fitting the data with equation (3.16) results in a root-mean squared barrier height $\varepsilon = 0.86 \pm 0.06$ kcal/mol (3.6 ± 0.2 kJ/mol) for (GS)₈ and of $\varepsilon = 0.94 \pm 0.03$ kcal/mol (3.9 ± 0.1 kJ/mol) for PV(EF). The intra-chain diffusion coefficient in the absence of roughness adopts values of $D^* = 288 \pm 83$ Å² ns⁻¹ for (GS)₈ and $D^* = 317 \pm 48$ Å² ns⁻¹ for PV(EF), which are about a factor of 10 larger than intra-chain diffusion at room temperature. Taking into account the error of the fitted values, they appear to be almost identical.

The root mean square end-to-end distance for both peptides is unaffected by a change of temperature (see Figure 3.36). However, a slight systematic change is observed when the entire distribution of end-to-end distances is considered, which can be especially seen from the values of the fitting parameters b and σ describing the end-to-end distance distribution

3. Results

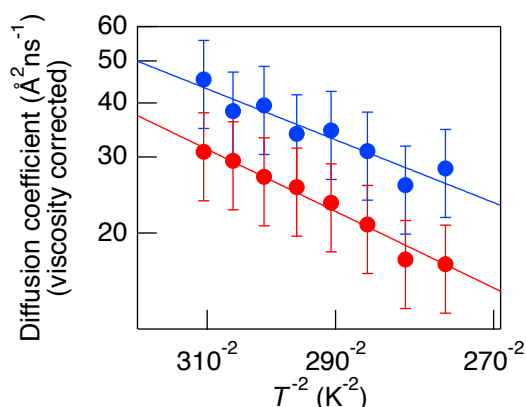


Figure 3.35.: Fits of the temperature dependence of the intra-chain diffusion coefficients with equation (3.16) for (GS)₈ (●) and PV(EF) (●). The resulting parameters are $\varepsilon = 0.86 \pm 0.06$ kcal/mol (3.6 ± 0.2 kJ/mol) and $D^* = 288 \pm 83 \text{ \AA}^2\text{ns}^{-1}$ for (GS)₈ as well as $\varepsilon = 0.94 \pm 0.03$ kcal/mol (3.9 ± 0.1 kJ/mol) and $D^* = 317 \pm 48 \text{ \AA}^2\text{ns}^{-1}$ for PV(EF)

according to equation (1.12). The equilibrium distance distribution broadens slightly with temperature. This effect is compensated by a shift of the parameter b towards smaller values (see Figure 3.36 A), so that the root mean square end-to-end distance remains unchanged.

3.5.3. Summary and discussions

The intra-chain diffusion coefficients of (GS)₈ and PV(EF) strongly increase with temperature. This increase is partly due to the decrease in solvent viscosity and partly caused by the rough free energy landscape of the unfolded state. The heights of the minor barriers of this rough potential are on the order of $1.5 \cdot k_B T$ at room temperature. The root mean square end-to-end distance is unaffected by temperature even though small systematic changes are observed in the end-to-end distance distribution.

No indication for a temperature induced collapse of the polypeptide chain is found in the present study as the root mean square end-to-end distance is unaffected by a change of temperature. However, the intra-chain diffusion coefficient strongly increases with temperature. Grupi and Haas have reported very similar results for the intra-chain diffusion coefficients and root mean square end-to-end distances of eight different fragments derived from the sequence of α -synuclein, however with much larger errors compared to the present study.¹⁷⁴ The intra-chain diffusion coefficient of seven of these fragments have been found to increase with temperature, while the root mean square end-to-end distances stay unaffected and the

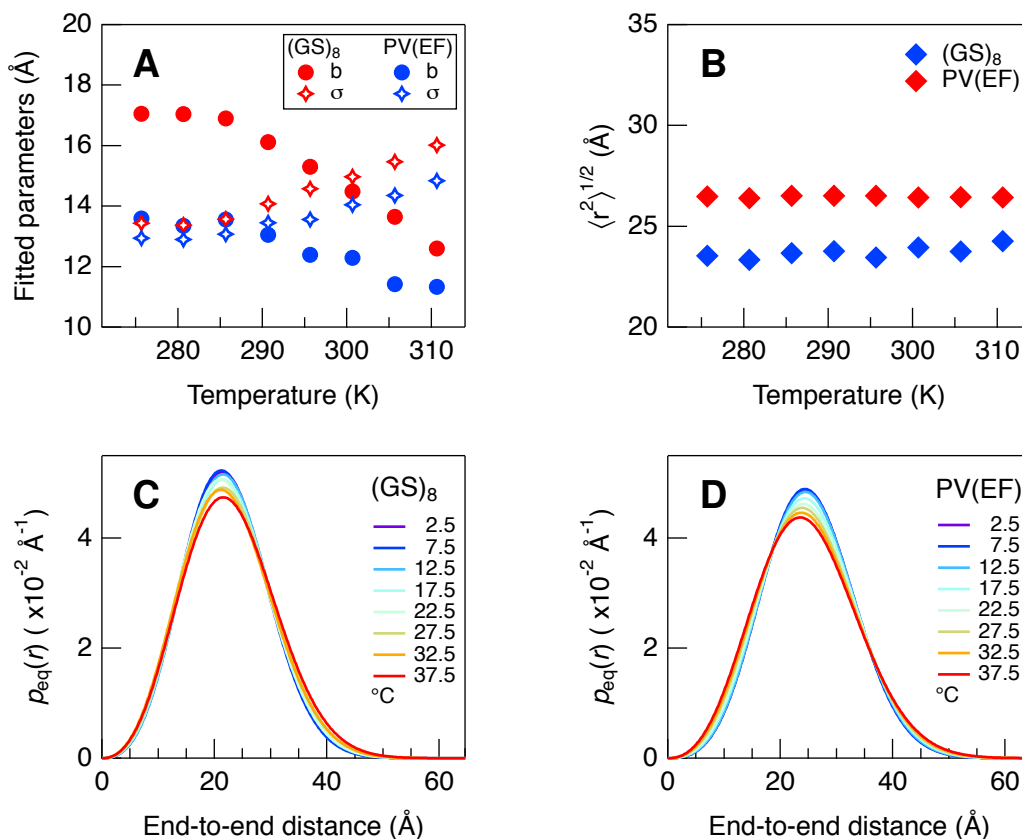


Figure 3.36.: Temperature dependence of end-to-end distance distribution. A: parameters b and σ describing the end-to-end distance distribution according to equation (1.12). B: root mean square end-to-end distance. C and D: end-to-end distance distribution $p_{eq}(r)$.

distributions slightly broaden.¹⁷⁴ However, the absolute values of the intra-chain diffusion coefficients of the present study range from 15.7–64.4 Å²ns⁻¹ for $(GS)_8$ and from 9.7–43.0 Å²ns⁻¹ for PV(EF) between 2.5°C and 37.5°C exceeding the values from 3–34 Å²ns⁻¹ between 5 and 40°C reported for the α -synuclein fragments¹⁷⁴. This may be due to the differences in sequence or to the lower accuracy of the reported values as only one single FRET-pair has been used.

The observed temperature-insensitive root mean square end-to-end distance is in contrast to a temperature induced collapse, which has been reported for acid denatured protein BBL from combined temperature jump and intensity based FRET experiments⁷⁹. In their study, Sadiq et al. used the FRET-pair Naph/Dans and interpreted a continuous increase in FRET efficiency as collapse of root mean square end-to-end distance of the protein. The findings

3. Results

of the present study, however, point out, that an increase in FRET efficiency with temperature is due to an increase of the intra-chain diffusion coefficient, which has been neglected by Sadiq et al. The influence of intra-chain diffusion on FRET efficiency is pronounced especially for long lifetime donors like Naph ($\tau = 38$ ns) as shown in Figure 3.1.

A temperature induced collapse of the proteins CspTm and ProT α in the presence of the denaturant GdmCl has also been reported.⁷⁷ Using single molecule FRET techniques, decreases between 13 % and 18 % in 0.5 and 7 M GdmCl have been found with temperatures increasing from 10–40°C. Intra-chain diffusion was not considered in the interpretation of the data leading to an over-interpretation of the collapse.

The intra-chain diffusion coefficient increases with temperature in agreement with theories like Zwanzig's diffusion in a rough potential (equation (3.16))³²³ or the Einstein-Smoluchowski equation (1.17)^{234,235}. However, Ahmad et al. observed a dramatic drop of the intra-chain diffusion coefficient with temperature based on the determination of tryptophan quenching rate constants for different loops of α -synuclein. This is in contrast to the findings of the present study, to findings of Grupi and Haas¹⁷⁴ and to theory.

The values found for the heights of the minor barrier in the rough free energy landscape of $\varepsilon = 0.86$ kcal/mol and $\varepsilon = 0.94$ kcal/mol ($1.45 \cdot k_B T$ and $1.58 \cdot k_B T$ at 25°C) are consistent with values reported in the literature.

Nettels et al. have determined a barrier height of $\varepsilon = 1.3 \cdot k_B T$ for viscosity corrected intra-chain diffusion of Csp⁸⁸ based on the assumption, that D^* equals the intra-chain diffusion coefficient in of Csp in 8 M GdmCl. In the present study, the values determined for D^* are about a factor of 10 higher than the values for the intra-chain diffusion coefficient in 8 M GdmCl of $D = 33 \text{ \AA}^2 \text{ ns}^{-1}$ und $D = 29 \text{ \AA}^2 \text{ ns}^{-1}$ for (GS)₈ and PV(EF) respectively (see chapter 3.4). Waldauer et al. have calculated the barrier height to be ≈ 1 kcal/mol for unfolded protein L based on quenching rate constants of tryptophan triplet states, a process which is not diffusion controlled as discussed in the introduction in section 1.4.1.⁸⁹ In a later study on the same protein a barrier height of $\varepsilon = 2.6 \cdot k_B T$ is reported,⁸⁹ which is however based on the assumption that D^* is the lateral diffusion coefficient of the protein in contradiction to the definition reported by Zwanzig³²³.

A value of $\varepsilon = 0.8 \cdot k_B T$ is reported by Yang and Gruebele for the Trpzp2 hairpin based on T-jump measurements³²⁸, which is smaller than the value determined in the present study.

In studies using methods with longer observation time windows, also higher barriers are found. A value of $\varepsilon = 4\text{--}5 \cdot k_B T$ was determined by Milanesi et al. for N-terminal domain of phosphoglycerate kinase using thiyl radical recombination dynamics³³⁰ and a value of

$\varepsilon = 5.7 \cdot k_B T$ was determined by Nevo et al. using atomic force microscopy^{326,327}.

The roughness of the free energy landscape was determined to have a root mean square barrier height of $\approx 1.5 \cdot k_B T$ or ≈ 0.9 kcal/mol. Minor barriers in the free energy landscape can originate from favorable or unfavorable interactions or from steric repulsion between different moieties of the polypeptide chain. The processes leading to the observed barrier heights can be identified by the magnitude of their energetic contributions.

Aromatic-aromatic interactions have been shown to contribute an energy between -1 and -2 kcal/mol⁷⁵ while surface exposed salt bridges have been shown to contribute between -0.21 and -1.25 kcal/mol to protein stability³³⁵. The energy of an intra-chain hydrogen bond has been reported to be between -1 kcal/mol in α -helices and -3 kcal/mol for buried bonds in water³³⁶. An additional source for barriers influencing the dynamics in the unfolded state is the bond rotation energy, because a change of the end-to-end distance is only possible through a change of the ϕ - ψ -angles. The barriers in bond rotation potentials have been proposed by Scheraga to be of the order of 0.5-1.3 kcal/mol³³⁷ or of the order of up to 3 kcal/mol as estimated by Ramachandran et al. from bond rotation energies in ethane³³⁸. A more recent study by Bacharov based on semi-empirical bond-bond interaction calculations, has suggested that the barriers hindering rotation are indeed much higher in the range of about 14–18 kcal/mol for N- C_α rotations and in the range of 5–7 kcal/mol for rotations of C_α -C³³⁹ in good agreement with other works based on ab initio quantum mechanical calculations (citations). The figures in reference³³⁹ however indicate the existence of smaller barriers between the major basin (poly-proline II, β -strand and α -helix) in the Ramachandran plot of glycine. These barriers are on the order of 0.5-1.5 kcal/mol and are absent for all other amino acids tested in reference³³⁹.

The dynamics in the (GS)₈-peptide can only be limited by hydrogen bonds or bond rotation energies as no charged or hydrophobic side chains are present between the FRET-labels. Taking into account the similarity between the root mean square barrier heights ε in both peptides studied, both these effects seem to be the major source of the barriers also in the PV(EF)-peptide. Both, the energies of intra-chain hydrogen bonds and bond rotation potentials are in quantitative excellent agreement with the values of $\varepsilon = 0.86$ kcal/mol in (GS)₈ and of $\varepsilon = 0.94$ kcal/mol in PV(EF).

In conclusion, both, bond rotation energies and hydrogen bonds, are possible origins for the rough energy landscape, which limits the dynamics in the unfolded state at physiological temperatures as tested by the two model peptides.

4. Summary

The aim of this study was to characterize conformational properties and the dynamics of unfolded polypeptides. To achieve this, an experimental approach based on FRET was employed, which yields the end-to-end distance distribution and the intra-chain diffusion coefficient of unfolded peptides with high accuracy. This method provided the basis for studies of the end-to-end distance distribution and the intra-chain diffusion coefficient under a variety of different solvent conditions.

Determination of the end-to-end distance distribution and the intra-chain diffusion coefficient by time-resolved FRET measurements

Time-resolved FRET measurements have been demonstrated to simultaneously yield the end-to-end distance distribution and the intra-chain diffusion coefficient of unfolded polypeptide chains. However, only a limited accuracy of the determined parameters could be achieved in previous works. In the present study, it could be shown, that high accuracy in the determination of the intra-chain diffusion coefficient and the end-to-end distance distribution is obtained, if FRET measurements involving donors with very long and short lifetime τ are globally analyzed. A comparative analysis including the properties of 3 different FRET pairs revealed, that the combination of AMCA ($\tau \approx 5$ ns) and Pyr ($\tau \approx 150$ ns) is the best choice for investigating unfolded peptides and that no significant improvement of parameter precision is obtained by including Naph ($\tau \approx 40$ ns) as a third donor.

Unfolded model peptides do not behave like random coils

A (Gly-Ser)-peptide and a native loop sequence from the protein β -carp parvalbumin were chosen as models for the unfolded state. Both peptides lack a folded structure in solution. The end-to-end distance distribution of both model peptides was determined by global analysis of time-resolved FRET measurements and compared to the distance distribution expected for the random coil model. For both model peptides, the resulting end-to-end distance distributions are well described by the distance distribution of the excluded volume

4. Summary

chain model. In contrast, the end-to-end distance distribution of a random coil could not be used to adequately describe the experimental data. Comparison of the experimental results with MC simulations indicates, that the shape of the distance distribution originates from limitations of backbone ϕ - ψ -angles due to steric clashes rather than from polymer-like characteristics of the polypeptide chain.

Internal friction not significantly contribute to intra-chain diffusion in the unfolded model peptides

The dynamics in unfolded polypeptides are limited by frictional forces, which can originate from the viscosity of the solvent (solvent friction) or from peptide intrinsic properties (internal friction). The coefficient of friction can be directly calculated from the intra-chain diffusion coefficient. Accordingly, the viscosity dependence of the intra-chain diffusion coefficient and the frictional coefficient were determined by time-resolved FRET measurements in the presence of the viscous co-solute glycerol. The intra-chain diffusion coefficient follows a non-ideal dependence $D \propto \eta^{-\beta}$ with β -values slightly above 1 for both model peptides as determined using glycerol as viscous co-solute. In order to separate solvent friction from internal friction, the viscosity dependence of the coefficient of friction was extrapolated to zero viscosity yielding the coefficient of internal friction. Values close to zero were found for the coefficient of internal friction for both, the (Gly-Ser)-peptide and the EF loop. Consequently, friction in the dynamics of unfolded polypeptides is dominated by the sequence-dependent coefficient of solvent friction as internal friction could not be detected.

The end-to-end distance distribution of unfolded peptides is strongly affected by GdmCl but not affected by sarcosine

The effect of the denaturant GdmCl and the stabilizing osmolyte sarcosine was determined to depend on the effect of the co-solutes on the unfolded state. The root mean square end-to-end distance strongly increases with increasing GdmCl concentration, while it remains unaffected by sarcosine. The viscosity corrected intra-chain diffusion is not significantly affected by both these co-solutes. MC simulations indicate, that an increase of the end-to-end distance may be related to an increased content of ppII and β -strand dihedral angles and to a decrease in the number of backbone H-bonds in the unfolded conformations. The simulations further indicate, that the end-to-end distance distribution is not significantly affected by an increased content of α -helical dihedral angles in the unfolded ensemble.

The energy landscape of the unfolded state can be approximated by a rough potential

To determine the shape of the energy landscape, the temperature dependence of the intra-chain diffusion coefficient obtained from time-resolved FRET measurements was investigated. The temperature dependence of the viscosity corrected intra-chain diffusion coefficients for both model peptides agrees well with the temperature dependence expected for a diffusional process in a rough potential. The root means square barrier heights in the rough energy landscape were determined to be of the order of $1.5 \cdot k_B T$ at room temperature and the intra-chain diffusion coefficient in absence of these barriers would be increased by a factor of 10 compared to intra-chain diffusion at room temperature. The barrier heights is similar in magnitude to barriers separating basins in the Ramachandran map of glycine or to the energies of H-bonds. The root mean square end-to-end distance is unaffected by temperature changes.

A. Materials and methods

All chemicals used throughout this work were purchased either from Carl Roth (Karlsruhe, Germany) or Sigma Aldrich (St. Louis, MO, USA), if not stated otherwise. Fmoc-protected amino acids were purchased from either Novabiochem (Hohenbrunn, Germany) or Iris Biotech (Marktredwitz, Germany). Resins for peptide synthesis were purchased from Rapp Polymers.

A.1. Peptide synthesis and purification

A.1.1. Peptide synthesis

All peptides used were synthesized by solid-phase peptide synthesis (SPPS) on a ABI 433 A peptide synthesizer with UV monitoring (Applied Biosystems, CA, USA) using standard Fluorenylmethoxycarbonyl (Fmoc) chemistry (Fastmoc). Both peptides were synthesized using R RAM resins (Rapp Polymers, Marktredwitz, Germany) with Fmoc deprotection by 20 % piperidine and activation by O-(benzotriazol-1-yl)-N,N,N',N'-tetramethyluronium hexafluorophosphate (HBTU). Amino acid side chains were protected by standard protecting groups according to the following sequences:

(GS)₈:

Gly, Ser(tBu), Gly, Ser(tBu), Gly-Ser(Me,Me pro), Gly, Ser(tBu),
Gly, Ser(tBu), Gly, Ser(tBu), Gly-Ser(Me,Me pro), Gly, Ser(tBu),
Dpr(Mtt), Ser(tBu), Arg(Pbf), Gly, Tentagel R RAM.

PV(EF):

Leu, Lys(Boc), Ala, Gly, Asp(OtBu), Ser(tBu), Asp(OtBu)-(Dmb)Gly,
Asp(OtBu)-(Dmb)Gly, Lys(Boc), Ile, Gly, Val, Asp(OtBu), Dpr(Mtt),
Ser(tBu), Gly, Tentagel R RAM.

For (GS)₈, several pseudo-prolines (Gly-Ser(Me,Me pro)) were used to prevent aggregation during synthesis. For PV(EF), Asp(OtBu)-(Dmb)Gly doublebuildingblocks were used to avoid aspartamide formation.

A.1.2. Labeling with chromophores

Chromophores were coupled on the resin either directly to the N-terminus or to the amine of α, β -Diaminopropionic acid (Dpr). Selective labeling of peptides with two different chromophores was achieved by using the selectively removable side chain protecting group 4-Methyltrityl (Mtt) for Dpr, which can be selectively removed using 3 % trifluoroacetic acid (TFA) and 5 % triethylsilan (TES) in dichloromethan (DCM). In case of donor-only or acceptor-only peptides, the second reactive site (either the N-terminus or the amine side chain of Dpr), were acetylated using acetic acid, (benzotriazol-1-yloxy)tripyrrolidino-phosphonium hexafluorophosphate (PyBOP) and N,N-Diisopropylethylamine (DIPEA) in Dimethylformamid (DMF). Acceptors (5-Dimethylaminonaphthalene-1-sulfonyl chloride (Dansyl Chloride) or 1-Fluoro-2,4-dinitrobenzene) were coupled to the N-terminus using DIPEA in DMF. Donors (7-Amino-4-methyl-3-coumarinylacetic acid (AMCA), 1-Pyrene-acetic acid or 1-Naphthylacetic acid) were coupled to the amine side chain of Dpr using PyBOP and 4-Methylmorpholine (NMM) in DMF.

A.1.3. Purification

Cleavage from the resin was obtained by 90 % TFA, 5 % TES and 5 % H₂O. After cleavage from the resin peptides were purified by reversed phase high-performance liquid chromatography (HPLC) using C8 preparative or C12 semi-preparative columns (Jupiter proteo, Phenomenex) with a water/acetonitril gradient. Samples were purified to ≥ 97 % for donor-acceptor-peptides and ≥ 95 % for donor- and acceptor-only peptides according to fluorescence monitoring by analytical HPLC; peptide identity was verified by matrix-assisted laser desorption/ionization time of flight mass spectrometry (MALDI-TOF-MS). Clean fractions were pooled, aliquoted and lyophilized.

A.2. Fluorescence measurements

Fluorescence spectra were recorded on a fluorolog τ spectrofluorometer (Jobin Yvon, Horiba, USA) or a fluoromax 3 spectrofluorometer (Jobin Yvon, Horiba, USA) both equipped with a correction file to obtain physical spectra provided by the manufacturer or on an ATF 105 spectrofluorometer (AVIV Biomedicals, NJ, USA) with a self written correction file obtained using a Spectral Fluorescence Standard Kit (certified by BAM, purchased from

Sigma Aldrich). Absorbance spectra were recorded on a double beam spectrophotometer (AVIV model 14 UV-VIS, AVIV Biomedicals, NJ, USA). The refractive indices were determined using an automatic refractometer (AR700, Reichert Inc. Depew, NY, USA).

Fluorescence lifetime measurements were carried out on a FluoTime200 Time-resolved spectrometer (PicoQuant, Berlin, Germany) with a PDL 800-D Picosecond pulsed diode laser driver unit, a PicoHarp TCSPC system and PMA-182 photo multiplier tubes (based on H5783 photosensor modules, Hamamatsu). The setup was equipped with polarizers and a cuvette holder (TLC 50, Quantum northwest, WA, USA) with Peltier temperature regulation and magnetic stirrer. Excitation of the fluorescence donor was achieved using a pulsed diode laser or a pulsed light-emitting diode (LED).

A.2.1. Sample preparation

All buffers were prepared using ultra pure bottled water (chromasolv plus, Sigma Aldrich) to keep residual fluorescence from contaminations as low as possible. GdmCl (AA-Grade, NIGU Chemie GmbH, Waldkraiburg, Germany) and glycerol (Merck), as well as all potassium and phosphate salts were used as received. Sarcosine (bioXtra, Sigma Aldrich) was further purified because a yellowish color of the solution was observed leading to very high residual fluorescence. A highly concentrated sarcosine solution of ≈ 6.5 M was prepared at room temperature, mixed with ultra pure activated charcoal (Norit, Sigma Aldrich), and filtered through syringe filters (anatotop, 0.2 μm (Whatman)) within the next hours. The pH and salt concentration of the so obtained clear solution was adjusted and the sarcosine concentration determined through the refractive index of the solution. GdmCl (10 mM potassium phosphate) and Sarcosine (10 mM potassium phosphate) buffers were first prepared at high concentrations and then diluted with 10 mM potassium phosphate buffer to match the desired co-solute concentration. Glycerol buffers were prepared with help of a balance to adjust the desired glycerol concentrations. Viscosities of the buffers were checked using a falling ball viscosimeter (Thermo Fischer Scientific). Buffers were aliquoted, frozen in liquid nitrogen and stored at -80°C prior to use.

For all measurements, lyophilized peptide was directly dissolved in the measuring buffer and in case of poor solubility subject to ultra-sonification for 10 min. Undissolved peptide was removed by centrifugation at 13000g for 10 min and use of the supernatant. This step was repeated two times. Filtering was observed to not be an adequate method to remove undissolved peptide, as washing out of unwanted fluorescent contaminations from almost

A. Materials and methods

Table A.1.: Fluorescence quantum yields of standard fluorophores.

name of standard	solvent	quantum yield	donor
9-anthracenecarbonitrile	methanol	0.80 ³⁴⁰	AMCA
1,4 bis(5-phenyloxazol-yl)benzene (POPOP)	cyclohexane	0.97 ³⁴⁰	Pyr,AMCA
9,10-Diphenylanthracene (DPA)	cyclohexane	1.00 ³⁴⁰	AMCA,Pyr
2,5-Diphenyloxazole (PPO)	cyclohexane	0.94 ³⁴⁰	Naph,Pyr
N-acetyl-L-tryptophanamide (NATA)	H ₂ O	0.14 ³⁴⁰	Naph
perylene	cyclohexane	0.94 ³⁴¹	AMCA
p-Terphenyl	cyclohexane	0.89 ³⁴⁰	Naph

all available filters was observed. The oxygen concentration of the buffer was observed to critically influence fluorescence decay times. To keep the oxygen concentrations of the buffers reproducible, all samples were equilibrated at the desired temperature and then 10 times bubbled with air by help of a pipette. When a nitrogen flow was used to prevent water condensation at the outer surface of the cuvettes, cuvettes were sealed after the air bubbling and prior to exposure to the nitrogen flow.

A.2.2. Determination of fluorescence quantum yields and Förster-distances

The Förster distance depends on the fluorescence quantum yield and the spectral overlap between donor fluorescence and acceptor extinction (see equation (1.24)).

The fluorescence quantum yields were determined from the fluorescence spectra relative to the spectra of fluorescence standards of known quantum yield.

$$Q = Q_{st} \cdot \frac{F A_{st} n^2}{F_{st} A n_{st}^2} \quad (\text{A.1})$$

With the integrated fluorescence intensities F , the optical density at the excitation wavelength A and the refractive index n of the solvent. The index st refers to the properties of the solutions of the fluorescence standards. To increase accuracy, three different standards were used in each case (see Table A.1) and the resulting quantum yields were averaged. Concentrated solutions of standard fluorophores as well as of donor-only labeled peptides were prepared with an optical density between 0.2 AU and 0.4 AU. After determination

of the exact optical density at the desired wavelength λ_{ex} from the absorbance spectrum recorded with the desired bandwidth $\Delta\lambda_{\text{ex}}$ the solutions were diluted to match the target optical density A or A_{st} in the range from 0.01–0.03 AU. In case of volatile solvents (for standard fluorophores) dilutions were made by help of a balance and measurements were carried out in cuvettes with a teflon plug.

The acceptor extinction spectra were determined by dissolving acceptor only labeled peptide in the corresponding buffer with a maximum optical density in the desired wavelength range of ≈ 0.2 – 0.3 AU. Extinction spectra $\varepsilon_{\text{A}}(\lambda)$ were obtained from the absorbance spectra by using the extinction coefficients ε_{A} from the literature ($\varepsilon_{\text{A}}^{\text{max}} = 4300 \text{ M}^{-1}\text{cm}^{-1}$ ³⁴² for Dans³⁴² and $\varepsilon_{\text{A}}^{\text{max}} = 15,900 \text{ M}^{-1}\text{cm}^{-1}$ for DNP³⁴³). The overlap integrals of donor fluorescence with acceptor extinction were calculated according to the following equation:

$$J = \frac{\int_0^{\infty} F_{\text{D}}(\lambda)\varepsilon_{\text{A}}(\lambda)\lambda^4 d\lambda}{\int_0^{\infty} F_{\text{D}}(\lambda)d\lambda} \quad (\text{A.2})$$

To calculate the Förster distances R_0 using equation (1.24), the overlap integrals J , the determined donor quantum yields Q and $\kappa^2 = 2/3$ were used. The relative error of the R_0 was estimated to be 2.5 %, based on the an error of 10 % for Q and ε_{A} .

A.2.3. Fluorescence lifetime measurements

Fluorescence lifetime measurements were carried out to determine the donor lifetimes τ of the donor only labeled peptides and for donor-acceptor labeled peptides to obtain the data

Table A.2.: Settings for fluorescence lifetime measurements.

	AMCA	AMCA/DNP	Pyr	Pyr/Dans	Naph	Naph/Dans
light source	diode laser	diode laser	LED	LED	LED	LED
puls width	44 ps	44 ps	500 ps	500 ps	635 ps	635 ps
λ_{ex}	372 nm	372 nm	298 nm	298 nm	281 nm	281 nm
$\Delta\lambda_{\text{ex}}$	1.1 nm	1.1 nm	14 nm	14 nm	12.5 nm	12.5 nm
λ_{em}	445 nm	445 nm	398 nm	398 nm	348 nm	348 nm
$\Delta\lambda_{\text{em}}$	4 nm	4 nm	16 nm	16 nm	16 nm	16 nm
time-res.	64 ps	16 ps	256 ps	256 ps	128 ps	64 ps
filters	LP398	LP398	LP345	LP345	LP296	LP296

A. Materials and methods

for global analysis.

The experimental settings are summarized in Table A.2. Sample concentrations were adjusted to a maximum optical density of 0.1 AU at the excitation wavelength. Polarizers were set to magic angle conditions. Samples were measured until the peak counts/channel reached a value of 10,000 counts. The fluorescence of the corresponding buffer under the same conditions was then measured for the same amount of time. The instrumental response function (IRF) was collected at the emission wavelength identical to the excitation wavelength by light scattering from dilute Ludox (Sigma Aldrich) solutions up to a peak value between 20,000 and 40,000 counts/channel.

Fits of the donor-only fluorescence lifetime measurements were performed using the software FluoFit (PicoQuant, Berlin, Germany). The shape of the IRF was considered through a re-convolution fit of a sum of exponentials according to the following equation:

$$F(t) = \int_{-\infty}^t \text{IRF}(t') \sum_{i=1}^n A_i \exp\left(-\frac{t-t'}{\tau_i}\right) dt' \quad (\text{A.3})$$

with the amplitude A_i and the lifetime τ_i of the i th component in the fit. For all donor-only measurements, one single donor lifetime τ was determined.

A.3. Global data analysis

The routines used for data analysis were programmed using Matlab (The MathWorks) release R2011b and run on a Mac Pro.

A.3.1. Data input

The successful calculation of a fit depends on several inputs. Firstly, the time-resolved raw data of the lifetime measurements of the double labeled peptides, the raw data of the fluorescence from the buffer and the raw data from measuring the IRF need to be imported. Secondly, the properties of the chromophores, i.e. the donor lifetime τ and the Förster distance R_0 , and the properties of the peptide, i.e. the total length r_{\max} (number of amino acids times 3.8 Å) and the minimum distance between the chromophores ($r_{\min} = 0$ was used here), need to be fixed. At last, the starting parameters (D , b and σ) of the global fit need to

be set.

A.3.2. Solution of PDE

The solution of the partial differential equation (PDE) (see equation (3.1)) for the parameters D, b, σ, τ, R_0 is then estimated by help of the Matlab function `pdepe`, which numerically solves initial-boundary value problems for parabolic-elliptic PDEs. The solution is calculated for 100 distances r between r_{\min} and r_{\max} with reflective boundary conditions at the limits r_{\min} and r_{\max} . The initial condition was $\bar{p}^*(r, t = 0) = 1$, which corresponds to the assumption, that at time zero, all donor fluorophores are in the excited state. The solution of the PDE, $\bar{p}^*(r, t_i)$, is calculated for each time point t_i in the raw data. The donor fluorescence is obtained from the two-dimensional solution according to equation (3.3), which reduces the solution to a one-dimensional time trace $F_{D,b,\sigma}^{\text{norm.}}(t_i)$, which is normalized to a maximum value of 1.

A.3.3. Re-convolution fit

Due to the finite width of the excitation pulse used in TCSPC when collecting the raw data and due to the non-ideal shape of the pulse of a diode laser or LED, the experimental fluorescence decay traces depend on the shape of the IRF. To account for this non-ideality in the raw data, a so called re-convolution fit is used. In a re-convolution fit, the theoretical decay function is convoluted with the experimentally determined IRF to obtain the final fit function. In the present study, the convolution between two traces is calculated through fast Fourier transformation of both traces, the multiplication of the obtained results and the inverse fast Fourier transformation of the product. The final fit function depend on several more parameters, which are however of inferior importance as they account for the experimental settings and for side-effects imposed by the chromophores. The total fit function for the re-convolution fit is thus:

$$F_{\text{total}}^{\text{fit}}(t_i) = \text{IRF}(t_j - t_0) \otimes \{A_1 \cdot F_{D,b,\sigma}^{\text{norm.}}(t_j) + A_n \cdot \exp(\tau_n t_j) + A_{\text{dc}}\} \quad (\text{A.4})$$

With t_0 accounting for the position of the fluorescence signal in the time trace, the amplitude of the main phase A_1 , the difference in dark counts between the raw data and the time trace of the buffer fluorescence and additional phases with amplitudes A_n and time constants τ_n which can account for photobleaching of the acceptor if $\tau_n = \tau$, for scattered light, or for

A. Materials and methods

stacking of the chromophores. As time base for this fit a linear spacing is chosen for the early time points, while for the later time points, a logarithmic spacing is achieved by successively discarding time points. This should give more reasonable weighting to the important early time points, which account for the D , b , σ -dependent fluorescence and remove weight from the later time points, which mainly account for the dark counts or the slow donor only phase of up to 2 % caused for example by acceptor bleaching. The total number of time points considered is 1000. In conclusion, the re-convolution fit determines the parameters t_0 , A_1 , A_n and A_{dc} , while not changing D , b and σ . In the optimization for the re-convolution fit, a reduced χ_R^2 is used.

$$\chi_R^2 = \frac{1}{n-p} \sum_{i=1}^{n=1000} \frac{(F_{\text{exp}}(t_i) - F_{\text{total}}^{\text{fit}}(t_i))^2}{(\Delta F_{\text{exp}}(t_i))^2} \quad (\text{A.5})$$

F_{exp} corresponds to the raw data of the double labeled peptide with the fluorescence of the buffer subtracted, and $\Delta F_{\text{exp}}(t_i)$ is the error of each the signal F_{exp} at the time point t_i which includes the error of the raw data point as well as the error of the buffer fluorescence signal. Therefore, values for χ_R^2 close to 1 are found for good fits.

A.3.4. Global fit

For each FRET pair in the analysis, the corresponding PDE was solved, and the corresponding re-convolution fit performed to obtain an individual χ_R^2 for each FRET pair. To obtain the global $\chi_{R, \text{global}}^2$, the arithmetic average of the χ_R^2 s for all FRET pairs were used.

The minimum of the value $\chi_{R, \text{global}}^2(D, b, \sigma)$ was then searched for by unconstrained non-linear optimization using Matlabs derivation free algorithm `fminsearch`. Each iteration in the minimization starts from a triple (D, b, σ) and solves the PDEs for all used FRET pairs followed by the corresponding re-convolution fits and thus $\chi_{R, \text{global}}^2(D, b, \sigma)$. The value of $\chi_{R, \text{global}}^2$ is then compared to the previous ones, and, based on the Nelder-Mead simplex direct search, a new guess for a triple (D', b', σ') is made. This is repeated until a minimum in $\chi_{R, \text{global}}^2(D, b, \sigma)$ is found.

A.3.5. Support plane analysis

As the algorithm `fminsearch` is not able to return errors of the fit parameters, support plane analysis is used to estimate the error. Firstly, a (global) fit is performed resulting in $D_{\text{fit}}, b_{\text{fit}}$ and σ_{fit} . Based on the obtained value for one of the fit parameters, secondary fits of the other two parameters are performed in order to calculate the projection of the $\chi_{R, \text{global}}^2(D, b, \sigma)$

hyper surface on one fit parameter. In case of the fitting parameter D for example, the projection is a function $\chi_{\text{R}}^2(D, b_{\text{fit}}(D), \sigma_{\text{fit}}(D))$. The confidence levels defining the lower and upper error estimates, were estimated by help of F-statistics according to the following equation:

$$\frac{\chi_{\text{R,global}}^2(D, b_{\text{fit}}(D), \sigma_{\text{fit}}(D))}{\chi_{\text{R,global}}^2(D_{\text{fit}}, b_{\text{fit}}, \sigma_{\text{fit}})} = F(p, \nu, P) \quad (\text{A.6})$$

with the the F-value F , p parameters, ν degrees of freedom and the probability $P = 0.67$. Values of ≈ 1.03 and ≈ 1.02 were obtained for the analysis of one single trace (1000 data points) and the global analysis of two traces (2×1000 data points) respectively.

List of Abbreviations

AMCA	7-amino-4-methyl-3-coumarin
CD	circular dichroism
Dans	dansyl
DCM	dichlormethan
DIPEA	N,N-Diisopropylethylamine
DMF	Dimethylformamid
DNP	2,4-dinitrophenyl
Dpr	α, β -Diaminopropionic acid
Fmoc	Fluorenylmethoxycarbonyl
FCS	fluorescence correlation spectroscopy
FRET	fluorescence resonance energy transfer
GdmCl	guanidinium chloride
Gly-Ser	poly-glycine-serine peptide
(GS)₈	(glycine-serine) ₈ peptide
H-bonds	hydrogen bonds
HBTU	O-(benzotriazol-1-yl)-N,N,N',N'-tetramethyluronium hexafluorophosphate
HPLC	high-performance liquid chromatography
IDP	intrinsically disordered protein
IRF	instrumental response function
LED	light-emitting diode
MALDI-TOF-MS	matrix-assisted laser desorption/ionization time of flight mass spectrometry
MC	Monte Carlo
MD	molecular dynamics
Naph	naphthalene
NMM	4-Methylmorpholine
NMR	nuclear magnetic resonance

List of Abbreviations

Mtt	4-Methyltrityl
PDE	partial differential equation
ppII	polyproline II
PV(EF)	EF-loop from the protein β -carp parvalbumin
PyBOP	(benzotriazol-1-yloxy)tripyrrolidinophosphonium hexafluorophosphate
Pyr	pyrene
SASA	solvent accessible surface area
SAXS	small angle X-ray scattering
sm	single molecule
SPPS	solid-phase peptide synthesis
TCSPC	time-correlated single photon counting
TES	triethylsilan
TFA	trifluoroacetic acid
TTET	triplet-triplet energy transfer

List of Figures

1.1	Ramachandran plot	2
1.2	Random coil model	8
1.3	Wormlike chain and excluded volume model	12
1.4	Protein dynamics	14
1.5	Energy landscape of protein dynamics	16
1.6	Jablonski diagram	20
1.7	Effect of broad distance distribution on FRET	23
3.1	Effect of intra-chain diffusion on FRET efficiency	32
3.2	Effect of intra-chain diffusion on fluorescence lifetime	33
3.3	Support plane analysis	37
3.4	Raw data, global fit and residuals	38
3.5	Sensitivity analysis I	40
3.6	Sensitivity analysis II	41
3.7	Glycine-serine repeat peptide with FRET labels	46
3.8	EF-loop from β -carp Parvalbumin with FRET labels	46
3.9	End-to-end distance distribution	47
3.10	Starting Ramachandran plots for simulation	50
3.11	End-to-end distance distribution obtained from MC simulations	51
3.12	Radius of gyration versus end-to-end distance	52
3.13	Raw data and residuals of global fit for (GS) ₈ at various glycerol contents	58
3.14	Raw data and residuals of global fit for PV(EF) at various glycerol contents	59
3.15	Effect of glycerol on intra-chain diffusion	60
3.16	Effect of glycerol on the end-to-end distance distribution	61
3.17	Determination of internal friction	62
3.18	Tanforde's transfer energy model	69
3.19	Raw data and residuals of (GS) ₈ in GdmCl	70

List of Figures

3.20	Raw data and residuals of PV(EF) in GdmCl	71
3.21	Effect of GdmCl on intra-chain diffusion	72
3.22	Effect of GdmCl on end-to-end distance distributions	73
3.23	Raw data and residuals of (GS) ₈ in sarcosine	74
3.24	Raw data and residuals of PV(EF) in sarcosine	75
3.25	Intra-chain diffusion in the presence of sarcosine	76
3.26	End-to-end distance distribution in the presence of sarcosine	77
3.27	Ramachandran plots of MC simulated (GS) ₈	78
3.28	Ramachandran plots of MC simulated PV(EF)	79
3.29	Backbone hydrogen bonds	80
3.30	Probability distribution of ϕ - ψ -angles with increased α -helical content	81
3.31	Backbone hydrogen bonds and radius of gyration	81
3.32	Raw data and residuals of global fit for (GS) ₈ at various temperatures	87
3.33	Raw data and residuals of global fit for PV(EF) at various temperatures	88
3.34	Temperature dependence of the intra-chain diffusion coefficient	89
3.35	Determination of roughness of free energy landscape	90
3.36	Temperature dependence of end-to-end distance distribution	91

List of Tables

3.1	Overview of FRET chromophores	36
3.2	Results of global fit	39
3.3	Results global fit and Edwards model	49
A.1	Fluorescence quantum yields of standard fluorophores	102
A.2	Fluorescence lifetime measurements	103

Bibliography

- [1] Crick, F. H. (1968) The origin of the genetic code. *Journal of Molecular Biology*, **38**, 367–379. 1
- [2] Orgel, L. E. L. (1968) Evolution of the genetic apparatus. *Journal of Molecular Biology*, **38**, 381–393.
- [3] Ambrogelly, A., Palioura, S., and Söll, D. (2006) Natural expansion of the genetic code. *Nature Chemical Biology*, **3**, 29–35. 1
- [4] EMBL-EBI, Uniprot - current release statistics. <http://www.ebi.ac.uk/uniprot/TREMBLstats/>, accessed May 5, 2013. 1
- [5] Anfinsen, C. B. C. (1973) Principles that govern the folding of protein chains. *Science (New York, N.Y.)*, **181**, 223–230. 1, 3
- [6] Uversky, V. N. V. (2002) Natively unfolded proteins: a point where biology waits for physics. *Protein science : a publication of the Protein Society*, **11**, 739–756. 1
- [7] Fink, A. L. A. (2005) Natively unfolded proteins. *Current opinion in structural biology*, **15**, 7–7.
- [8] Tompa, P. (2002) Intrinsically unstructured proteins. *Trends in Biochemical Sciences*, **27**, 527–533. 1
- [9] Walter, S. and Buchner, J. (2002) Molecular chaperones - Cellular machines for protein folding. *Angewandte Chemie International Edition*, **41**, 1098–1113. 1
- [10] Lang, K. K., Schmid, F. X. F., and Fischer, G. G. (1987) Catalysis of protein folding by prolyl isomerase. *Nature*, **329**, 268–270. 1
- [11] Messens, J. and Collet, J.-F. (2006) Pathways of disulfide bond formation in *Escherichia coli*. *International Journal of Biochemistry & Cell Biology*, **38**, 1050–1062. 1
- [12] Cabrita, L. D., Dobson, C. M., and Christodoulou, J. (2010) Protein folding on the ribosome. *Current opinion in structural biology*, **20**, 13–13. 1
- [13] Knight, A. M., Culviner, P. H., Kurt-Yilmaz, N., Zou, T., Ozkan, S. B., and Cavagnero, S. (2013) Electrostatic Effect of the Ribosomal Surface on Nascent Polypeptide Dynamics. *ACS Chemical Biology*, pp. –. 1

- [14] Dobson, C. M. C. (2003) Protein folding and misfolding. *Nature*, **426**, 884–890. 1
- [15] Ramachandran, G. N., Ramakrishnan, C., and Sasisekharan, V. V. (1963) Stereochemistry of polypeptide chain configurations. *Journal of Molecular Biology*, **7**, 95–99. 2
- [16] MacArthur, M. W. M. and Thornton, J. M. J. (1991) Influence of proline residues on protein conformation. *Journal of Molecular Biology*, **218**, 397–412. 2
- [17] Scherer, G., Kramer, M. L., Schutkowski, M., Reimer, U., and Fischer, G. (1998) Barriers to rotation of secondary amide peptide bonds. *J. Am. Chem. Soc.*, **120**, 5568–5574. 2
- [18] Stewart, D. E., Sarkar, A., and Wampler, J. E. (1990) Occurrence and role of cis peptide bonds in protein structures. *Journal of Molecular Biology*, **214**, 253–260. 2
- [19] Jabs, A. A., Weiss, M. S. M., and Hilgenfeld, R. R. (1999) Non-proline Cis peptide bonds in proteins. *Journal of Molecular Biology*, **286**, 14–14. 2
- [20] Avbelj, F. and Baldwin, R. L. (2004) Origin of the neighboring residue effect on peptide backbone conformation. *Proceedings of the National Academy of Sciences of the United States of America*, **101**, 10967–10972. 3
- [21] Porter, L. L. and Rose, G. D. (2011) Redrawing the Ramachandran plot after inclusion of hydrogen-bonding constraints. *Proceedings of the National Academy of Sciences*, **108**, 109–113. 3
- [22] Kendrew, J. C., Bodo, G., Dintzis, H. M., Parrish, R. G., Wyckoff, H., and Phillips, D. C. (1958) A three-dimensional model of the myoglobin molecule obtained by x-ray analysis. *Nature*, **181**, 662–666. 3
- [23] Matthews, B. W. (1976) X-ray Crystallographic Studies of Proteins. *Annual review of physical chemistry*, **27**, 493–493. 3
- [24] Clore, G. M. and Gronenborn, A. M. (1990) Two-, three-, and four-dimensional NMR methods for obtaining larger and more precise three-dimensional structures of proteins in solution. *Biophysics and Biophysical Chemistry*, **20**, 29–63. 3
- [25] Wuthrich, K. (1989) Protein structure determination in solution by nuclear magnetic resonance spectroscopy. *Science (New York, N.Y.)*, **243**, 45–50. 3
- [26] Banavar, J. R. and Maritan, A. (2007) Physics of proteins. *Annual review of biophysics and biomolecular structure*, **36**, 261–280. 3
- [27] Rose, G. D., Fleming, P. J., Banavar, J. R., and Maritan, A. (2006) A backbone-based theory of protein folding. *Proceedings of the National Academy of Sciences of the United States of America*, **103**, 16623–16633. 3

- [28] Pace, C. N. C., Shirley, B. A. B., McNutt, M. M., and Gajiwala, K. K. (1995) Forces contributing to the conformational stability of proteins. *FASEB Journal*, **10**, 75–83. 3
- [29] Jackson (1998) How do small single-domain proteins fold? *Folding and Design*, **3**, 0–0. 4, 5
- [30] Privalov, P. L. (1979) Stability of Proteins Small Globular Proteins. C B Anfinsen, J. T. E. and Richards, F. M. (eds.), *Advances in Protein Chemistry*, pp. 167–241 T2 –, Academic Press. 4
- [31] Dill, K. A. K. and Maccallum, J. L. J. (2012) The protein-folding problem, 50 years on. *Science (New York, N.Y.)*, **338**, 1042–1046. 4
- [32] Bowman, G. R. G., Voelz, V. A. V., and Pande, V. S. V. (2011) Taming the complexity of protein folding. *Current opinion in structural biology*, **21**, 8–8. 4
- [33] Gibbs, J. W. (1875) On the Equilibrium of Heterogeneous Substances. *Transactions of the Connecticut Academy*, **III**, 108–248. 4
- [34] Pace, C. N. (1986) [14]Determination and analysis of urea and guanidine hydrochloride denaturation curves. *sciencedirect.com*, pp. 266–280, Elsevier. 4
- [35] Gong, H., Porter, L. L., and Rose, G. D. (2011) Counting peptide-water hydrogen bonds in unfolded proteins. *Protein science : a publication of the Protein Society*, **20**, 417–427. 4
- [36] Levinthal, C. (1969) How to fold Graciously. *Mossbauer Spectroscopy in Biological Systems: Proceedings of a meeting held at Allerton House, Monticello, Illinois*, pp. 22–24. 4
- [37] Kubelka, J. J., Hofrichter, J. J., and Eaton, W. A. W. (2004) The protein folding 'speed limit'. *Current opinion in structural biology*, **14**, 13–13. 4
- [38] Wagner, C. and Kiefhaber, T. (1999) Intermediates can accelerate protein folding. *Proceedings of the National Academy of Sciences of the United States of America*, **96**, 6716–6721. 5
- [39] Zwanzig, R. R., Szabo, A. A., and Bagchi, B. B. (1991) Levinthal's paradox. *Proceedings of the National Academy of Sciences of the United States of America*, **89**, 20–22. 5
- [40] Schmid, F. X. F. and Baldwin, R. L. R. (1979) The rate of interconversion between the two unfolded forms of ribonuclease A does not depend on guanidinium chloride concentration. *Journal of Molecular Biology*, **133**, 285–287. 5

- [41] Kim, P. S. and Baldwin, R. L. (1982) Specific Intermediates in the Folding Reactions of Small Proteins and the Mechanism of Protein Folding. *Biochemistry*, **51**, 459–489. 5
- [42] Ptitsyn, O. B. and Uversky, V. N. (1994) The Molten Globule Is a 3rd Thermodynamical State of Protein Molecules. *Febs Letters*, **341**, 15–18. 5
- [43] Guo, Z. and Thirumalai, D. (1995) Kinetics of protein folding: Nucleation mechanism, time scales, and pathways. *Biopolymers*, **36**, 83–102. 5
- [44] KARPLUS, M. and WEAVER, D. L. (1976) Protein-Folding Dynamics. *Nature*, **260**, 404–406. 5
- [45] Karplus, M. and Weaver, D. L. (1979) Diffusion–collision model for protein folding. *Biopolymers*, **18**, 1421–1437. 5
- [46] Fersht, A. R. (1995) Optimization of rates of protein folding: the nucleation-condensation mechanism and its implications. *Proceedings of the National Academy of Sciences of the United States of America*, **92**, 10869–10873. 5
- [47] Itzhaki, L. S. L., Otzen, D. E. D., and Fersht, A. R. A. (1995) 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**, 29–29. 5
- [48] Sánchez, I. E. I. and Kiefhaber, T. T. (2003) Evidence for Sequential Barriers and Obligatory Intermediates in Apparent Two-state Protein Folding. *Journal of Molecular Biology*, **325**, 10–10. 5
- [49] Bachmann, A. and Kiefhaber, T. (2001) Apparent two-state tendamistat folding is a sequential process along a defined route. *Journal of Molecular Biology*, **306**, 375–386. 5
- [50] Pirchi, M. M., Ziv, G. G., Riven, I. I., Cohen, S. S. S., Zohar, N. N., Barak, Y. Y., and Haran, G. G. (2010) Single-molecule fluorescence spectroscopy maps the folding landscape of a large protein. *Nature Communications*, **2**, 493–493. 5
- [51] Stigler, J. J., Ziegler, F. F., Gieseke, A. A., Gebhardt, J. C. M. J., and Rief, M. M. (2011) The complex folding network of single calmodulin molecules. *Science (New York, N.Y.)*, **334**, 512–516. 5
- [52] Fersht, A. R. (2008) From the first protein structures to our current knowledge of protein folding: delights and scepticisms. *Nature Reviews: Molecular Cell Biology*, **9**, 650–654. 5

- [53] Dill, K. A., Ozkan, S. B., Shell, M. S., and Weikl, T. R. (2008) The Protein Folding Problem. *Annual Review of Biophysics*, **37**, 289–316. 5
- [54] Kiefhaber, T. and Bieri, O. (2000) Kinetic models in protein folding. Pain, R. H. (ed.), *Mechanism of Protein Folding*, Oxford University Press. 5
- [55] Eyring, H. (1935) The Activated Complex in Chemical Reactions. *The Journal of Chemical Physics*, **3**, 107–115. 5
- [56] Leffler, J. E. (1953) Parameters for the Description of Transition States. *Science (New York, N.Y.)*, **117**, 340–341. 6
- [57] Matouschek, A. A., Kellis, J. T. J., Serrano, L. L., and Fersht, A. R. A. (1989) Mapping the transition state and pathway of protein folding by protein engineering. *Nature*, **340**, 122–126. 6
- [58] Fersht, A. R. A., Matouschek, A. A., and Serrano, L. L. (1992) The folding of an enzyme. I. Theory of protein engineering analysis of stability and pathway of protein folding. *Journal of Molecular Biology*, **224**, 771–782.
- [59] Sánchez, I. E. and Kiefhaber, T. (2003) Origin of Unusual ϕ -values in Protein Folding: Evidence Against Specific Nucleation Sites. *Journal of Molecular Biology*, **334**, 1077–1085.
- [60] Fersht, A. R. and Sato, S. (2004) Φ -Value analysis and the nature of protein-folding transition states. *Proceedings of the National Academy of Sciences of the United States of America*, **101**, 7976–7981. 6
- [61] Sánchez, I. E. I. and Kiefhaber, T. T. (2003) 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**, 18–18. 6
- [62] Shortle, D. D. (1995) The denatured state (the other half of the folding equation) and its role in protein stability. *FASEB Journal*, **10**, 27–34. 6
- [63] Kiefhaber, T. T., Bachmann, A. A., and Jensen, K. S. K. (2012) Dynamics and mechanisms of coupled protein folding and binding reactions. *Current opinion in structural biology*, **22**, 9–9. 7
- [64] Huang, Y. and Liu, Z. (2009) Kinetic Advantage of Intrinsically Disordered Proteins in Coupled Folding–Binding Process: A Critical Assessment of the “Fly-Casting” Mechanism. *Journal of Molecular Biology*, **393**, 1143–1159. 7
- [65] Wright, P. E. and Dyson, H. J. (1999) Intrinsically unstructured proteins: re-assessing the protein structure-function paradigm. *Journal of Molecular Biology*, **293**, 321–331. 7

- [66] Schellman, J. A. (2002) Fifty years of solvent denaturation. *Biophysical chemistry*, **96**, 91–101. 7, 67
- [67] Schellman, J. A. (1978) Solvent denaturation. *Biopolymers*, **17**, 1305–1322. 7, 67
- [68] Greenfield, N. J. and Fasman, G. D. (1969) Computed circular dichroism spectra for the evaluation of protein conformation. *Biochemistry*, **8**, 4108–4116. 7
- [69] Adler, A. J. A., Hoving, R. R., Potter, J. J., Wells, M. M., and Fasman, G. D. G. (1968) Circular dichroism of polypeptides. Poly(hydroxyethyl-L-glutamine) compared to poly(L-glutamic acid). *J. Am. Chem. Soc.*, **90**, 4736–4738.
- [70] Johnson, W. C. (1990) Protein secondary structure and circular dichroism: A practical guide. *Proteins: Structure, Function, and Bioinformatics*, **7**, 205–214.
- [71] Kelly, S. M., Jess, T. J., and Price, N. C. (2005) How to study proteins by circular dichroism. *Biochimica et biophysica acta*, **1751**, 119–139. 7
- [72] Flory, P. J. (1969) *Statistical Mechanics of Chain Molecules*. John Wiley & Sons, Inc. 7, 8
- [73] Doi, M. and Edwards, S. F. (1986) *The Theory of Polymer Dynamics*. International series of monographs on physics, Oxford University Press. 7, 8
- [74] Debye, P. (1946) The Intrinsic Viscosity of Polymer Solutions. *The Journal of Chemical Physics*, **14**, 636–639. 8
- [75] Soranno, A., Buchli, B., Nettels, D., Cheng, R. R., Müller-Späh, S., Pfeil, S. H., Hoffmann, A., Lipman, E. A., Makarov, D. E., and Schuler, B. (2012) Quantifying internal friction in unfolded and intrinsically disordered proteins with single-molecule spectroscopy. *Proceedings of the National Academy of Sciences*. 8, 15, 53, 56, 64, 93
- [76] Sherman, E. E. and Haran, G. G. (2006) Coil-globule transition in the denatured state of a small protein. *Proceedings of the National Academy of Sciences of the United States of America*, **103**, 11539–11543. 82, 84
- [77] Nettels, D., et al. (2009) Single-molecule spectroscopy of the temperature-induced collapse of unfolded proteins. *Proceedings of the National Academy of Sciences*, **106**, 20740–20745. 92
- [78] Müller-Späh, S., Soranno, A., Hirschfeld, V., Hofmann, H., Rügger, S., Reymond, L., Nettels, D., and Schuler, B. (2010) From the Cover: Charge interactions can dominate the dimensions of intrinsically disordered proteins. *Proceedings of the National Academy of Sciences*, **107**, 14609–14614. 82

- [79] Sadqi, M., Lapidus, L. J., and Munoz, V. (2003) How fast is protein hydrophobic collapse? *Proceedings of the National Academy of Sciences of the United States of America*, **100**, 12117–12122. 91
- [80] Nettels, D., Hoffmann, A., and Schuler, B. (2008) Unfolded protein and peptide dynamics investigated with single-molecule FRET and correlation spectroscopy from picoseconds to seconds. *The Journal of Physical Chemistry B*, **112**, 6137–6146.
- [81] Sherman, E. and Haran, G. (2011) Fluorescence correlation spectroscopy of fast chain dynamics within denatured protein L. *ChemPhysChem*, **12**, 696–703. 84
- [82] Borgia, M. B., Borgia, A., Best, R. B., Steward, A., Nettels, D., Wunderlich, B., Schuler, B., and Clarke, J. (2011) Single-molecule fluorescence reveals sequence-specific misfolding in multidomain proteins. *Nature*, **474**, 662–665.
- [83] Gopich, I. V., Nettels, D., Schuler, B., and Szabo, A. (2009) Protein dynamics from single-molecule fluorescence intensity correlation functions. *The Journal of Chemical Physics*, **131**, 095102. 15
- [84] Schuler, B. and Eaton, W. A. (2008) Protein folding studied by single-molecule FRET. *Current opinion in structural biology*, **18**, 16–26.
- [85] Waldauer, S. A., et al. (2008) Ruggedness in the folding landscape of protein L. *HFSP journal*, **2**, 388–395. 85, 86
- [86] Merchant, K. A. K., Best, R. B. R., Louis, J. M. J., Gopich, I. V. I., and Eaton, W. A. W. (2007) Characterizing the unfolded states of proteins using single-molecule FRET spectroscopy and molecular simulations. *Proceedings of the National Academy of Sciences of the United States of America*, **104**, 1528–1533. 82
- [87] Lum, J. K., Neuweiler, H., and Fersht, A. R. (2012) Long-Range Modulation of Chain Motions within the Intrinsically Disordered Transactivation Domain of Tumor Suppressor p53. *Journal of the American Chemical Society*, **134**, 1617–1622.
- [88] Nettels, D. D., Gopich, I. V. I., Hoffmann, A. A., and Schuler, B. B. (2007) Ultrafast dynamics of protein collapse from single-molecule photon statistics. *Proceedings of the National Academy of Sciences of the United States of America*, **104**, 2655–2660. 15, 85, 86, 92
- [89] Waldauer, S. A., Bakajin, O., and Lapidus, L. J. (2010) Extremely slow intramolecular diffusion in unfolded protein L. *Proceedings of the National Academy of Sciences of the United States of America*, **107**, 13713–13717. 85, 92
- [90] Schuler, B. and Hofmann, H. (2013) Single-molecule spectroscopy of protein folding dynamics-expanding scope and timescales. *Current opinion in structural biology*, pp. -. 14

- [91] Hofmann, H. H., Soranno, A. A., Borgia, A. A., Gast, K. K., Nettels, D. D., and Schuler, B. B. (2012) Polymer scaling laws of unfolded and intrinsically disordered proteins quantified with single-molecule spectroscopy. *Proceedings of the National Academy of Sciences of the United States of America*, **109**, 16155–16160. 11
- [92] Sherman, E., Itkin, A., Kuttner, Y. Y., Rhoades, E., Amir, D., Haas, E., and Haran, G. (2008) Using fluorescence correlation spectroscopy to study conformational changes in denatured proteins. *Biophysical Journal*, **94**, 4819–4827.
- [93] Singh, V. R. V., Kopka, M. M., Chen, Y. Y., Wedemeyer, W. J. W., and Lapidus, L. J. L. (2007) Dynamic similarity of the unfolded states of proteins L and G. *Biochemistry*, **46**, 10046–10054. 14
- [94] Hoffmann, A., et al. (2007) Mapping protein collapse with single-molecule fluorescence and kinetic synchrotron radiation circular dichroism spectroscopy. *Proceedings of the National Academy of Sciences of the United States of America*, **104**, 105–110.
- [95] Lee, J. C., Langen, R., Hummel, P. A., Gray, H. B., and Winkler, J. R. (2004) α -synuclein structures from fluorescence energy-transfer kinetics: Implications for the role of the protein in Parkinson's disease. *Proceedings of the National Academy of Sciences of the United States of America*, **101**, 16466–16471. 8, 53
- [96] Flory, P. J. (1949) The Configuration of Real Polymer Chains. *The Journal of Chemical Physics*, **17**, 303. 9
- [97] Le Guillou, J. and Zinn-Justin, J. (1977) Critical Exponents for the n-Vector Model in Three Dimensions from Field Theory. *Physical Review Letters*, **39**, 95–98. 9
- [98] Bryngelson, J. D. and Wolynes, P. G. (1990) A simple statistical field theory of heteropolymer collapse with application to protein folding. *Biopolymers*, **30**, 177–188. 9
- [99] Sanchez, I. C. (1979) Phase Transition Behavior of the Isolated Polymer Chain. *Macromolecules*, **12**, 980–988. 9
- [100] Dima, R. I. and Thirumalai, D. (2004) Asymmetry in the Shapes of Folded and Denatured States of Proteins. *The Journal of Physical Chemistry B*, **108**, 6564–6570. 9
- [101] Tanford, C., Kawahara, K., and Lapanje, S. (1966) Proteins in 6-M guanidine hydrochloride. Demonstration of random coil behavior. *The Journal of biological chemistry*, **241**, 1921–1923. 9, 52
- [102] Millett, I. S., Doniach, S., and Plaxco, K. W. (2002) Toward a taxonomy of the denatured state: Small angle scattering studies of unfolded proteins. Rose, G. D. (ed.), *Unfolded Proteins*, pp. 241–262, Academic Press. 9

- [103] Wilkins, D. K. D., Grimshaw, S. B. S., Receveur, V. V., Dobson, C. M. C., Jones, J. A. J., and Smith, L. J. L. (1999) Hydrodynamic radii of native and denatured proteins measured by pulse field gradient NMR techniques. *Biochemistry*, **38**, 16424–16431. 9
- [104] Goldenberg, D. P. D. (2003) Computational simulation of the statistical properties of unfolded proteins. *Journal of Molecular Biology*, **326**, 1615–1633. 9, 52
- [105] Aune, K. C., Salahuddin, A., Zarlengo, M. H., and Tanford, C. (1967) Evidence for residual structure in acid- and heat-denatured proteins. *The Journal of biological chemistry*, **242**, 4486–4489. 9
- [106] Kohn, J. E., et al. (2004) Random-coil behavior and the dimensions of chemically unfolded proteins. *Proceedings of the National Academy of Sciences of the United States of America*, **101**, 12491–12496. 9
- [107] NERI, D., BILLETER, M., WIDER, G., and WUTHRICH, K. (1992) Nmr Determination of Residual Structure in a Urea-Denatured Protein, the 434-Repressor. *Science (New York, N.Y.)*, **257**, 1559–1563. 9
- [108] Ohnishi, S. S. and Shortle, D. D. (2003) Effects of denaturants and substitutions of hydrophobic residues on backbone dynamics of denatured staphylococcal nuclease. *Protein science : a publication of the Protein Society*, **12**, 1530–1537.
- [109] Zhang, O. O. and Forman-Kay, J. D. J. (1997) NMR studies of unfolded states of an SH3 domain in aqueous solution and denaturing conditions. *Biochemistry*, **36**, 3959–3970. 9
- [110] Kazmirski and Daggett (1998) Simulations of the structural and dynamical properties of denatured proteins: the "molten coil" state of bovine pancreatic trypsin inhibitor. *Journal of Molecular Biology*, **277**, 20–20. 9
- [111] Shortle, D. D. and Ackerman, M. S. M. (2001) Persistence of native-like topology in a denatured protein in 8 M urea. *Science (New York, N.Y.)*, **293**, 487–489. 9
- [112] Ohnishi, S., Lee, A. L., Edgell, M. H., and Shortle, D. (2004) Direct demonstration of structural similarity between native and denatured eglin C. *Biochemistry*, **43**, 4064–4070.
- [113] Zagrovic, B. B., Snow, C. D. C., Khaliq, S. S., Shirts, M. R. M., and Pande, V. S. V. (2002) Native-like mean structure in the unfolded ensemble of small proteins. *Journal of Molecular Biology*, **323**, 153–164.
- [114] Lietzow, M. A. M., Jamin, M. M., Dyson, H. J. H. J., and Wright, P. E. P. (2002) Mapping long-range contacts in a highly unfolded protein. *Journal of Molecular Biology*, **322**, 655–662.

- [115] Tcherkasskaya, O. O. and Uversky, V. N. V. (2001) Denatured collapsed states in protein folding: example of apomyoglobin. *Proteins*, **44**, 244–254.
- [116] Gillespie, J. R. and Shortle, D. (1997) Characterization of long-range structure in the denatured state of staphylococcal nuclease. I. paramagnetic relaxation enhancement by nitroxide spin labels. *Journal of Molecular Biology*, **268**, 158–169.
- [117] Gillespie, J. R. and Shortle, D. (1997) Characterization of long-range structure in the denatured state of staphylococcal nuclease. II. distance restraints from paramagnetic relaxation and calculation of an ensemble of structures. *Journal of Molecular Biology*, **268**, 170–184.
- [118] Kazmirski, S. L., Wong, K. B., Freund, S. M., Tan, Y. J., Fersht, A. R., and Daggett, V. (2001) Protein folding from a highly disordered denatured state: the folding pathway of chymotrypsin inhibitor 2 at atomic resolution. *Proceedings of the National Academy of Sciences of the United States of America*, **98**, 4349–4354. 9
- [119] Wong, K. B., Clarke, J., Bond, C. J., Neira, J. L., Freund, S. M., Fersht, A. R., and Daggett, V. (2000) Towards a complete description of the structural and dynamic properties of the denatured state of barnase and the role of residual structure in folding. *Journal of Molecular Biology*, **296**, 26–26. 9
- [120] Fitzkee, N. C. and Rose, G. D. (2004) Reassessing random-coil statistics in unfolded proteins. *Proceedings of the National Academy of Sciences of the United States of America*, **101**, 12497–12502. 9, 52, 83
- [121] Tanford, C. (1968) Protein denaturation. *Advances in protein chemistry*, **23**, 121–282. 9
- [122] Jha, A. K., Colubri, A., Freed, K. F., and Sosnick, T. R. (2005) Statistical coil model of the unfolded state: Resolving the reconciliation problem. *Proceedings of the National Academy of Sciences of the United States of America*, **102**, 13099–13104. 10, 52
- [123] Shen, Y. and Bax, A. (2007) Protein backbone chemical shifts predicted from searching a database for torsion angle and sequence homology. *Journal of Biomolecular Nmr*, **38**, 289–302. 10
- [124] Neal, S. S., Nip, A. M. A., Zhang, H. H., and Wishart, D. S. D. (2003) Rapid and accurate calculation of protein ¹H, ¹³C and ¹⁵N chemical shifts. *Journal of Biomolecular Nmr*, **26**, 215–240. 10
- [125] Kjaergaard, M., Brander, S., and Poulsen, F. M. (2011) Random coil chemical shift for intrinsically disordered proteins: effects of temperature and pH. *Journal of Biomolecular Nmr*, **49**, 139–149. 10

- [126] Kjaergaard, M. and Poulsen, F. M. (2011) Sequence correction of random coil chemical shifts: correlation between neighbor correction factors and changes in the Ramachandran distribution. *Journal of Biomolecular Nmr*, **50**, 157–165.
- [127] Kjaergaard, M. M., Nørholm, A.-B. A., Hendus-Altenburger, R. R., Pedersen, S. F. S., Poulsen, F. M. F., and Kragelund, B. B. B. (2010) Temperature-dependent structural changes in intrinsically disordered proteins: formation of alpha-helices or loss of polyproline II? *Protein science : a publication of the Protein Society*, **19**, 1555–1564. 11
- [128] Tamiola, K. K., Acar, B. B., and Mulder, F. A. A. F. (2010) Sequence-specific random coil chemical shifts of intrinsically disordered proteins. *J. Am. Chem. Soc.*, **132**, 18000–18003.
- [129] Camilloni, C., De Simone, A., Vranken, W. F., and Vendruscolo, M. (2012) Determination of Secondary Structure Populations in Disordered States of Proteins Using Nuclear Magnetic Resonance Chemical Shifts. *Biochemistry*, **51**, 2224–2231.
- [130] Marsh, J. A., Singh, V. K., Jia, Z., and Forman-Kay, J. D. (2006) Sensitivity of secondary structure propensities to sequence differences between alpha- and gamma-synuclein: Implications for fibrillation. *Protein science : a publication of the Protein Society*, **15**, 2795–2804.
- [131] De Simone, A., Cavalli, A., Hsu, S.-T. D., Vranken, W., and Vendruscolo, M. (2009) Accurate Random Coil Chemical Shifts from an Analysis of Loop Regions in Native States of Proteins. *J. Am. Chem. Soc.*, **131**, 16332–+. 10
- [132] Lipsitz, R. S. R. and Tjandra, N. N. (2003) Residual dipolar couplings in NMR structure analysis. *Annual review of biophysics and biomolecular structure*, **33**, 387–413. 10
- [133] Bax, A. A., Kontaxis, G. G., and Tjandra, N. N. (2000) Dipolar couplings in macromolecular structure determination. *Methods in enzymology*, **339**, 127–174.
- [134] Bernadó, P. P., Blanchard, L. L., Timmins, P. P., Marion, D. D., Ruigrok, R. W. H. R., and Blackledge, M. M. (2005) A structural model for unfolded proteins from residual dipolar couplings and small-angle x-ray scattering. *Proceedings of the National Academy of Sciences of the United States of America*, **102**, 17002–17007. 10, 49, 53
- [135] Clore, G. M. G. and Iwahara, J. J. (2009) Theory, practice, and applications of paramagnetic relaxation enhancement for the characterization of transient low-population states of biological macromolecules and their complexes. *Chemical Reviews*, **109**, 4108–4139. 10

- [136] Salmon, L., Nodet, G., Ozenne, V., Yin, G., Jensen, M. R., Zweckstetter, M., and Blackledge, M. (2010) NMR characterization of long-range order in intrinsically disordered proteins. *Journal of the American Chemical Society*, **132**, 8407–8418. 10, 53
- [137] Huang, J.-r. and Grzesiek, S. (2010) Ensemble Calculations of Unstructured Proteins Constrained by RDC and PRE Data: A Case Study of Urea-Denatured Ubiquitin. *J. Am. Chem. Soc.*, **132**, 694–705. 10, 11
- [138] Guinier, A. (1939) *Diffraction des rayons X aux très petits angles: application à l'étude des phénomènes ultramicroscopiques...* Masson. 10
- [139] Svergun, D. I. and Koch, M. H. J. (2003) Small-angle scattering studies of biological macromolecules in solution. *Reports on Progress in Physics*, **66**, 1735–1782. 10
- [140] Svergun, D. I., Barberato, C., and Koch, M. H. J. (1995) CRY SOL - a Program to Evaluate X-ray Solution Scattering of Biological Macromolecules from Atomic Coordinates. *J. Appl. Cryst.*, **28**, 768–773. 10
- [141] Chacón, P. P., Morán, F. F., Díaz, J. F. J., Pantos, E. E., and Andreu, J. M. J. (1998) Low-Resolution Structures of Proteins in Solution Retrieved from X-Ray Scattering with a Genetic Algorithm. *Biophysical Journal*, **74**, 16–16.
- [142] Mertens, H. D. T. and Svergun, D. I. (2010) Structural characterization of proteins and complexes using small-angle X-ray solution scattering. *Journal of structural biology*, **172**, 128–141.
- [143] Bernadó, P. P., Mylonas, E. E., Petoukhov, M. V. M., Blackledge, M. M., and Svergun, D. I. D. (2007) Structural characterization of flexible proteins using small-angle X-ray scattering. *J. Am. Chem. Soc.*, **129**, 5656–5664. 10, 53
- [144] Grishaev, A. A., Guo, L. L., Irving, T. T., and Bax, A. A. (2010) Improved fitting of solution X-ray scattering data to macromolecular structures and structural ensembles by explicit water modeling. *J. Am. Chem. Soc.*, **132**, 15484–15486.
- [145] Bernadó, P. and Svergun, D. I. (2012) Structural analysis of intrinsically disordered proteins by small-angle X-ray scattering. *Molecular bioSystems*, **8**, 151–167. 10, 53
- [146] Receveur-Brechot, V. and Durand, D. (2012) How Random are Intrinsically Disordered Proteins? A Small Angle Scattering Perspective. *Current Protein & Peptide Science*, **13**, 55–75. 10
- [147] Meier, S., Blackledge, M., and Grzesiek, S. (2008) Conformational distributions of unfolded polypeptides from novel NMR techniques. *The Journal of Chemical Physics*, **128**, –052204. 10

- [148] Jensen, M. R., Markwick, P. R. L., Meier, S., Griesinger, C., Zweckstetter, M., Grzesiek, S., Bernadó, P., and Blackledge, M. (2009) Quantitative Determination of the Conformational Properties of Partially Folded and Intrinsically Disordered Proteins Using NMR Dipolar Couplings. *Structure*, **17**, 1169–1185. 53
- [149] Jensen, M. R., Ruigrok, R. W., and Blackledge, M. (2013) Describing intrinsically disordered proteins at atomic resolution by NMR. *Current opinion in structural biology*, pp. –. 10
- [150] Jensen, M. R., Salmon, L., Nodet, G., and Blackledge, M. (2010) Defining conformational ensembles of intrinsically disordered and partially folded proteins directly from chemical shifts. *Journal of the American Chemical Society*, **132**, 1270–1272. 10, 53
- [151] Ozenne, V. V., Schneider, R. R., Yao, M. M., Huang, J.-r. J., Salmon, L. L., Zweckstetter, M. M., Jensen, M. R. M., and Blackledge, M. M. (2012) Mapping the potential energy landscape of intrinsically disordered proteins at amino acid resolution. *J. Am. Chem. Soc.*, **134**, 15138–15148. 10, 11
- [152] Bernadó, P., Modig, K., Grela, P., Svergun, D. I., Tchorzewski, M., Pons, M., and Akke, M. (2010) Structure and Dynamics of Ribosomal Protein L12: An Ensemble Model Based on SAXS and NMR Relaxation. *Biophysical Journal*, **98**, 2374–2382. 10
- [153] Choy, W. Y. and Forman-Kay, J. D. (2001) Calculation of ensembles of structures representing the unfolded state of an SH3 domain. *Journal of Molecular Biology*, **308**, 1011–1032.
- [154] Marsh, J. A., Neale, C., Jack, F. E., Choy, W.-Y., Lee, A. Y., Crowhurst, K. A., and Forman-Kay, J. D. (2007) Improved Structural Characterizations of the drkN SH3 Domain Unfolded State Suggest a Compact Ensemble with Native-like and Non-native Structure. *Journal of Molecular Biology*, **367**, 17–17.
- [155] Marsh, J. A. and Forman-Kay, J. D. (2009) Structure and Disorder in an Unfolded State under Nondenaturing Conditions from Ensemble Models Consistent with a Large Number of Experimental Restraints. *Journal of Molecular Biology*, **391**, 359–374. 10
- [156] Fitzkee, N. C., Fleming, P. J., and Rose, G. D. (2005) The Protein Coil Library: A structural database of nonhelix, nonstrand fragments derived from the PDB. *Proteins: Structure, Function, and Bioinformatics*, **58**, 852–854. 10, 50
- [157] Lovell, S. C., Davis, I. W., Arendall, W. B., de Bakker, P. I. W., Word, J. M., Prisant, M. G., Richardson, J. S., and Richardson, D. C. (2003) Structure validation by Calpha geometry: phi,psi and Cbeta deviation. *Proteins*, **50**, 437–450. 10

- [158] Uversky, V. N. (2009) Intrinsically disordered proteins and their environment: effects of strong denaturants, temperature, pH, counter ions, membranes, binding partners, osmolytes, and macromolecular crowding. *Protein Journal*, **28**, 305–325. 11
- [159] Kratky, O. and Porod, G. (1949) Röntgenuntersuchung gelöster Fadenmoleküle. *Revue des Travaux Chimiques des Pays-Bas*, **68**, 1106–1122. 11
- [160] Thirumalai, D. and Ha, B. Y. (1997) Statistical Mechanics of Semiflexible Chains: A meanfield variational approach. *Theoretical and Mathematical Models in Polymer Research*, pp. 1–35, Academic Press. 11
- [161] Bouchiat, C., Wang, M. D., Allemand, J. F., Strick, T., Block, S. M., and Croquette, V. (1999) Estimating the Persistence Length of a Worm-Like Chain Molecule from Force-Extension Measurements. *Biophysical Journal*, **76**, 409–413. 11
- [162] Puchner, E. M. E., Franzen, G. G., Gautel, M. M., and Gaub, H. E. H. (2008) Comparing Proteins by Their Unfolding Pattern. *Biophysical Journal*, **95**, 9–9.
- [163] Berkemeier, F. F., Bertz, M. M., Xiao, S. S., Pinotsis, N. N., Wilmanns, M. M., Gräter, F. F., and Rief, M. M. (2011) Fast-folding alpha-helices as reversible strain absorbers in the muscle protein myomesin. *Proceedings of the National Academy of Sciences of the United States of America*, **108**, 14139–14144. 11
- [164] Zhou, H.-X. (2001) Loops in Proteins Can Be Modeled as Worm-Like Chains. *The Journal of Physical Chemistry B*, **105**, 6763–6766. 11
- [165] Lapidus, L. J., Steinbach, P. J., Eaton, W. A., Szabo, A., and Hofrichter, J. (2002) Effects of Chain Stiffness on the Dynamics of Loop Formation in Polypeptides. Appendix: Testing a 1-Dimensional Diffusion Model for Peptide Dynamics. *The Journal of Physical Chemistry B*, **106**, 11628–11640. 12, 14
- [166] Buscaglia, M., Lapidus, L. J., Eaton, W. A., and Hofrichter, J. (2006) Effects of denaturants on the dynamics of loop formation in polypeptides. *Biophysical Journal*, **91**, 276–288. 12, 14
- [167] Edwards, S. F. (1965) The statistical mechanics of polymers with excluded volume. *Proceedings of the Physical Society*, **85**, 613–624. 12
- [168] Navon, A., Ittah, V., Landsman, P., Scheraga, H. A., and Haas, E. (2000) Distributions of Intramolecular Distances in the Reduced and Denatured States of Bovine Pancreatic Ribonuclease A. Folding Initiation Structures in the C-Terminal Portions of the Reduced Protein. *Biochemistry*, **40**, 105–118. 12, 53
- [169] Haas, E., Wilchek, M., Katchalski-Katzir, E., and Steinberg, I. Z. (1975) Distribution of end-to-end distances of oligopeptides in solution as estimated by energy transfer.

- Proceedings of the National Academy of Sciences of the United States of America*, **72**, 1807–1811.
- [170] Haas, E., Katchalski-Katzir, E., and Steinberg, I. Z. (1978) Brownian motion of the ends of oligopeptide chains in solution as estimated by energy transfer between the chain ends. *Biopolymers*, **17**, 11–31. 15, 31, 42, 56
- [171] Beechem, J. M. and Haas, E. (1989) Simultaneous determination of intramolecular distance distributions and conformational dynamics by global analysis of energy transfer measurements. *Biophysical Journal*, **55**, 1225–1236. 15
- [172] Möglich, A., Joder, K., and Kiefhaber, T. (2006) 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**, 12394–12399. 15, 17, 43, 45, 82
- [173] Grupi, A. and Haas, E. (2011) Segmental Conformational Disorder and Dynamics in the Intrinsically Disordered Protein alpha-Synuclein and Its Chain Length Dependence. *Journal of Molecular Biology*, **405**, 1267–1283. 15, 17, 42
- [174] Grupi, A. and Haas, E. (2011) Time-Resolved FRET Detection of Subtle Temperature-Induced Conformational Biases in Ensembles of alpha-Synuclein Molecules. *Journal of Molecular Biology*, **411**, 234–247. 15, 17, 42, 53, 90, 91, 92
- [175] Huang, F., et al. (2009) Multiple conformations of full-length p53 detected with single-molecule fluorescence resonance energy transfer. *Proceedings of the National Academy of Sciences*, **106**, 20758–20763. 12, 15, 17, 42
- [176] Bieri, O., Wirz, J., Hellrung, B., Schutkowski, M., Drewello, M., and Kiefhaber, T. (1999) 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**, 9597–9601. 14, 56, 57
- [177] Sisamakis, E., Valeri, A., Kalinin, S., Rothwell, P. J., and Seidel, C. A. M. (2010) Accurate single-molecule FRET studies using multiparameter fluorescence detection. *Methods in enzymology*, **475**, 455–514. 14
- [178] Choi, U. B. U., McCann, J. J. J., Weninger, K. R. K., and Bowen, M. E. M. (2011) Beyond the random coil: stochastic conformational switching in intrinsically disordered proteins. *Structure*, **19**, 566–576. 14
- [179] Dexter, D. L. (1953) A Theory of Sensitized Luminescence in Solids. *The Journal of Chemical Physics*, **21**, 836–850. 14

- [180] Satzger, H., Schmidt, B., Root, C., Zinth, W., Fierz, B., Krieger, F., Kiefhaber, T., and Gilch, P. (2004) Ultrafast Quenching of the Xanthone Triplet by Energy Transfer: New Insight into the Intersystem Crossing Kinetics. *Journal of Physical Chemistry A*, **108**, 10072–10079. 14
- [181] Heinz, B., Schmidt, B., Root, C., Satzger, H., Milota, F., Fierz, B., Kiefhaber, T., Zinth, W., and Gilch, P. (2006) On the unusual fluorescence properties of xanthone in water. *Physical Chemistry Chemical Physics*, **8**, 3432–3439. 14
- [182] Krieger, F. F., Fierz, B. B., Bieri, O. O., Drewello, M. M., and Kiefhaber, T. T. (2003) Dynamics of Unfolded Polypeptide Chains as Model for the Earliest Steps in Protein Folding. *Journal of Molecular Biology*, **332**, 265–274. 14, 17, 45
- [183] Chang, I.-J., Lee, J. C., Winkler, J. R., and Gray, H. B. (2003) The protein-folding speed limit: intrachain diffusion times set by electron-transfer rates in denatured Ru(NH₃)₅(His-33)-Zn-cytochrome c. *Proceedings of the National Academy of Sciences of the United States of America*, **100**, 3838–3840. 14
- [184] Lapidus, L. J., Eaton, W. A., and Hofrichter, J. (2000) 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. 14
- [185] Lapidus, L. J. L., Eaton, W. A. W., and Hofrichter, J. J. (2001) Dynamics of intramolecular contact formation in polypeptides: distance dependence of quenching rates in a room-temperature glass. *Physical Review Letters*, **87**, 258101–258101.
- [186] Lapidus, L. J. L., Eaton, W. A. W., and Hofrichter, J. J. (2002) Measuring dynamic flexibility of the coil state of a helix-forming peptide. *Journal of Molecular Biology*, **319**, 19–25.
- [187] Buscaglia, M. M., Schuler, B. B., Lapidus, L. J. L., Eaton, W. A. W., and Hofrichter, J. J. (2003) Kinetics of Intramolecular Contact Formation in a Denatured Protein. *Journal of Molecular Biology*, **332**, 4–4.
- [188] Buscaglia, M., Kubelka, J., Eaton, W. A., and Hofrichter, J. (2005) Determination of Ultrafast Protein Folding Rates from Loop Formation Dynamics. *Journal of Molecular Biology*, **347**, 8–8. 14
- [189] Neuweiler, H., Schulz, A., Böhmer, M., Enderlein, J., and Sauer, M. (2003) Measurement of submicrosecond intramolecular contact formation in peptides at the single-molecule level. *J. Am. Chem. Soc.*, **125**, 5324–5330. 14
- [190] Neuweiler, H. and Sauer, M. (2004) Using photoinduced charge transfer reactions to study conformational dynamics of biopolymers at the single-molecule level. *Current Pharmaceutical Biotechnology*, **5**, 285–298. 14

- [191] Huang, F. F. and Nau, W. M. W. (2003) A conformational flexibility scale for amino acids in peptides. *Angewandte Chemie International Edition*, **42**, 2269–2272. 14
- [192] Huang, F., Hudgins, R. R., and Nau, W. M. (2004) Primary and secondary structure dependence of peptide flexibility assessed by fluorescence-based measurement of end-to-end collision rates. *J. Am. Chem. Soc.*, **126**, 16665–16675.
- [193] Hudgins, R. R., Huang, F., Gramlich, G., and Nau, W. M. (2002) 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.
- [194] Roccatano, D., Sahoo, H., Zacharias, M., and Nau, W. M. (2007) Temperature Dependence of Looping Rates in a Short Peptide. *The Journal of Physical Chemistry B*, **111**, 2639–2646. 14, 17
- [195] Haran, G., Haas, E., Szpikowska, B. K., and Mas, M. T. (1992) Domain motions in phosphoglycerate kinase: determination of interdomain distance distributions by site-specific labeling and time-resolved fluorescence energy transfer. *Proceedings of the National Academy of Sciences of the United States of America*, **89**, 11764–11768. 15
- [196] Gryczynski, I. I., Wicz, W. W., Johnson, M. L. M., Cheung, H. C. H., Wang, C. K. C., and Lakowicz, J. R. J. (1988) Resolution of end-to-end distance distributions of flexible molecules using quenching-induced variations of the Forster distance for fluorescence energy transfer. *Biophysical Journal*, **54**, 577–586. 15
- [197] Lakowicz, J. R. J., Kuśba, J. J., Szmacki, H. H., Gryczynski, I. I., Eis, P. S. P., Wicz, W. W., and Johnson, M. L. M. (1991) Resolution of end-to-end diffusion coefficients and distance distributions of flexible molecules using fluorescent donor-acceptor and donor-quencher pairs. *Biopolymers*, **31**, 1363–1378. 15, 43
- [198] Maliwal, B. P. B., Kuśba, J. J., Wicz, W. W., Johnson, M. L. M., and Lakowicz, J. R. J. (1993) End-to-end diffusion coefficients and distance distributions from fluorescence energy transfer measurements: enhanced resolution by using multiple acceptors with different Förster distances. *Biophysical chemistry*, **46**, 273–281. 15, 43
- [199] Gryczynski, I., Lakowicz, J. R., and Kuśba, J. (1995) End-to-end diffusion coefficients and distance distributions from fluorescence energy transfer measurements: Enhanced resolution by using multiple donors with different lifetimes. *Journal of Fluorescence*, **5**, 195–203. 15, 43
- [200] Maliwal, B. P. B., Lakowicz, J. R. J., Kupryszewski, G. G., and Rekowski, P. P. (1993) Fluorescence study of conformational flexibility of RNase S-peptide: distance-distribution, end-to-end diffusion, and anisotropy decays. *Biochemistry*, **32**, 12337–12345. 15, 42

- [201] Lakowicz, J. R. J., Nair, R. R., Piszczek, G. G., and Gryczynski, I. I. (2000) End-to-end diffusion on the microsecond timescale measured with resonance energy transfer from a long-lifetime rhenium metal-ligand complex. *Photochemistry and Photobiology*, **71**, 157–161. 15, 42
- [202] Jacob, M. H. M., Dsouza, R. N. R., Ghosh, I. I., Norouzy, A. A., Schwarzlose, T. T., and Nau, W. M. W. (2013) Diffusion-enhanced Förster resonance energy transfer and the effects of external quenchers and the donor quantum yield. *J. Phys. Chem. B*, **117**, 185–198. 15, 43
- [203] Ahmad, B. B., Chen, Y. Y., and Lapidus, L. J. L. (2012) Aggregation of α -synuclein is kinetically controlled by intramolecular diffusion. *Proceedings of the National Academy of Sciences of the United States of America*, **109**, 2336–2341. 15, 17
- [204] Kalinin, S., Sisamakris, E., Magennis, S. W., Felekyan, S., and Seidel, C. A. M. (2010) On the origin of broadening of single-molecule FRET efficiency distributions beyond shot noise limits. *The Journal of Physical Chemistry B*, **114**, 6197–6206. 15
- [205] Gopich, I. I. and Szabo, A. A. (2004) Theory of photon statistics in single-molecule Förster resonance energy transfer. *The Journal of Chemical Physics*, **122**, 14707–14707.
- [206] Nir, E. E., Michalet, X. X., Hamadani, K. M. K., Laurence, T. A. T., Neuhauser, D. D., Kovchegov, Y. Y., and Weiss, S. S. (2006) Shot-noise limited single-molecule FRET histograms: comparison between theory and experiments. *J. Phys. Chem. B*, **110**, 22103–22124. 15
- [207] Kuzmenkina, E. V., Heyes, C. D., and Nienhaus, G. U. (2005) Single-molecule Förster resonance energy transfer study of protein dynamics under denaturing conditions. *Proceedings of the National Academy of Sciences of the United States of America*, **102**, 15471–15476. 15
- [208] Chodera, J. D. J., Singhal, N. N., Pande, V. S. V., Dill, K. A. K., and Swope, W. C. W. (2007) Automatic discovery of metastable states for the construction of Markov models of macromolecular conformational dynamics. *The Journal of Chemical Physics*, **126**, 155101–155101. 15
- [209] Noe, F., Schutte, C., Vanden-Eijnden, E., Reich, L., and Weikl, T. R. (2009) Constructing the equilibrium ensemble of folding pathways from short off-equilibrium simulations. *Proceedings of the National Academy of Sciences of the United States of America*, **106**, 19011–19016.
- [210] Bowman, G. R., Beauchamp, K. A., Boxer, G., and Pande, V. S. (2009) Progress and challenges in the automated construction of Markov state models for full protein systems. *The Journal of Chemical Physics*, **131**, 124101–124101–11.

- [211] Voelz, V. A. V., Bowman, G. R. G., Beauchamp, K. K., and Pande, V. S. V. (2010) Molecular simulation of ab initio protein folding for a millisecond folder NTL9(1-39). *J. Am. Chem. Soc.*, **132**, 1526–1528.
- [212] Beauchamp, K. A., Ensign, D. L., Das, R., and Pande, V. S. (2011) Quantitative comparison of villin headpiece subdomain simulations and triplet-triplet energy transfer experiments. *Proceedings of the National Academy of Sciences*, **108**, 12734–12739. 15
- [213] Noe, F., Doose, S., Daidone, I., Lollmann, M., Sauer, M., Chodera, J. D., and Smith, J. C. (2011) Dynamical fingerprints for probing individual relaxation processes in biomolecular dynamics with simulations and kinetic experiments. *Proceedings of the National Academy of Sciences of the United States of America*, **108**, 4822–4827. 16
- [214] Schulz, J. C. F. J., Schmidt, L. L., Best, R. B. R., Dzubiella, J. J., and Netz, R. R. R. (2012) Peptide chain dynamics in light and heavy water: zooming in on internal friction. *J. Am. Chem. Soc.*, **134**, 6273–6279. 16, 57
- [215] Frauenfelder, H., Chen, G., Berendzen, J., Fenimore, P. W., Jansson, H., McMahon, B. H., Stroe, I. R., Swenson, J., and Young, R. D. (2009) A unified model of protein dynamics. *Proceedings of the National Academy of Sciences of the United States of America*, **106**, 5129–5134. 16
- [216] Fenimore, P. W., Frauenfelder, H., McMahon, B. H., and Young, R. D. (2004) Bulk-solvent and hydration-shell fluctuations, similar to alpha- and beta-fluctuations in glasses, control protein motions and functions. *Proceedings of the National Academy of Sciences of the United States of America*, **101**, 14408–14413. 16
- [217] Lubchenko, V. V., Wolynes, P. G. P., and Frauenfelder, H. H. (2005) Mosaic energy landscapes of liquids and the control of protein conformational dynamics by glass-forming solvents. *J. Phys. Chem. B*, **109**, 7488–7499. 16
- [218] Ansari, A. A., Berendzen, J. J., Bowne, S. F. S., Frauenfelder, H. H., Iben, I. E. I., Sauke, T. B. T., Shyamsunder, E. E., and Young, R. D. R. (1985) Protein states and proteinquakes. *Proceedings of the National Academy of Sciences of the United States of America*, **82**, 5000–5004. 16
- [219] Fierz, B., Satzger, H., Root, C., Gilch, P., Zinth, W., and Kiefhaber, T. (2007) Loop formation in unfolded polypeptide chains on the picoseconds to microseconds time scale. *Proceedings of the National Academy of Sciences*, **104**, 2163–2168. 16, 45
- [220] Henzler-Wildman, K. and Kern, D. (2007) Dynamic personalities of proteins. *Nature*, **450**, 964–972. 16

- [221] Krieger, F., Möglich, A., and Kiefhaber, T. (2005) Effect of Proline and Glycine Residues on Dynamics and Barriers of Loop Formation in Polypeptide Chains. *Journal of the American Chemical Society*, **127**, 3346–3352. 17, 45
- [222] Daidone, I., Neuweiler, H., Doose, S., Sauer, M., and Smith, J. C. (2010) Hydrogen-bond driven loop-closure kinetics in unfolded polypeptide chains. *PLoS computational biology*, **6**, e1000645.
- [223] Neuweiler, H. H., Löllmann, M. M., Doose, S. S., and Sauer, M. M. (2007) Dynamics of Unfolded Polypeptide Chains in Crowded Environment Studied by Fluorescence Correlation Spectroscopy. *Journal of Molecular Biology*, **365**, 14–14. 17, 56, 64
- [224] Krieger, Fierz, Axthelm, Joder, Meyer, and Kiefhaber (2004) Intrachain diffusion in a protein loop fragment from carp parvalbumin. *Chemical Physics*, **307**, 7–7. 17, 45, 47
- [225] Fierz, B. and Kiefhaber, T. (2007) End-to-end vs interior loop formation kinetics in unfolded polypeptide chains. *Journal of the American Chemical Society*, **129**, 672–679. 17, 45
- [226] Neuweiler, H., Johnson, C. M., and Fersht, A. R. (2009) Direct observation of ultrafast folding and denatured state dynamics in single protein molecules. *Proceedings of the National Academy of Sciences*, **106**, 18569–18574. 17
- [227] Möglich, A., Krieger, F., and Kiefhaber, T. (2005) Molecular Basis for the Effect of Urea and Guanidinium Chloride on the Dynamics of Unfolded Polypeptide Chains. *Journal of Molecular Biology*, **345**, 153–162. 17, 45, 64
- [228] Chen, Y., Wedemeyer, W. J., and Lapidus, L. J. (2010) A General Polymer Model of Unfolded Proteins under Folding Conditions. *The Journal of Physical Chemistry B*, **114**, 15969–15975. 17
- [229] Langevin, P. (1908) Sur la théorie du mouvement brownien. *Acad. Sci.*, **146**, 530–533. 17
- [230] Lemons, D. S. and Gythiel, A. (1997) Paul Langevin's 1908 paper "On the Theory of Brownian Motion" ["Sur la théorie du mouvement brownien," C. R. Acad. Sci. (Paris) 146, 530-533 (1908)]. *American Journal of Physics*, **65**, 1079–1081. 17
- [231] von Smoluchowski, M. (1915) Über Brown'sche Molekularbewegung unter Einwirkung äusserer Kräfte und deren Zusammenhang mit der verallgemeinerten Diffusionsgleichung. *Annalen der Physik*, **4**, 1103–1112. 18
- [232] Schulten, K., Non-equilibrium statistical mechanics - lecture notes. <http://www.ks.uiuc.edu/~kosztin/PHYCS498NSM/LectureNotes.html>, accessed May 28, 2013. 18

- [233] Nyquist, H. (1928) Thermal Agitation of Electric Charge in Conductors. *Physical Review*, **32**, 110–113. 18
- [234] Einstein, A. (1905) Über die von der molekularkinetischen Theorie der Wärme geforderte Bewegung von in ruhenden Flüssigkeiten suspendierten Teilchen. *Annalen der Physik*, **322**, 549–560. 18, 55, 92
- [235] von Smoluchowski, M. (1906) Zur kinetischen Theorie der Brownschen Molekularbewegung und der Suspensionen. *Annalen der Physik*, **326**, 756–780. 18, 55, 92
- [236] Wilemski, G. and Fixman, M. (1974) Diffusion-controlled intrachain reactions of polymers. I Theory. *The Journal of Chemical Physics*, **60**, 866–877. 19
- [237] Wilemski, G. and Fixman, M. (1974) Diffusion-controlled intrachain reactions of polymers. II Results for a pair of terminal reactive groups. *The Journal of Chemical Physics*, **60**, 878–890.
- [238] Szabo, A., Schulten, K., and Schulten, Z. (1985) First passage time approach to diffusion controlled reactions. *Journal of Chemical Physics*, **72**, 4350–4357. 19, 64, 84
- [239] Kramers, H. A. (1940) Brownian motion in a field of force and the diffusion model of chemical reactions. *Physica*, **7**, 284–304. 19, 56
- [240] Stokes, G. G. (1852) On the Change of Refrangibility of Light. *Philosophical Transactions of the Royal Society of London*, **142**, 463–562. 20
- [241] Kasha, M. (1950) Characterization of electronic transitions in complex molecules. *Discussions of the Faraday Society*, **9**, 14–19. 21
- [242] Förster, T. (1948) Zwischenmolekulare Energiewanderung und Fluoreszenz. *Annalen der Physik*, **6**, 55–75. 21
- [243] Haugland, R. P., J. Y., and Stryer, L. (1969) Dependence of Kinetics of Singlet-Singlet Energy Transfer on Spectral Overlap. *Proceedings of the National Academy of Sciences of the United States of America*, **63**, 23–&. 21
- [244] Steinberg, I. Z. (1971) Long-range nonradiative transfer of electronic excitation energy in proteins and polypeptides. *Annual Review of Biochemistry*, **40**, 83–114. 22
- [245] Langhals, H., Esterbauer, A. J., Walter, A., Riedle, E., and Pugliesi, I. (2010) Förster resonant energy transfer in orthogonally arranged chromophores. *Journal of the American Chemical Society*, **132**, 16777–16782. 22
- [246] Stryer, L. L. and Haugland, R. P. R. (1967) Energy transfer: a spectroscopic ruler. *Proceedings of the National Academy of Sciences of the United States of America*, **58**, 719–726. 22

- [247] Stryer, L. L. (1977) Fluorescence energy transfer as a spectroscopic ruler. *Biochemistry*, **47**, 819–846. 22
- [248] Amir, D. and Haas, E. (1986) Determination of Intramolecular Distance Distributions in a Globular Protein by Nonradiative Excitation-Energy Transfer Measurements. *Biopolymers*, **25**, 235–240. 23
- [249] Gryczynski, Z., Tenenholz, T., and Bucci, E. (1992) Rates of energy transfer between tryptophans and hemes in hemoglobin, assuming that the heme is a planar oscillator. *Biophysical Journal*, **63**, 648–653.
- [250] Elofsson, A., Rigler, R., Nilsson, L., Roslund, J., Krause, G., and Holmgren, A. (1991) Motion of aromatic side chains, picosecond fluorescence, and internal energy transfer in *Escherichia coli* thioredoxin studied by site-directed mutagenesis, time-resolved fluorescence spectroscopy, and molecular dynamics simulations. *Biochemistry*, **30**, 9648–9656.
- [251] dos Remedios, C. G. C. and Moens, P. D. P. (1995) Fluorescence Resonance Energy Transfer Spectroscopy Is a Reliable 'Ruler' for Measuring Structural Changes in Proteins - Dispelling the Problem of the Unknown Orientation Factor. *Journal of structural biology*, **115**, 11–11. 23
- [252] Sindbert, S., Kalinin, S., Nguyen, H., Kienzler, A., Clima, L., Bannwarth, W., Appel, B., Müller, S., and Seidel, C. A. M. (2011) Accurate distance determination of nucleic acids via Förster resonance energy transfer: implications of dye linker length and rigidity. *Journal of the American Chemical Society*, **133**, 2463–2480. 24
- [253] Lakowicz, J. R., Kuśba, J., Gryczynski, I., Wiczak, W., Szymanski, H., and Johnson, M. L. (1991) End-to-end diffusion and distance distributions of flexible donor-acceptor systems observed by intramolecular energy transfer and frequency-domain fluorometry; enhanced resolution by global analysis of externally quenched and non-quenched samples. *The Journal of Physical Chemistry*, **95**, 9654–9660. 43
- [254] Fierz, B., Joder, K., Krieger, F., and Kiefhaber, T. (2007) Using triplet-triplet energy transfer to measure conformational dynamics in polypeptide chains. *Methods in molecular biology (Clifton, N.J.)*, **350**, 169–187. 45
- [255] Krieger, F. (2004) *Dynamics in Unfolded Polypeptide Chains as Model for Elementary Steps in Protein Folding*. Ph.D. thesis, edoc.unibas.ch. 47
- [256] Joder, K. (2011) *Intramolecular and intermolecular diffusion processes in protein folding and assembly*. Ph.D. thesis, TUM. 47, 57, 64
- [257] Bondi, A. (1964) van der Waals Volumes and Radii. *The Journal of Physical Chemistry*, **68**, 441–451. 49

- [258] flexible meccano. <http://www.ibs.fr/science-213/scientific-output/software/flexible-meccano?lang=en>, accessed May 20, 2013. 49
- [259] The protein coil library. <http://roselab.jhu.edu/coil/>, accessed April 27, 2013. 50
- [260] Swindells, M. B. M., MacArthur, M. W. M., and Thornton, J. M. J. (1995) Intrinsic phi, psi propensities of amino acids, derived from the coil regions of known structures. *Nature Structural & Molecular Biology*, **2**, 596–603. 50
- [261] Manke, C. W. and Williams, M. C. (1985) Internal viscosity of polymers and the role of solvent resistance. *Macromolecules*, **18**, 2045–2051. 55
- [262] Stokes, G. G. (1901) On the Effect of the Internal Friction of Fluids on the Motion of Pendulums. *Mathematical and Physical Papers.*, **3**, 1–10. 55
- [263] Kuhn, W. and Kuhn, H. (1945) Bedeutung beschränkt freier Drehbarkeit für die Viskosität und Strömungsdoppelbrechung von Fadenmolekellösungen I. *Helvetica Chimica Acta*, **28**, 1533–1579. 56
- [264] Kuhn, W. and Kuhn, H. (1946) Bedeutung beschränkt freier Drehbarkeit für die Viskosität und Strömungsdoppelbrechung von Fadenmolekellösungen II. *Helvetica Chimica Acta*, **29**, 71–94.
- [265] Kuhn, W. and Kuhn, H. (1946) Modellmässige Deutung der inneren Viskosität (der Formzähigkeit) von Fadenmolekeln II. *Helvetica Chimica Acta*, **29**, 830–858.
- [266] Kuhn, W. and Kuhn, H. (1946) Modellmässige Deutung der inneren Viskosität (der Formzähigkeitskonstante) von Fadenmolekeln I. *Helvetica Chimica Acta*, **29**, 609–626. 56
- [267] Cerf, R. (1957) La macromolécule en chaîne dans un champ hydrodynamique. Théorie générale. Propriétés dynamo-optiques. *Journal of Polymer Science*, **23**, 125–150. 56
- [268] Peterlin, A. (1966) Dynamic viscosity of polymer solutions. *Kolloid-Zeitschrift & Zeitschrift für Polymere*, **209**, 181–181.
- [269] Peterlin, A. (1967) Frequency dependence of intrinsic viscosity of macromolecules with finite internal viscosity. *Journal of Polymer Science Part A-2: Polymer Physics*, **5**, 179–193. 56, 63
- [270] Cerf, R. (1960) Sur la variation du coefficient de diffusion de translation avec la viscosité du solvant: terme correctif à la loi d'Einstein. *Comptes rendus de l'académie des sciences*, **250**, 3599–3561. 56, 63

- [271] Cerf, R. (1971) Dynamique conformationnelle des macro-molécules. Nouvelle équation de diffusion avec viscosité interne contenant les paramètres stochastiques. *Comptes rendus de l'académie des sciences*, **272**, 56
- [272] Cerf, R. (1972) On the stochastic approach to irreversible statistical mechanics of chain macromolecules. *Chemical Physics Letters*, **16**, 42–44. 56
- [273] Cerf, R. (1973) Possible existence of two classes of conformational changes in linear polymeric chain molecules. *Chemical Physics Letters*, **22**, 613–615. 56
- [274] Peterlin, A. (1972) Origin of internal viscosity in linear macromolecules. *Journal of Polymer Science Part B: Polymer Letters*, **10**, 101–105. 56
- [275] Ansari, A., Jones, C. M., Henry, E. R., Hofrichter, J., and Eaton, W. A. (1992) The role of solvent viscosity in the dynamics of protein conformational changes. *Science (New York, N.Y.)*, **256**, 1796–1798. 56
- [276] Chrnyk, B. A. B. and Matthews, C. R. C. (1990) Role of diffusion in the folding of the alpha subunit of tryptophan synthase from *Escherichia coli*. *Biochemistry*, **29**, 2149–2154.
- [277] Plaxco, K. W. and Baker, D. (1998) 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. 56
- [278] Jacob, M. M., Geeves, M. M., Holtermann, G. G., and Schmid, F. X. F. (1999) Diffusional barrier crossing in a two-state protein folding reaction. *Nature Structural & Molecular Biology*, **6**, 923–926.
- [279] Cellmer, T., Henry, E. R., Hofrichter, J., and Eaton, W. A. (2008) Measuring internal friction of an ultrafast-folding protein. *Proceedings of the National Academy of Sciences*, **105**, 18320–18325. 56
- [280] Qiu, L. and Hagen, S. J. (2004) A limiting speed for protein folding at low solvent viscosity. *J. Am. Chem. Soc.*, **126**, 3398–3399.
- [281] Pabit, S. A., Roder, H., and Hagen, S. J. (2004) Internal friction controls the speed of protein folding from a compact configuration. *Biochemistry*, **43**, 12532–12538.
- [282] Wensley, B. G. B., Batey, S. S., Bone, F. A. C. F., Chan, Z. M. Z., Tumelty, N. R. N., Steward, A. A., Kwa, L. G. L., Borgia, A. A., and Clarke, J. J. (2010) Experimental evidence for a frustrated energy landscape in a three-helix-bundle protein family. *Nature*, **463**, 685–688. 56

- [283] Jas, G. S., Eaton, W. A., and Hofrichter, J. (2001) Effect of Viscosity on the Kinetics of α -Helix and β -Hairpin Formation. *The Journal of Physical Chemistry B*, **105**, 261–272. 56, 57
- [284] Borgia, A. A., Wensley, B. G. B., Soranno, A. A., Nettels, D. D., Borgia, M. B. M., Hoffmann, A. A., Pfeil, S. H. S., Lipman, E. A. E., Clarke, J. J., and Schuler, B. B. (2012) Localizing internal friction along the reaction coordinate of protein folding by combining ensemble and single-molecule fluorescence spectroscopy. *Nature Communications*, **3**, 1195–1195. 56
- [285] Narayanan, R., Pelakh, L., and Hagen, S. J. (2009) Solvent friction changes the folding pathway of the tryptophan zipper TZ2. *Journal of Molecular Biology*, **390**, 538–546. 63
- [286] Nyffenegger, C. (2012) *Effects of Osmolytes and other Co-Solutes on the Dynamics of Loop Formation in Proteins and Peptides*. Ph.D. thesis, Technische Universität München. 64
- [287] Auton, M., Rösgen, J., Sinev, M., Holthauzen, L. M. F., and Bolen, D. W. (2011) Osmolyte effects on protein stability and solubility: A balancing act between backbone and side-chains. *Biophysical chemistry*, **159**, 90–99. 64, 68
- [288] Yancey, P. H. P., Clark, M. E. M., Hand, S. C. S., Bowlus, R. D. R., and Somero, G. N. G. (1982) Living with water stress: evolution of osmolyte systems. *Science (New York, N.Y.)*, **217**, 1214–1222. 67
- [289] Santoro, M. M. M., Liu, Y. Y., Khan, S. M. S., Hou, L. X. L., and Bolen, D. W. D. (1992) Increased thermal stability of proteins in the presence of naturally occurring osmolytes. *Biochemistry*, **31**, 5278–5283. 67
- [290] Arakawa, T. T. and Timasheff, S. N. S. (1985) The stabilization of proteins by osmolytes. *Biophysical Journal*, **47**, 411–414. 67
- [291] Baskakov, I. I., Wang, A. A., and Bolen, D. W. D. (1998) Trimethylamine-N-Oxide Counteracts Urea Effects on Rabbit Muscle Lactate Dehydrogenase Function: A Test of the Counteraction Hypothesis. *Biophysical Journal*, **74**, 8–8. 67
- [292] Baskakov, I. V. I., Kumar, R. R., Srinivasan, G. G., Ji, Y. S. Y., Bolen, D. W. D., and Thompson, E. B. E. (1999) Trimethylamine N-oxide-induced cooperative folding of an intrinsically unfolded transcription-activating fragment of human glucocorticoid receptor. *The Journal of biological chemistry*, **274**, 10693–10696. 67
- [293] Kumar, R., Lee, J. C., Bolen, D. W., and Thompson, E. B. (2001) The conformation of the glucocorticoid receptor af1/tau1 domain induced by osmolyte binds co-regulatory proteins. *The Journal of biological chemistry*, **276**, 18146–18152. 67

Bibliography

- [294] Schellman, J. A. (1975) Macromolecular binding. *Biopolymers*, **14**, 999–1018. 67
- [295] Schellman, J. A. (1987) Selective binding and solvent denaturation. *Biopolymers*, **26**, 549–559. 67
- [296] Schellman, J. A. (1994) The thermodynamics of solvent exchange. *Biopolymers*, **34**, 1015–1026. 67, 68
- [297] Timasheff, S. N. S. (2002) Thermodynamic binding and site occupancy in the light of the Schellman exchange concept. *Biophysical chemistry*, **101-102**, 99–111. 67
- [298] Myers, J. K. J., Pace, C. N. C., and Scholtz, J. M. J. (1995) Denaturant m values and heat capacity changes: relation to changes in accessible surface areas of protein unfolding. *Protein science : a publication of the Protein Society*, **4**, 2138–2148. 68
- [299] Tanford, C. (1970) Protein Denaturation: Part C. Theoretical Models for The Mechanism of Denaturation. C B Anfinsen, J. T. E. and Richards, F. M. (eds.), *Advances in Protein Chemistry*, pp. 1–95 T2 –, Academic Press. 68
- [300] Tanford, C. (1964) Isothermal Unfolding of Globular Proteins in Aqueous Urea Solutions. *Journal of the American Chemical Society*, **86**, 2050–2059. 68
- [301] Nozaki, Y. and Tanford, C. (1963) The Solubility of Amino Acids and Related Compounds in Aqueous Urea Solutions. *Journal of Biological Chemistry*, **238**, 4074–4081. 68
- [302] Lewis, G. N. and Randall, M. (1923) *Thermodynamics and the free energy of chemical substances*. McGraw-Hill. 68
- [303] Cohn, E. J. and Edsall, J. T. (1965) *Proteins, amino acids and peptides as ions and dipolar ions*. Monograph series, Hafner Pub. Co. 68
- [304] Liu, Y. and Bolen, D. W. (1995) The peptide backbone plays a dominant role in protein stabilization by naturally occurring osmolytes. *Biochemistry*, **34**, 12884–12891. 68
- [305] Auton, M. and Bolen, D. W. (2004) Additive transfer free energies of the peptide backbone unit that are independent of the model compound and the choice of concentration scale. *Biochemistry*, **43**, 1329–1342.
- [306] Auton, M. and Bolen, D. W. (2005) Predicting the energetics of osmolyte-induced protein folding/unfolding. *Proceedings of the National Academy of Sciences of the United States of America*, **102**, 15065–15068.
- [307] Auton, M. and Bolen, D. W. (2007) Application of the transfer model to understand how naturally occurring osmolytes affect protein stability. *Methods in enzymology*, **428**, 397–418. 68, 69, 83, 84

- [308] Auton, M., Holthauzen, L. M. F., and Bolen, D. W. (2007) Anatomy of energetic changes accompanying urea-induced protein denaturation. *Proceedings of the National Academy of Sciences of the United States of America*, **104**, 15317–15322. 68
- [309] Mezei, M., Fleming, P. J., Srinivasan, R., and Rose, G. D. (2004) Polyproline II helix is the preferred conformation for unfolded polyalanine in water. *Proteins*, **55**, 502–507. 68, 83
- [310] Hu, K.-N., Havlin, R. H., Yau, W.-M., and Tycko, R. (2009) Quantitative determination of site-specific conformational distributions in an unfolded protein by solid-state nuclear magnetic resonance. *Journal of Molecular Biology*, **392**, 1055–1073. 83
- [311] Cortajarena, A. L., Lois, G., Sherman, E., O’Hern, C. S., Regan, L., and Haran, G. (2008) Non-random-coil Behavior as a Consequence of Extensive PPII Structure in the Denatured State. *Journal of Molecular Biology*, **382**, 203–212. 68, 83
- [312] Fleming, P. J., Fitzkee, N. C., Mezei, M., Srinivasan, R., and Rose, G. D. (2005) A novel method reveals that solvent water favors polyproline II over beta-strand conformation in peptides and unfolded proteins: conditional hydrophobic accessible surface area (CHASA). *Protein science : a publication of the Protein Society*, **14**, 111–118. 68, 83
- [313] Qu, Y. Y., Bolen, C. L. C., and Bolen, D. W. D. (1998) Osmolyte-driven contraction of a random coil protein. *Proceedings of the National Academy of Sciences of the United States of America*, **95**, 9268–9273. 68, 83
- [314] Holthauzen, L. M. F., Rösgen, J., and Bolen, D. W. (2010) Hydrogen bonding progressively strengthens upon transfer of the protein urea-denatured state to water and protecting osmolytes. *Biochemistry*, **49**, 1310–1318. 68, 83
- [315] Latypov, R. F., Liu, D., Jacob, J., Harvey, T. S., Bondarenko, P. V., Kleemann, G. R., Brems, D. N., and Raibekas, A. A. (2009) Denaturant-dependent conformational changes in a beta-trefoil protein: global and residue-specific aspects of an equilibrium denaturation process. *Biochemistry*, **48**, 10934–10947. 82
- [316] Whittington, S. J. S., Chellgren, B. W. B., Hermann, V. M. V., and Creamer, T. P. T. (2005) Urea promotes polyproline II helix formation: implications for protein denatured states. *Biochemistry*, **44**, 6269–6275. 83
- [317] Elam, W. A. W., Schrank, T. P. T., Campagnolo, A. J. A., and Hilser, V. J. V. (2013) Temperature and urea have opposing impacts on polyproline II conformational bias. *Biochemistry*, **52**, 949–958. 83
- [318] Cho, S. S., Reddy, G., Straub, J. E., and Thirumalai, D. (2011) Entropic stabilization of proteins by TMAO. *The Journal of Physical Chemistry B*, **115**, 13401–13407. 83, 84

- [319] Uversky, V. N. and Fink, A. L. (2002) The chicken-egg scenario of protein folding revisited. *Febs Letters*, **515**, 79–83. 83
- [320] Muñoz, V. and Serrano, L. (1994) Elucidating the folding problem of helical peptides using empirical parameters. *Nature structural biology*, **1**, 399–409. 84
- [321] Muñoz, V. and Serrano, L. (1995) Elucidating the folding problem of helical peptides using empirical parameters. III. Temperature and pH dependence. *Journal of Molecular Biology*, **245**, 297–308.
- [322] Muñoz, V. and Serrano, L. (1995) Elucidating the folding problem of helical peptides using empirical parameters. II. Helix macrodipole effects and rational modification of the helical content of natural peptides. *Journal of Molecular Biology*, **245**, 275–296. 84
- [323] Zwanzig, R. R. (1988) Diffusion in a rough potential. *Proceedings of the National Academy of Sciences of the United States of America*, **85**, 2029–2030. 85, 92
- [324] Bryngelson, J. D. and Wolynes, P. G. (1989) Intermediates and barrier crossing in a random energy model (with applications to protein folding). *The Journal of Physical Chemistry*, **93**, 6902–6915. 85
- [325] Pollak, E., Auerbach, A., and Talkner, P. (2008) Observations on rate theory for rugged energy landscapes. *Biophysical Journal*, **95**, 4258–4265. 85
- [326] Nevo, R., Brumfeld, V., Kapon, R., Hinterdorfer, P., and Reich, Z. (2005) Direct measurement of protein energy landscape roughness. *EMBO reports*, **6**, 482–486. 85, 93
- [327] Hyeon, C. and Thirumalai, D. (2003) Can energy landscape roughness of proteins and RNA be measured by using mechanical unfolding experiments? *Proceedings of the National Academy of Sciences of the United States of America*, **100**, 10249–10253. 93
- [328] Yang, W. Y. and Gruebele, M. (2004) Detection-Dependent Kinetics as a Probe of Folding Landscape Microstructure. *Journal of the American Chemical Society*, **126**, 7758–7759. 92
- [329] Ihalainen, J. A., Bredenbeck, J., Pfister, R., Helbing, J., Chi, L., van Stokkum, I. H. M., Woolley, G. A., and Hamm, P. (2007) Folding and unfolding of a photo-switchable peptide from picoseconds to microseconds. *Proceedings of the National Academy of Sciences of the United States of America*, **104**, 5383–5388.
- [330] Milanese, L., Waltho, J. P., Hunter, C. A., Shaw, D. J., Beddard, G. S., Reid, G. D., Dev, S., and Volk, M. (2012) Measurement of energy landscape roughness

- of folded and unfolded proteins. *Proceedings of the National Academy of Sciences*, **109**, 19563–19568. 85, 92
- [331] Oliveira, R. J., Whitford, P. C., Chahine, J., Leite, V. B., and Wang, J. (2010) Coordinate and time-dependent diffusion dynamics in protein folding. *Methods*, **52**, 8–8. 85
- [332] Wang, J. (2004) The complex kinetics of protein folding in wide temperature ranges. *Biophysical Journal*, **87**, 2164–2171.
- [333] Chavez, L. L. L., Onuchic, J. N. J., and Clementi, C. C. (2004) Quantifying the roughness on the free energy landscape: entropic bottlenecks and protein folding rates. *J. Am. Chem. Soc.*, **126**, 8426–8432.
- [334] Liu, F., Nakaema, M., and Gruebele, M. (2009) The transition state transit time of WW domain folding is controlled by energy landscape roughness. *The Journal of Chemical Physics*, **131**, 195101–195101. 85
- [335] Horovitz, A. A., Serrano, L. L., Avron, B. B., Bycroft, M. M., and Fersht, A. R. A. (1990) Strength and co-operativity of contributions of surface salt bridges to protein stability. *Journal of Molecular Biology*, **216**, 1031–1044. 93
- [336] Bolen, D. W. and Rose, G. D. (2008) Structure and Energetics of the Hydrogen-Bonded Backbone in Protein Folding. *Annual Review of Biochemistry*, **77**, 339–362. 93
- [337] Scheraga, H. A. (1968) Calculations of conformations of polypeptides. *Advances in Physical Organic Chemistry*, **6**, 103–184. 93
- [338] Ramachandran, G. N. G. and Sasisekharan, V. V. (1967) Conformation of polypeptides and proteins. *Advances in protein chemistry*, **23**, 283–438. 93
- [339] Basharov, M. A. M. (2011) The internal rotational barriers about $\text{NC}\alpha$ and $\text{C}\alpha\text{C}$ backbone bonds of polypeptides. *European Biophysics Journal*, **41**, 53–61. 93
- [340] Boens, N., et al. (2007) Fluorescence lifetime standards for time and frequency domain fluorescence spectroscopy. *Analytical Chemistry*, **79**, 2137–2149. 102
- [341] Berlman, I. B. (1971) *Handbook of Fluorescence Spectra of Aromatic Molecules*. Academic Press. 102
- [342] Weber, G. (1952) Polarization of the fluorescence of macromolecules. II. Fluorescent conjugates of ovalbumin and bovine serum albumin. *Biochemical Journal*, **51**, 155–167. 103

Bibliography

- [343] Knight, C. G., Willenbrock, F., and Murphy, G. (1992) A novel coumarin-labelled peptide for sensitive continuous assays of the matrix metalloproteinases. *Febs Letters*, **296**, 263–266. 103

Acknowledgements

This work was carried out between October 2008 and November 2013 in the Group for Biophysical Chemistry at the department of chemistry, TUM. Financial support was provided (within other) by the *SFB 863 - Forces in Biomolecular Systems* and the graduate school *Materials Science of Complex Interfaces (CompInt)*.

At first, I want to thank **Prof. Dr. Thomas Kiefhaber** for giving me the great possibility to work in his group, for his supervision and important scientific input and for the independency I experienced during my thesis.

I want to acknowledge all former and current group members for the wonderful atmosphere in- and outside the lab: **Dr. Tobias Aumüller, Dr. Annett Bachmann, Maren Büttner, Dominik Bucher, Markus Guder, Susanne Halbritter, Dr. Michael Hösl, Dr. Kerstin Hoffmann-Jacobsen, Peter Kämmerer, Richard Kil, Sabine Kullick, Natalie Merk, Sabine Neumaier, Dr. Christian Nyffenegger, Dr. Alexander Ogrodnik, Florian Praetorius, Dr. Andreas Reiner, Dr. Tobias Schümmer, Lena Schwarzer, Ursula Seidel, Dr. Karin Stecher, Matthias Stecher, Kristine Steen Jensen (Ph.D.), Jeremy Stoan, Traudl Wenger, Stefan Wicht, Dr. Daniel Winter, Daniela Xhindoli**

Many thanks to **Dr. Tobias Aumüller** for all his suggestions and advice concerning fluorescence labeling and fluorescence methods and for scientific discussions.

Many thanks to **Dr. Kerstin Hoffmann-Jacobsen** for providing me all her detailed insights when handing over the project.

I want to thank **Prof. George Rose (Ph.D.)** for discussions of parts of this results and his advice and ideas on modeling peptide conformations.

For scientific discussions on related results of the group, I want to especially thank **Dr. Christian Nyffenegger** and **Dr. Karin Stecher**. Thanks to **Florian Praetorius** and **Alex Hogrebe**, who worked together with me on the temperature dependence during their master thesis and internship .

Many, many thanks to my **mother, father and sister** and my **partner** for all their extraordinary support, advice and encouragements during the years of working on this thesis.

