Dynamic and Structural Properties of Amyloid Forming and Intrinsically Disordered Polypeptide Chains Determined by Time-Resolved FRET Measurements

Dissertation

zur Erlangung des
Doktorgrades der Naturwissenschaften (Dr. rer. nat.)

der

Naturwissenschaftlichen Fakultät I –Biowissenschaften–

der Martin-Luther-Universität
Halle-Wittenberg,

vorgelegt

von M. Sc. Peter Enke

geb. am 26.04. 1985 in Plauen

Gutachter:

Prof. Dr. Thomas Kiefhaber

Prof. Dr. Wolfgang Paul

Prof. Dr. Claus Seidel

Tag der öffentlichen Verteidigung: 12.03.2021

Table of Contents

1. Introduction

1.1. Proteins

Proteins represent the most common functional macromolecules in the cell. The total protein concentration ranges within 200 to 300 g/l or 20% to 30% (w/v) so that proteins even realize a crowded environment [1]. Obviously, proteins perform more important functions as to realize a crowded environment *e.g.* from cell signaling, scaffolding of the cells shape, cell organization and organelle transport up to the function of an enzyme to catalyze specific reactions. Form the largest known protein *titin,* with more as 3000 kDa [2], to the smallest known *tarsal-less* protein, with only 11 amino acids [3], [4], all of them share the same set of only 20 different α-L amino acids. As is well known, most proteins exhibit their full function due to their three-dimensional structure. The native three-dimensional structure of proteins is mainly encoded by the amino acid sequence, which is determined in the gene. By the transcription, the polymerases translated the gene into mRNA. At the next step, the translation, the mRNA-encoded information is converted into the amino acid sequence by the large ribosome complex. During the translation into a linear amino acid sequence the spontaneously and self-organized folding occurs in a stable and functional three-dimensional structure [5]. Indeed, the folding process is a complex mechanism, which has not been completely understood. The folding mechanism is often described by the controversially discussed energy landscape, which consists of many local minima as shown in Figure 1.1. As described in the energy landscape, there are many different folding pathways besides functional folding, such as misfolding into aggregates or assembly into oligomers or higher-ordered fibrils. This illustrate that the folding process is not only complex but also an extremely critical process for the cell (Figure 1.1.). For example, misfolding of proteins becomes important in the case of same prominent affiliations like Alzheimer's disease, Parkinson's disease, spongiform encephalopathy, type II diabetes as well as Huntington's disease [6].

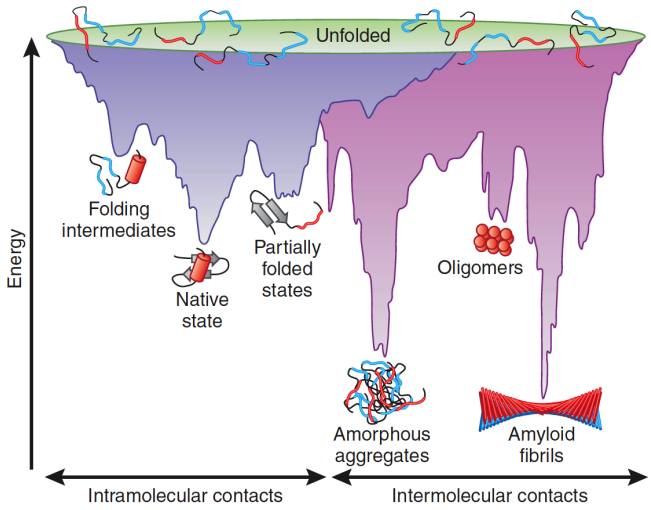


Figure 1.1: Energy landscape scheme of the folding process [7]. The complex folding process may proceed vice kinetic or/and thermodynamic intermediates. Besides folding, additional pathways of aggregation and assembly to oligomers or higher-ordered fibrils are depicted. The main driving force of the described pathways is primarily determined by the intramolecular and intermolecular contacts of the amino acid chain.

1.2. Protein folding

Folding can simply describe as a reversible transition from disordered to an ordered state [8]. Simply proteins are larger polymers consisting of more or less branched amino acids, which undergo folding as a self-assembly process from the unfolded state to the functional three-dimensional tertiary structure. Nevertheless, to understand the folding mechanism remains challenging at a molecular level. At a first glance the spontaneously folding of proteins simply depends directly on the amino acid sequence. But that leads to Levinthal's paradox [9]. If all possible conformations have to be passed to find the correct folded state, a folding process will take more than the age of the universe for a protein consisting of only 100 amino acids. However, the folding process is known to proceed on a millisecond to second timescale [10]. Consequently, there must be specific pathways for the folding process. A possibility to solve this problem is to divide the folding process into sub-processes or to predetermined structures [11], [12]. Three fundamental theories are described to overcome the folding problem.

The *sequential folding* model predicts first the hierarchical formation of secondary structural elements and then the association to the final tertiary structure [13]–[16] in a diffusion-collision controlled mechanism [17]. In the *nucleation-growth* model, local secondary elements are first formed, which serve as a nucleus for rapid growth [18], [19]. In the *hydrophobic collapse* model a "molten globule" is formed as a tertiary hydrophobic structure, which serves as a starting point for the initiation of the secondary structure formation [20], [21]. Of course, a combination of the predicted models could be useful. In fact, for proteins in general different folding models seem to be real, so that the development of a uniform model becomes challenging. Nevertheless, intermediate states as defined for the sequential folding model as well as for the hydrophobic collapse model are often discussed for protein folding and were further predicted to influence the rate of protein folding [22]. A detailed understanding of different intermediates can help to understand the complicated folding mechanism of a protein. Intermediates are generally defined as transient, partially structured and less populated forms in a folding pathway [15], [22]–[26]. Especially the folding of small proteins can be described by a simple two-state mechanism, where the folding proceeds from an unfolded state to a folded state without detectable intermediates [15], [27]. However, detailed investigations verified intermediates also for apparent two-state folders [28], [29], and underline the complex problem of protein folding studies. It seems to be straightforward to investigate the first step of folding, the unfolded state, which has not been fully understood, yet.

1.3. The unfolded state and intrinsically disordered proteins

The protein functions are directly connected to the three-dimensional global structure of proteins, as explained by the lock-end-key model about 100 years ago [30]. On the other hand, events like ligand binding affinities, allosteric transitions and catalytic mechanism cannot fully explained by the folded state of proteins [31]. The unfolding state of proteins in cells is fundamental considering *e.g.* proteolysis as well as protein turnover [19], [32]. Thus, the protein structure-function-paradigm developed in the 20th century, will get a re-evaluation considering the importance of the non-native conformation. This paradigm can be briefly explained as a workflow starting from an

amino acid sequence over the formation of the prominent three-dimensional structure and finishing in the protein function [33], [34]. The observation of chemical denaturation of proteins, first done by Kauzmann *et al.* about 70 years ago, was a cornerstone of protein research with a wide biological significance [35], [36]. For example, the determination of protein stability is inevitable for the biophysical characterization of proteins. The protein stability can only be energetically considered by the comparison of the folded (N) and the unfolded (U) state. In the simplest model, the protein folding pathway can be described by the two-state mechanism, which exclude intermediates (see also Introduction 1.2.)

$$U \underset{k_u}{\overset{k_f}{\rightleftharpoons}} N \qquad \text{(Eq. 1.1)}$$

Folding and unfolding can be described by the folding rate constant k_f and the unfolding rate constant k_u, respectively. Generally, the non-native conformation or unfolded state is defined as an ensemble of many poorly structured conformations or microstates, which behave like a self-avoiding random coil [31], [37], [38]. The same definition is true for natively unfolded proteins or segments, which lack ordered structure under physiological conditions [34]. More than 30% of the eukaryotic proteins reveal wholly or partially disordered regions [33], [39]–[41] and for up to 80% of cancer-associated proteins are predicted to contain disordered segments longer than 30 amino acids [42], [43]. The so-called intrinsically disordered regions or proteins (IDR/IDP) fulfill various functions [34], [44], [45]. Up to 28 different functions are distinguished for intrinsically disordered regions, which can categorized in six functional classes, such as entropic chains (direct function as linker or spacer), display sites (sites of posttranslational modification), chaperones (assist folding), effectors (modulate activity of partner molecules), assemblers (assemble complexes) and scavengers (store and/or neutralize small ligands) [45]–[48]. A coarse classification divides the IDP's into (i) regulation, (ii) scaffolding and recruitment (iii) conformational variability and (iv) adaptability [48], [49].

Interestingly, it was shown that phosphorylation, which is the major regulatory mechanism in cells, is significantly prominent in intrinsically disordered regions [50]. Computational amino acid analysis revealed some disorder-promoting amino acids, especially glutamic acid, glutamine, lysine, arginine, glycine, serine and proline [47],

[51]. It was discussed that the high number of different functions of IDPs is caused by their conformational plasticity [52]. Due the disordered structure a reduced affinity is proposed without loss of affinity [40]. Moreover, the random conformation allows to bind a wide variety of proteins and nucleic acids [40], [53]. The "fly-casting mechanism" explains the potential enhancement of the IDP's association rate to binding partners based on the bigger capture radius resulting from the disordered structure, which additionally allows a relatively larger distance for binding [52], [54]–[56]. IDPs undergo either a significant folding in a secondary structure, or a disorder-to-order transition on binding of a specific binding partner [44], [46], [51], [57]. Here, the binding partners can be ions, *e.g.* calcium ions in the case of β-carp parvalbumin [56], [58], DNA in the case of the DNA-binding domain from brinker [56], [59], intrinsically disordered homologs, *e.g.* of the trimeric foldon domain of bacteriophage T4 fibritin [56], [60], or globular proteins, *e.g.* the pKID/KIX domain of the eukaryotic transcriptional cofactor of CREB-binding protein, respectively [56], [61]. On example is the tumor suppressor p53, which reveals 37% intrinsically disordered structure. It has been shown that this protein can fold into completely different secondary structures on interaction with different binding partners [43], [55]. The p53 folds into an α-helical secondary structure on binding the S100B(ββ) protein then [62] and into a β-sheet structure on binding the Sir2 protein [63].

It becomes clear that an understanding of folding by binding of IDRs/IDPs is an important issue to understand folding in general as well as different regulations and functions in cells. Two different mechanisms for the folding of IDPs on binding of interaction partners are discussed in literature. First, the conformational selection, where IDPs fold before, and second the induced folding, where IDPs fold after binding [55], [56], [64]. The methods for investigation of natively or non-natively unfolded states are numerous. Methods such as near and far UV circular dichroism spectroscopy (CD), multidimensional nuclear magnetic resonance spectroscopy (NMR), small-angle X-ray scattering (SAXS), infrared spectroscopy (FTIR), Raman optical activity spectroscopy (ROA), dynamic light scattering (DLS), single molecule fluorescence as well as different computer simulations are powerful techniques to measure unfolded proteins or peptide chains under non-native (denatured) or native (physiological) conditions. Of course, these methods have different limitations often related to complicated physical models [34], [47], [65]. The unfolded state, whether non-native or native, can be defined as the initial state for the protein folding process

and becomes more and more important to understand folding and broad ranges of cell functions, but remains still challenging to understood in mechanistic detail [38], [66].

1.4 Polymer models for the characterization of the unfolded state

As mentioned above, a detailed understanding of an unfolded state is still challenging. The unfolded state of polypeptide chains consists of an ensemble of conformers, which differ in dimension and dynamics. The distribution of conformers can be described by theoretical polymer models, which try to capture all specific conformers in the measured system. The simplest polymer model corresponds to the ideal chain, where interactions of segments are ignored. Of course, this situation is obviously never realized for a real polymer chain. Nevertheless, the *ideal chain* model is the starting point for further developed models and, additionally, polymers can show nearly ideal chain behavior under certain conditions [67].

For the very simplest ideal chain, the polymer consists of n number of segments with a bond length l. All segments n can be freely oriented independently of each other. This model is known as the *freely jointed chain* model of an ideal chain [68]. The *freely jointed chain* model in the limit of long chains and small deformations is mathematically equivalent to the *Gaussian chain* model. Thus, for a ideal chain the probability density distribution function for an end-to-end distance corresponds to a *Gaussian* function [67]–[69]:

$$p'_{eq}(r) = \left(\frac{3}{2\pi \left\langle r^2 \right\rangle} \right)^{3/2} \exp\left(-\frac{3r^2}{2\left\langle r^2 \right\rangle} \right) \qquad \text{(Eq. 1.2)}$$

The actual probability to find the both chain ends at the distance r needs the consideration of spherical symmetry:

$$p_{eq}(r) = 4\pi r^2 \cdot p'_{eq}(r) \qquad \text{(Eq. 1.3)}$$

The distribution of end-to-end distance within a sphere can thus be normalized as:

$$\int_0^\infty p_{eq}(r)\cdot 4\pi r^2 dr = 1 \qquad \text{(Eq. 1.4)}$$

So that the *r*-dependent probability density of the end-to-end distance of a *Gaussian random walk* can be described as follows [69], [70]:

$$p_{eq}(r) = \left(\frac{3}{2\pi\left\langle r^2\right\rangle}\right)^{3/2} 4\pi r^2 \exp\left(-\frac{3r^2}{2\left\langle r^2\right\rangle}\right) \qquad \text{(Eq. 1.5)}$$

or

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

For the *Gaussian chain* model the mean square end-to-end distance $\left\langle r^2\right\rangle$ is thus calculated as:

$$\left\langle r^2\right\rangle = \int_{r=0}^{\infty} p_{eq}\left(r\right) r^2 4\pi r^2 dr \qquad \text{(Eq. 1.6)}$$

For an ideal chain or *Gaussian chain* the root mean square end-to-end distance $\left\langle r^2\right\rangle^{1/2}$ is averaged over all different conformations of the linear polymer chain. Where *n* denotes for number of beads and *l* denotes for bonds of length.

$$\left\langle r^2\right\rangle^{1/2} = l\cdot n^{1/2} \qquad \text{(Eq. 1.7)}$$

Another important parameter to evaluate the sizes of a polymer is the radius of gyration r_G, which can be determined experimentally *e.g.* by small angle X-ray scattering (SAXS). For a *Gaussian* chain the relationship of the radius of gyration and the mean square end-to-end distance is given as:

$$\left\langle r_G^2\right\rangle = \left\langle r^2\right\rangle / 6 \qquad \text{(Eq. 1.8)}$$

If the assumption of equation 1.4 is universally correct, the size of the polymer chains will scale with a power-law exponent of 0.5 with increasing number of segments. Different experiments show a deviation from the exponential factor of 0.5, especially for SAXS and NMR measurements of chemically unfolded proteins as well as IDPs, which often describes the dimension with a radius of gyration R_G. [71]–[82]. For an ideal chain an infinity small beads are assumed. But real random-coil polymer chains exhibit finite size of beads, which alter the distance distribution towards larger distances. Also, in the simple *Gaussian chain* model the excluded volume, which restrict for overlapping of segments or for general interactions, is ignored. Therefore the scaling exponent ν differs from the theoretically expected value of 0.5.

$$R_G = R_0 N^{\nu} \qquad \text{(Eq. 1.9)}$$

Where R_G is the radius of gyration, R_0 is a constant and depends on the microstructure of the polymer *i.e.* bond length, volume and rotational angles as well as the excluded volume force [69]. Under conditions where polymer-polymer and polymer-solvent interactions are equal, the polymer indeed behaves like an ideal chain and the *Gaussian chain* model becomes valid with a scaling exponent ν of 0.5. Such solvents are called θ-solvents. For a *poor* solvent, the linear polymer chain is more collapsed, since the interactions of polymer-polymer are more favored over the interactions of polymer-solvent. The lower limit of ν is thus 1/3, corresponding to a solid sphere. For *good* solvents the polymer chain is expanded, because the polymer-solvent interactions are more favorable as the polymer-polymer interaction. For such *good* solvents a ν of 3/5 was found, but precise calculations indicate a ν of 0.588, based on the normalization group theory of a self-avoiding walk [78], [83]. In other words, the excluded volume is important and the simple *Gaussian chain* or *freely jointed* model, with vanishing excluded volume, turns out to be unlikely [69], [72]. Nevertheless, *e.g.* in single-molecule spectroscopy the *Gaussian chain* model seems to be successfully applied to describe the distance distribution of the unfolded state or IDPs, respectively [84]–[89]. A theoretical extension of the *freely jointed* model is the *worm-like chain* model or *Kratky-Porod* model [67], [90], which describe rather stiff and almost rod-like polymers. One assumption is that the bond angle is very close to 180° and not fully freely rotatable like in the *freely jointed* model.

Due to an internal stiffness of a real polymer chain in the *wormlike chain* model the so-called persistence length is defined. The persistence length l_p gives an estimation of the length along with the correlations are lost in the direction of the tangent vectors at the chain. An additional assumption is that the segment length is negligibly small compared to the overall length of the polymer chain [91]. In this context, the overall real distance along the backbone of the polymer chain is called contour length L. The probability density to find a certain end-to-end distance is given as:

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

Where $p_{eq}(r)$ is the probability distribution of end-to-end distances and r is the end-to-end distance. l_p is the persistence length. The *wormlike chain* model is also applied in single molecular force spectroscopy [92]–[94]. The persistence length is used to investigate the effect of chain stiffness on the dynamics of loop formation in polypeptides under different conditions [95], [96] and is also compared to Monte Carlo simulations [97]. Nevertheless, the minimal representation of a real polymer chain that adequately describes the end-to-end distance distribution is found by the EDWARDS model [91]. Previous Monte Carlo simulation studies of the end-to-end distribution of polymer compared different polymer models (with an excluded volume) with numerical distributions of end-to-end distances of polymers of different lengths. The EDWARDS model successfully described the distribution with the highest confidence of 95 % [98] and has been used to describe the model polypeptide chain and unfolded proteins [99]–[106]. In the EDWARDS model the distribution function of a polymer of length L with the end-to-end distance r to find a segment is given as [107]:

$$p_{eq}(r) = C \cdot \exp\left(\frac{-27\left[r-\left(\frac{5}{3}\right)^{3/5}\left(\frac{\nu}{3\pi l}\right)^{1/5}\cdot L^{3/5}\right]^2}{20L\cdot l}\right) \qquad \text{(Eq. 1.11)}$$

Where $L=nl$ is described as the number of segments n and a length of the segment l. C is the normalization constant and v represents the excluded volume derived from the repulsive potential between two segments separated by the distance r [107]. After simplification with the free parameters b and σ the probability density to find both ends of the chain within a sphere is normalized to:

$$p_{eq}(r) \propto r^2 \cdot \exp\left(-\left(\frac{r-b}{\sigma}\right)^2\right) \qquad \text{(Eq. 1.12)}$$

with $b = \left(\frac{5}{3}\right)^{3/5} \cdot \left(\frac{v}{3\pi \cdot l}\right)^{1/5} \cdot (n \cdot l)^{3/5}$ and $\sigma^2 = \frac{20 \cdot n \cdot l^2}{27}$ (Eq. 1.12 a and b)

As shown in Figure 1.2 an increase of the parameter b mainly leads to a shift towards larger distances. An increase of the parameter σ mainly leads to a larger width of the end-to-end distribution.

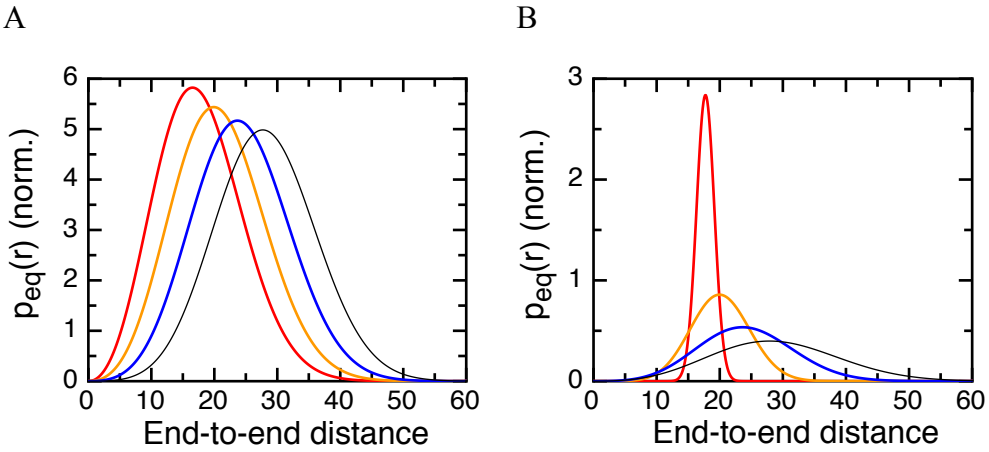


Figure 1.2: End-to-end distance distributions $p_{eq}(r)$ using the EDWARDS model. **A** An increase of the parameter b (from red over orange and blue to black) mainly leads to a shift towards larger distances. **B** An increase of the parameter σ (from red over orange and blue to black) mainly leads to a larger width of the end-to-end distribution. For the calculation of the end-to-end distribution the parameter b ranging in 5 Å steps from 7 Å (red) to 22 Å (black) and the parameter σ ranging in 5 Å steps from 2 Å (red) to 17 Å (black).

1.5. Methods for characterization of the unfolded state

Besides the end-to-end distance or the root mean square end-to-end distance, the radius often successfully describes the size of an unstructured peptide. To obtain information about the radius of a polymer the techniques are often based on the scattering behavior. Thereby, the size can be given as radius of gyration R_G or the radius of hydration R_H. Both radii can be converted to each other, but only with the assumption based on different polymer models like the *Gaussian chain* model and an appropriate hydrodynamic theory. The hydrodynamic radius R_H is defined as a radius of a hypothetical hard sphere, which exhibits the same diffusion coefficient like the examinated particles. Of course, in practice the investigated macromolecules are non-spherical, but dynamic and solvated. Techniques like dynamic light scattering (DLS), fluorescence correlation spectroscopy (FCS) and nuclear magnetic resonance spectroscopy (NMR) provide the translational diffusion coefficient to calculate the hydrodynamic radius assuming the validity of the Stokes-Einstein equation:

$$R_H = \frac{k_B T}{6\pi \cdot \eta \cdot D_t} \qquad \text{(Eq. 1.13)}$$

Where k_B is the Boltzmann constant, T the absolute temperature and η the viscosity of the solvent. D_t represents the translational diffusion coefficient, which is analyzed from the correlation function [108]–[110]. The estimation of the hydrodynamic radius R_H is correct using the self-diffusion coefficient only in the limit range of low concentrations, because in this limit the hydrodynamic coil-coil interactions do not influence the measured diffusion coefficient. It is therefore assumed that the cooperative diffusion coefficient determined in the DLS measurement results in the self-diffusion coefficient.

The radius of gyration can be experimentally determined by the small-angle neutron scattering (SANS) or small-angle X-ray scattering (SAXS) [111]–[113]. The square radius of gyration R_G^2 is defined as the average square distance between segments of any point in the macromolecule from the center of mass R_{CM} [67].

$$R_G^2 = \frac{1}{N}\sum_{i=1}^{N}\left(\vec{R}_i - \vec{R}_{CM}\right)^2 = \frac{1}{N^2}\sum_{i=1}^{N}\sum_{j=1}^{N}\left(\vec{R}_i - \vec{R}_j\right)^2 \qquad \text{(Eq. 1.14)}$$

with $$\vec{R}_{CM} = \frac{1}{N}\sum_{i=1}^{N}\vec{R}_i \qquad \text{(Eq. 1.14 b)}$$

Where N is the number of segments, $\vec{R}_i$ and $\vec{R}_j$ position vectors. The last equality is obtained after insertion of R_{CM} and some algebra. For an ideal linear chain the mean-square radius of gyration $\left\langle R_G^2 \right\rangle$ is related to the mean square end-to-end distance $\left\langle R^2 \right\rangle$ as follows:

$$\left\langle R_G^2 \right\rangle = \frac{b^2 N}{6} = \frac{\left\langle R^2 \right\rangle}{6} \qquad \text{(Eq. 1.15)}$$

Where b denotes the segment length and N is the number of segments. The described relationship depends on the ideal *Gaussian* chains and will differ for linear polymers such as ring polymers. Here, a correction factor of 1/6 is estimated [67]. The determination of R_G or the end-to-end distance of a polymer allows to understand more about its general behavior at different conditions such as temperature, co-solutes, denaturants etc., but does not grant insight into the dynamics of the polymer and the conformational distribution of flexible polymer chains.

Besides the dimension, it is also important to understand the dynamics of polymer chains. One dynamical parameter is represented by the translational diffusion coefficient as discussed for the hydrodynamic radius determination (equation 1.13), which generally describes the diffusion of the whole polymer chain in a solvent. However, also the translational diffusion coefficient does not depend upon internal fluctuations of the unstructured polymer chain. More detailed information on the internal kinetics of polymer chains is revealed in loop formation experiments, where a rate constant of intramolecular loop formation between two ends of an unfolded polypeptide chain is measured. The triplet-triplet energy transfer (TTET) from an excited donor to the energetically lower acceptor yields directly information about the rate constant of loop formation. Mainly two different kinds of TTET-experiments with fundamental differences exist in literature. One TTET-method based on the cysteine-induced quenching of the tryptophan triplet state [84], [95], [96], [114], [115], which is not fully diffusion controlled. Therefore, this method is strongly limited [116]. A successful triplet-quenching method is based on the xanthonic acid derivative as triplet donor and naphthylalanine as acceptor. It was shown that the

quenching of the xanthonic acid by the naphthylalanine is fully diffusion-controlled, irreversible and a fast (ps-range) photophysical processes, which need van-der-Waals contact between donor and acceptor [117]–[120]. Also electron transfer from a triplet-excited Zn-porphyrin to $Rn(NH_3)_5$(His-33) complex incorporated into denatured cytochrome *c* was published to yield the rate constant of loop formation [121]. Intrachain dynamics were also studied by fluorescence quenching methods with a long-lifetime fluorescence probe DBO (2,3-dazabicyclo[2.2.2.]oct-2-ene) as donor and tryptophan as acceptor [122], [123], which revealed comparable limitations for the cysteine-induced quenching of the triplet state of tryptophan [116]. A single-molecule-sensitive fluorescence method with an oxazine derivative as donor and tryptophan as quencher was published as a fluorescence method to measure rate constants of contact formation by fluorescence fluctuation [124], [125]. This was also used in combination with photoinduced electron transfer and fluorescence correlation spectroscopy (PET-FCS) to monitor fast folding dynamics with the help of the *Gaussian chain* polymer model [126], [127]. It was discussed that a rate constant of loop formation is critical to use for a direct determination of the chain dynamics due to the dependence of the dynamics on both internal diffusion as well as on the distance equilibrium distribution [128]. Förster resonance energy transfer at the single molecule level will provide some information on the end-to-end distance distribution of flexible polymer chains in combination with different polymer models as discussed before [129]–[131]. A fundamental idea to combine both the dynamics and the dimension of a flexible polymer chain was first given by Haas *et al.* [100]. Thereby, in time-resolved fluorescence experiments the excited FRET-donor diffuses during the donor-lifetime and changes the FRET efficiency as well as the distance of donor and acceptor. Global fitting of FRET decay curves in dependence of altered solvent viscosity [100] in the absence and presence of a fluorescence quencher [132] was performed to yield the intrachain diffusion. Especially the assumption that the intrachain diffusion coefficient D depends on the viscosity η with $D = 1/\eta$ turns out to be not fully correct [103], [133]. It also was shown that the alteration of the fluorescence donor properties by adding a quencher to yield some information about the intrachain diffusion coefficient is not valid [134].

Nevertheless, using time-resolved FRET seems to be straightforward to get an idea on both dimension and dynamics of flexible unfolded polypeptide chains.

1.6. The fundamental aspect of the time-resolved FRET method

The time-resolved fluorescence/Förster resonance energy transfer method (trFRET) method represents a powerful tool to characterize biophysical properties of unstructured proteins or polypeptide chains based on the chain dynamics and dimensions. The fundamental equation of Haas *et. al.* combines the dependence of the energy transfer efficiency with the dimensional and the dynamic influences of unstructured polypeptide chains on FRET. Thus, a standardized fluorescence lifetime setup can be used to simultaneously obtain information about the dynamics and dimension of unstructured polypeptide chains (Figure 1.3). The single fluorescence decay of a donor becomes multi-exponential by energy transfer to an acceptor due to the influence of both the dimensional distribution of unstructured polypeptide chains (ensemble) and the intrachain diffusion, as illustrated in Figure 1.4. For the experimental setup two FRET-pairs are each coupled to the model polypeptide chains, the dansyl-*pyrene* and the DNB-AMCA (acceptor-donor). Only the large difference of the donor lifetime of two FRET-pairs allows significant information about the dimension and dynamics of polypeptide chains, as discussed and illustrated below (chapter 1.7 and Figure 1.4). In addition to the fluorescence decay of the FRET-pair-coupled model polypeptide chains, also knowledge about the Förster distance, donor-only lifetime and other parameters are required (Material and Methods 6.5).

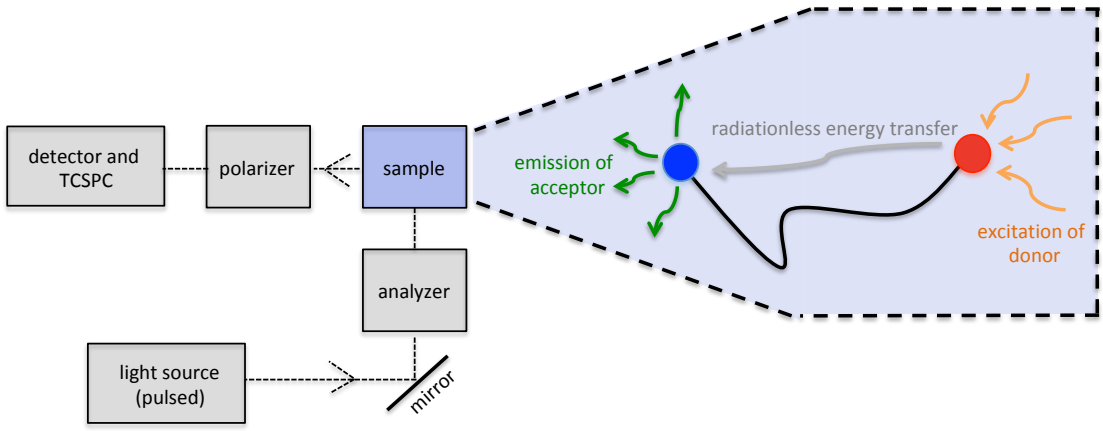


Figure 1.3: Scheme of an instrumental setup of a standardized fluorescence lifetime experiment.
A pulsed light source with a specific wavelength is used. The donor is excited and a radiationless energy transfer from the donor (red circle) to the acceptor (blue circle) occurs. The emission light of the acceptor is detected (PDA) and the signal is processed by the TCSPC system (see below). The polarizer and analyzer enable anisotropy experiments.

The fundamental equation by Haas *et al.* combines the dependence of the FRET efficiency to the distance between donor and acceptor as well as to the influence of the acceptor diffusion during donor lifetime. This provides information on both dimension and dynamics of unfolded polymers:

$$\frac{\partial \overline{p}^*(r,t)}{\partial t} = -\frac{1}{\tau_D} \cdot \overline{p}^*(r,t) - \frac{1}{\tau_D}\left(\frac{R_0}{r}\right)^6 \cdot \overline{p}^*(r,t) + \frac{1}{p_{eq}(r)} \frac{\partial}{\partial r}\left(p_{eq}(r) \cdot D \cdot \frac{\partial \overline{p}^*(r,t)}{\partial r}\right) \qquad \text{(Eq. 2.1)}$$

The partial differential equation (PDE) from Haas *et al.* denotes $\overline{p}^*(r,t)$ as the relative population of excited donors of an end-to-end distance *r* depending on the lifetime of the excited donor τ_D in the absence of an acceptor, the nonradiative energy transfer with R_0 as the Förster distance of the FRET and the influence of the intrachain diffusion *D*. The population of the end-to-end distance distribution of a flexible polymer chain at equilibrium is given as $p_{eq}(r)$. The first term on the right hand side of equation 2.1 describes the radiationless decay of the excited donor with the donor lifetime τ_D to the ground state. The second term represents the dependence of the energy transfer on the distance of the excited donor to the acceptor and the last right hand term represents the reestablishment of the end-to-end distance distribution by the intrachain diffusion. Based on the third term, no analytical solution of the PDE is possible. In order to solve the PDE, a numerical approximation of the solution is necessary. The relative population $\overline{p}^*(r,t)$ is described by the population of the excited donor molecules $p^*(r,t)$ and the population of the end-to-end distance distribution at equilibrium $p_{eq}(r)$, which can be specified by different polymer models (chapter 1.4).

$$\overline{p}^*(r,t) = \frac{p^*(r,t)}{p_{eq}(r)} \qquad \text{(Eq. 2.2)}$$

1.7. Comparison of steady-state FRET and time-resolved FRET method

The Förster resonance energy transfer (FRET) occurs when a fluorophore or chromophore, called donor, is excited and the resulting emission spectrum overlaps with the absorption spectrum of the so-called acceptor [135].
The non-radiative energy transfer from a donor to an acceptor takes place via dipole-dipole coupling and is length dependent. The rate of energy transfer k_T is given by:

$$k_T = \frac{1}{\tau_D}\left(\frac{R_0}{r}\right)^6 \qquad \text{(Eq. 2.3)}$$

Where *r* is the distance between the donor *D* and the acceptor *A* and τ_D is the lifetime of the donor in absence of any energy transfer to the acceptor. The R_0 represents the Förster distance and is defined as the distance between donor and acceptor where the FRET-efficiency is 50%.

$$R_0^6 = \frac{9000\ln(10)\kappa^2 Q_D}{128\pi^5 N_A n^4}\int_0^\infty F_D(\lambda)\varepsilon_A(\lambda)\lambda^4\, d\lambda \qquad \text{(Eq. 2.4)}$$

The Förster distance is described with the spectral overlap integral of the emission spectrum of the donor F_D and the absorption spectrum of the acceptor ε_A and is influenced by the refractive index *n* of the medium [135], [136]. N_A reflects the Avogadro's number. The quantum yield of the donor Q_D is defined as the ratio of the emitted photons and the absorbed photons. The κ^2 describes the relative orientation of the transition dipoles of the donor and acceptor. κ^2 can range from 0 to 4, whereby a value of $\kappa^2 = 0$ means a perpendicular orientation of both fluorophores and a value of $\kappa^2 = 4$ means a parallel orientation. If κ^2 is 0 the rate of energy transfer k_T becomes zero and no FRET can occur [137], [138]. For free and fast rotating fluorophores (dynamic averaging condition) the average orientation factor $\langle\kappa^2\rangle$ is 2/3 and for randomly distributed donor and acceptor $\langle\kappa^2\rangle$ is 0.476 [139], [140].

For the most experiments a value of 2/3 for κ^2 is used [141]–[145]. Interestingly, in previous work a sufficient FRET was even observed for fluorophore pair with a relative orientation factor of zero, which was explained by vibrational motion of the fluorophores [146]. Like the lifetime of donor and the quantum yield, the Förster distance R_0 can be strongly dependent on conditions like pH, temperature or co-solutes. Therefore these parameters should be always determined experimentally. Generally, FRET is a powerful tool for obtaining information on distances *e.g.* in proteins or peptides and is therefore correctly named as a "spectroscopic ruler" [143], [147]–[150]. The Förster distances are typically in the range of 10 to 60 Å [143]. The FRET is commonly used to determine distances between donor and acceptor, which are chemical linked at different position of proteins or macromolecules [151]–[157]. For fixed distances between both fluorophores, the FRET efficiency *E* depends on the lifetime τ, the quantum yield Q and an the fluorescence intensity F as follow:

$$E = 1 - \frac{\tau_{DA}}{\tau_D} = 1 - \frac{Q_{DA}}{Q_D} = 1 - \frac{F_{DA}}{F_D} \qquad \text{(Eq. 2.5)}$$

The index *DA* stands for donor-acceptor pair and the index *D* for donor. In steady-state fluorescence measurements the distance *r* between the two fluorophores *e.g.* in a folded protein can be calculated from the FRET efficiency *E* and the known R_0 as follows:

$$E = \frac{R_0^{\,6}}{R_0^{\,6} + r^6} \qquad \text{(Eq. 2.6)}$$

Besides steady-state fluorescence, time-resolved fluorescence represents a second fundamental method. In time-resolved fluorescence a pulsed light source is used and reveals time dependent intensity decays that primarily provide information about the lifetime of the fluorophore and its anisotropy (Figure 1.3). Such measurements require more complex instrumentation, such as a time-correlated single photon counting (TCSPC) system, but provide detailed molecular information that is not available using steady-state measurements. For time-resolved measurements in general, there is no strong concentration dependence due to the first order function of decay reactions [158].

In principle, the steady-state fluorescence intensity is simply the area of the fluorescence intensity decay of the time-resolved fluorescence measurement [158]:

$$F = \int_0^\infty F_0 \cdot e^{-t/\tau}\, dt = F_0 \cdot \tau \qquad \text{(Eq. 2.7)}$$

Equation 2.7 shows the steady-state fluorescence intensity F, which is proportional to the lifetime τ. Where F_0 can be considered as a parameter, which depends on the fluorophore concentration and on instrument parameters. For unfolded proteins or polypeptide chains, which are highly flexible and dynamic, the relationship of the FRET efficiency and the distance of the fluorophores become more complex. Not only a fixed distance between the fluorophores is expected, but rather an ensemble of differently fluctuating dyes, which represents a distribution of end-to-end distances. Thus, the average FRET efficiency $\langle E \rangle$ depends additionally on the distribution of distances $p_{eq}(r)$ and is directly related to the average fluorescence intensity $\langle F_{DA} \rangle$:

$$\langle E \rangle = \int_{r_{\min}}^{r_{\max}} p_{eq}(r) \frac{R_0^6}{R_0^6 + r^6} dr = 1 - \frac{\langle F_{DA} \rangle}{F_D} \qquad \text{(Eq. 2.8)}$$

The average FRET efficiency $\langle F_{DA} \rangle$ is also related to the average lifetime $\langle \tau_{DA} \rangle$ as shown for the FRET efficiency assuming a fixed distance (equation 2.5):

$$\langle E \rangle = 1 - \frac{\langle \tau_{DA} \rangle}{\tau_D} \qquad \text{(Eq. 2.9)}$$

In time-resolved fluorescence decay the fluorescence intensity $\langle F_{DA} \rangle$ is related to the sum of different lifetimes weighted by the broad distance distribution $p_{eq}(r)$:

$$\langle F_{DA} \rangle(t) = \int_{r_{\min}}^{r_{\max}} p_{eq}(r) \cdot F_0 e^{-t/\tau_{DA}(r)}\, dr \qquad \text{(Eq. 2.10)}$$

In order to determine the $p_{eq}(r)$ from time-resolved FRET data, the average fluorescence intensity $\langle F_{DA} \rangle$ can be calculated by integration over time in combination with equation 2.10. It should be noted that the influence of the intrachain diffusion is negligible:

$$\langle F_{DA} \rangle = \int_{r_{min}}^{r_{max}} \int_{t=0}^{\infty} p_{eq}(r) \cdot F_0 e^{-t/\tau_{DA}(r)} dt dr = F_0 \int_{r_{min}}^{r_{max}} p_{eq}(r) \cdot \tau_{DA}(r) dr = F_0 \langle \tau_{DA} \rangle \quad \text{(Eq. 2.11)}$$

The time-resolved fluorescence intensity decay can be calculated from $\overline{p}^*(r,t)$ of the partial differential equation from Haas *et al.* (equation 2.1.):

$$F(t) = \int_{r_{min}}^{r_{max}} \overline{p}^*(r,t) \cdot p_{eq}(r) dr \quad \text{(Eq. 2.12)}$$

The time-resolved fluorescence decay *F(t)* is influenced both by the distance distribution described by the polymer model and by the intrachain diffusion. In Figure 1.4 the influence of the *F(t)* on the lifetime of the donor is depicted. The donor fluorescence decay shows for both simulations single-exponential decay as expected (A and B in Figure 1.4, green dashed lines). Due to the distance distribution, however, a strong deviation from the single-exponential fluorescence decay is observed (A and B in Figure 1.4, blue lines). The simulated fluorescence decay curves clearly show the strong influences of the intrachain diffusion for the long donor lifetime (τ_D =160 ns) (B in Figure 1.4). A minor influence of the diffusion is observed for short donor lifetime (τ_D = 5 ns) (compare areas under the curves of A and B in Figure 1.4, right side). Additionally, the fluorescence decay for a donor with the long lifetime is extremely fast compared to the fluorescence decay in the absence of diffusion (B in Figure 1.4).

A

B

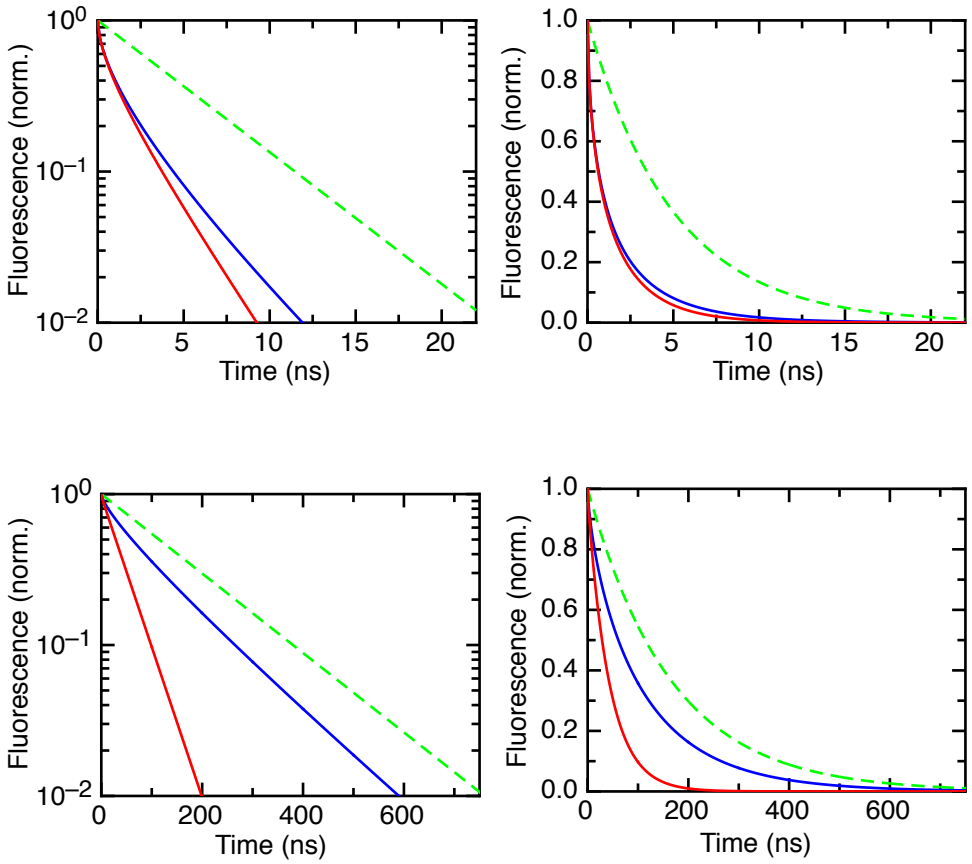


Figure 1.4: Simulated time-resolved fluorescence decay curves according to equation 2.1. and 2.12. For better comparison a semi-logarithmic y-scale on the left side and a linear y-scale on the right side of normalized fluorescence against the time is shown. For **A** a donor lifetime of 5 ns and for **B** 165 ns was used. The PDE is solved for both lifetimes with R_0 = 22 Å, b = 17.5 Å and σ = 12.6 Å. As explained in 2.1, the PDE consists of three terms, first the decay of the donor, second the decay of the donor in the presence of an acceptor (FRET) and third the influence of the intrachain diffusion. For the simulation of the fluorescence decay curves, the dashed green line represents the donor decay without FRET, the blue curve is generated with the first term and the second term of the PDE and the red line depicts the complete PDE with the additional influence of the diffusion coefficient (D_{DA} = 10 Å^2/ns).

For the donor with short lifetime, the fluorescence decay curve shows an only slight difference to the decay considering the diffusion term (areas under the curves, A in Figure 1.4 right side). The fluorescence decay curves of the donor with long lifetime show a high difference between the decay with and without the diffusion term (areas under the curves, B in Figure 1.4 right site). Previous results revealed that a high quality of the calculated intrachain diffusion and the distance distribution can be achieved by globally fitting of two fluorescence decay curves using two different

FRET pairs [103], [159]. Therefore, one FRET pair with a long donor lifetime and another with a short donor lifetime is applied. Nevertheless, the analysis of the fluorescence decay curves by a global fit requires a polymer model to describe the population of the end-to-end distance distribution $p_{eq}(r)$. As discussed in chapter 1.4, different polymer models were published to capture the ensemble of conformers in an end-to-end distance distribution. Previous works showed that the EDWARDS model was successfully used to analyze data of different unfolded polypeptide chains [103], [159]. To compare the different polymer models and give evidence for the preferred EDWARDS' model, MC simulations with different polypeptide chains were performed. The MC simulation was done by the software *flexible meccano* [160], which was developed as a tool for the generation of distance distribution of IDPs and unfolded proteins [161]–[163]. The software randomly samples the backbone dihedral ϕ-ψ-angle pairs from a compilation of 500 high-resolution X-ray structures (resolution <1.8 Å). In order to avoid clashes between the amino acids a residual specific excluded volume is used. A sphere placed at the C^{β} (or C^{α} atom of glycine) of the specified amino acid generates this specific excluded volume. In the case of an overlap between to different amino acids, the ϕ-ψ-pair is rejected and another ϕ-ψ-pair is used until no clash is found. For the simulation 300.000 conformers are sampled and fitted with the three different polymer models as explained in detail in chapter 1.4. Three different polypeptide chains with increasing stiffness were investigated, poly-(glycine-serine), polyserine, and polyglutamine. In order to prove the reliability of the polymer model these three polypeptide chains types with different chain lengths (8 aa, 16 aa and 24 aa) were simulated as shown in Figure 1.5. Also two natural sequences namely the two fragments of the natural IDP β-carp parvalbumin, the so-called DE- and EF-loop, were additionally used.

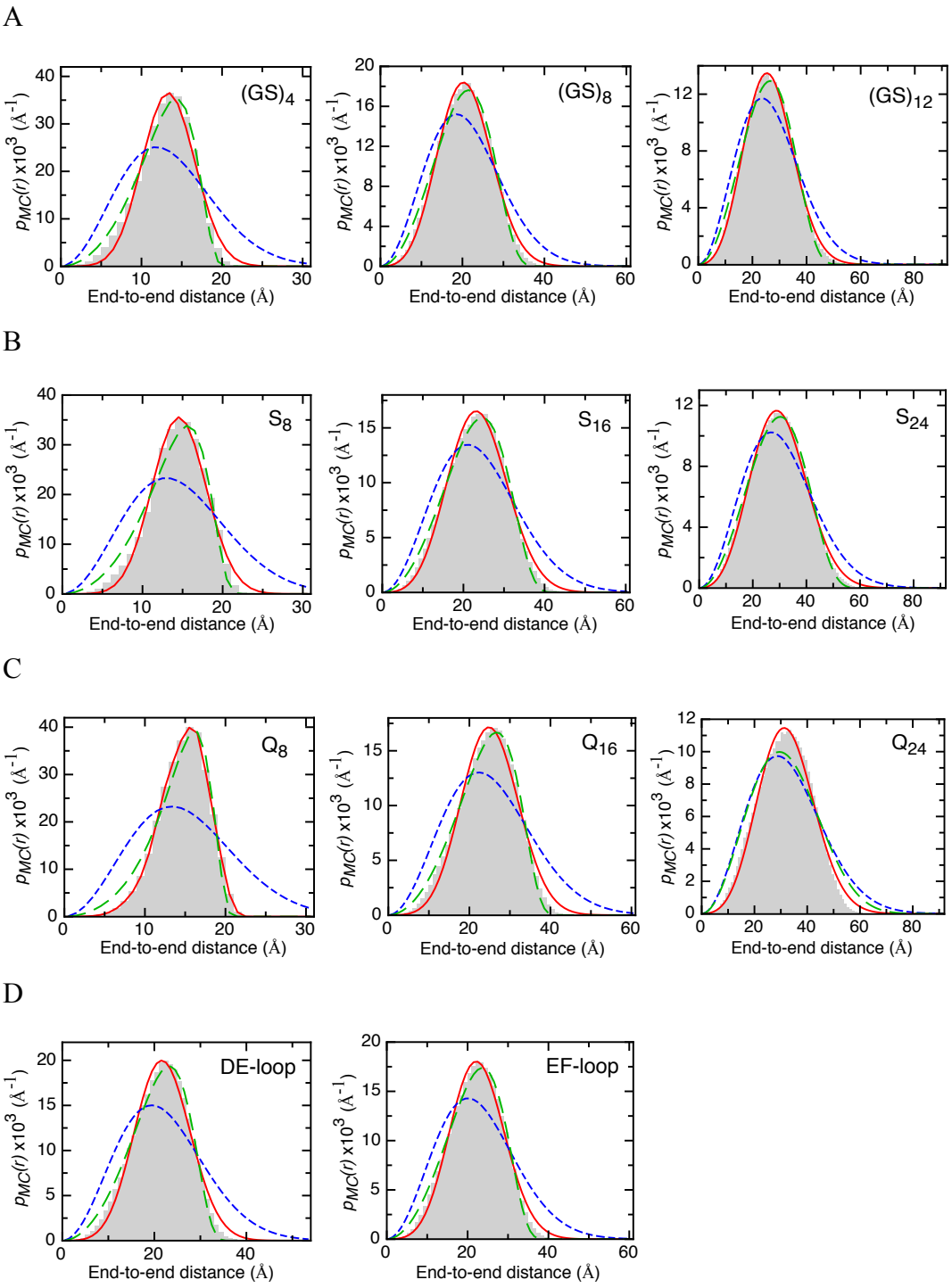


Figure 1.5: Histograms of simulated conformer ensemble generated by *flexible meccano* [160]. 300.000 conformers were sampled for **A** poly-(glycine-serine) **B** polyserine, **C** polyglutamine of different length (8 aa, 16 aa and 24 aa) and **D** natural IDP β-carp parvalbumin fragments DE- and EF-loop. The resulting populations of end-to-end distances were fitted using the *Gaussian chain* model (---) (equation 1.5), the *wormlike chain* model (---) (equation 1.10) or the EDWARDS model (—) (equation. 1.12).

As depicted in Figure 1.5 for all three polypeptide chains with a length of 16 aa, the EDWARDS model reveals a better approximation for shorter distances of the simulated conformer ensemble, while the *wormlike chain* model showed a better approximation for longer distances. Nevertheless, for the medium length of 16 aa both the EDWARDS model and the *wormlike chain* model showed a good approximation as observed for all three types of polypeptide chains. But with increasing length of the polypeptide chains the EDWARDS model shows a significantly better agreement with the simulated conformer ensemble compared to the two other polymer models. The same hold true for the shortest length of 8 aa, where the EDWARDS model gave the best description of the simulated conformer ensemble. In going from the very flexible poly-(glycine-serine) to the stiffer polyglutamine (increasing chain stiffness) the quality of the approximation of the EDWARDS model increased, while the *wormlike chain* model decreased. Additionally, the EDWARDS model showed a good approximation for the DE- and EF-loop fragment of the natural IDP β-carp parvalbumin. The simple *Gaussian chain* model could not fit the simulated conformer ensemble independent of the length and the type of the polypeptide chains. Hence, the EDWARDS model seems to be able to describe the simulated distance distribution with high quality independent of the chain length and stiffness. In conclusion, the time-resolved FRET method is a powerful tool for investigation of the dynamics and the dimensions of different types of unfolded polypeptide chains from natural as well as from unnatural repetitive sequences of different flexibility. In order to describe the conformer ensemble in a high accuracy the EDWARDS model is used.

2. Introduction of amyloidic polyglutamine and the disease-related behavior

In the chapter above, different methods are described to obtain information about the dynamics and dimension of unstructured polypeptide chains. The time-resolved FRET measurement, mainly used in my work, represents a powerful method for characterizing biophysical properties of unstructured polypeptide chains. This chapter discusses the mainly unstructured amyloidic polyglutamine chains, which belong to the IDPs. Polyglutamine chains are involved in several diseases, and the aim of my work is to characterize biophysical properties as a possible disease-related effect of polyglutamine chains.

Polyglutamine segments are found in a variety of proteins, especially in transcription factors and may act as transcriptional regulator domains on binding other transcription factors and co-factors [164], [165]. Polyglutamine tracts in proteins and peptides are discussed in the context of at least nine of expanded CAG repeat diseases [166]–[169], including the neurodegenerative disorders Kenneday's disease, Huntington's disease, spinocerebellar ataxia type (SCA)-1, -2, -3, -6, -7, -17 and dentatorubral pallidoluysian atrophy (DRPLA).

Table 1.: Neurodegenerative diseases caused by expanded polyglutamine [170]. HD: Huntington's disease; SBMA: spinal and bullar muscular atrophy; DRPLA: dentatorubral pallidoluysian atrophy; SCA: spinocerebellar ataxia.

Disease	Protein	Normal repeat length	Pathogenic repeat length
HD	Huntingtin (htt)	6 - 34	36 - 121
SBMA	Androgene receptor	9- 36	38 - 62
DRPLA	Atrophin 1	7- 34	49 - 88
SCA1	Ataxin 1	6 -39	40 - 82
SCA2	Ataxin 2	15 – 24	32 - 200
SCA3	Ataxin 3	13 – 36	61 - 84
SCA6	α1a voltage-dependent calcium channel subunit	4 – 20	20 - 29
SCA7	Ataxin 7	4 – 35	37 - 306
SCA17	TATA box binding protein	25- 42	47 - 63

Pathogenic polyglutamine containing proteins have an abnormal higher length of polyglutamine segments, which result in an accumulation of neuronal aggregates in brain regions and introduces dysfunction and cell death of specific neurons [166], [171]. The specific function of polyglutamine segments of normal length as in transcription factors and co-factors are often unknown or beyond interest [164]. Nevertheless, a well investigated example is the neurodegenerative disease spinocerebellar ataxia-3 or Machado-Joseph disease and the corresponding ataxin-3 protein. The non-pathogenic ataxin-3 protein is discussed to be involved in the ubiquitin-proteasome system and binds polyubiquitin [172]–[175]. Additionally, ataxin-3 regulates the aggresome formation and suppresses polyglutamine-induced toxicity [176], [177]. Ataxin-3 is also known as a transcription repressor via inhibition of histone acetyltransferase activity of *e.g.* cAMP response element-binding protein (CREB) and further transcription co-activators [175], [178]. Interestingly, ataxin-3 proteins undergo a two-step aggregation as depicted in Figure 2.1 [175]. The first step is the formation of amyloid protofibrils by intermolecular interaction of the N-terminal domain called Josephin domain. The second stage of long and straight amyloid fibrils is formed due to an abnormal expansion of polyglutamine segments of above 55 glutamines [179]. The formation of the amyloid fibrils occurs primarily due to the interaction of the abnormally expanded polyglutamine segment at the disordered C-terminal region of ataxin-3. Moreover, an inhibitor of the second stage, the so-called polyQ binding peptide 1 (QBP1), was found. The QBP1 binds to the disordered C-terminal region of ataxin-3 [179]. In general, QBP1 was an effective inhibitor of polyglutamine aggregation *in vitro* and *in vivo* and is discussed as a therapeutic drug against amyloid disease [164], [179].

The example of ataxin-3 illustrates the importance in understanding the molecular origin of polyglutamines to aggregate or forms specific amyloidic fibrils, especially relating to prospective therapeutics. The most prevalent and well-known pathogenic CAG triplet repeat expansion is Huntington's disease. The huntingtin contains a polyglutamine tract like in transcription factors. The 348 kDa human huntingtin is involved in organelle transport via interaction to the molecular motor machinery, mediates endocytosis, vesicle recycling and endosomal trafficking as well as regulates the selective autophagy [165]. More remarkable is the interaction of huntingtin with transcriptional activators and repressors on binding the polyglutamine tract [165].

Above a typical disease risk threshold of 35 - 45 polyglutamine repeats [164] the hungtingtin occurs formation of pathogenic amyloid fibrils.

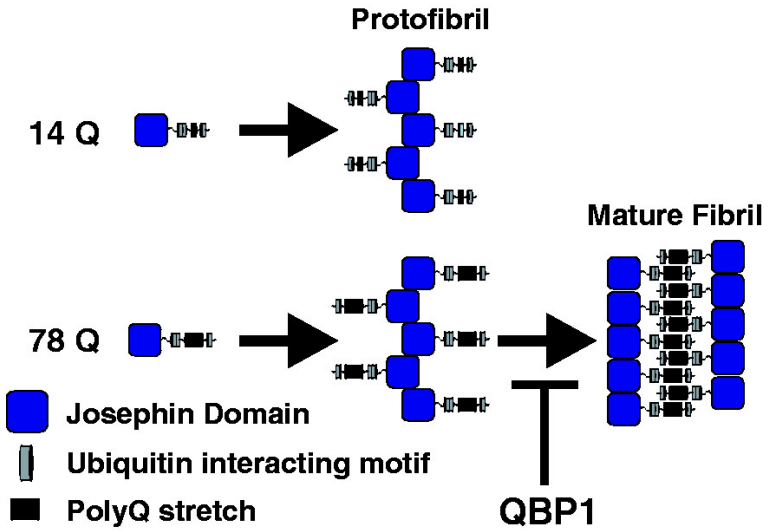


Figure 2.1: Ataxin-3 aggregation mechanism from Knight *et. al.* [179]. Both normal and abnormal ataxin-3 proteins (picture: 14 Q's and 76 Q's, respectively) undergo a first aggregation step and form amyloid protofibrils due to intermolecular interactions of the N-terminal Josephin domain. The second aggregation step is only observed for the abnormal expansion of the polyglutamine segment (picture: 76 Q's) and pathogenic long straight amyloid fibrils are formed. The polyQ binding peptide 1 (QBP1) serves as an inhibitor of the second aggregation step due to direct interaction with the disordered C-terminus.

The polyglutamine segment is located at the N-terminal region of the huntingtin surrounded by 17 amino acids long N-terminal part and the proline-rich C-terminal segment [170] (Figure 2.2).

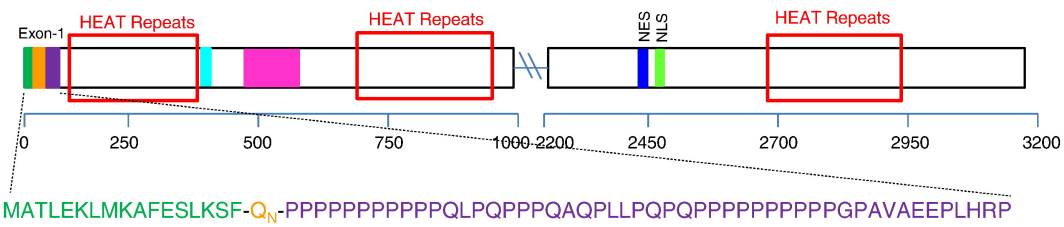


Figure 2.2: Scheme of the huntingtin from Wetzel *et. al.* [164]. The 348 kDa huntingtin consists of several HEAT repeats for protein-protein interaction as well as a nuclear export and a localization signal (NES and NLS). The N-terminus contains the polyglutamine segment (orange), which is flanked by a 17 amino acid N-terminal region (green) and a proline-rich domain (purple). Above a threshold of 35 - 45 polyglutamine repeats the Huntingtin proteins tend to aggregate into toxic amyloid fibrils.

The 17 amino acids in the N-terminal part form an α-helical structure, whereas the polyglutamine stretch is considered as a flexible region. The proline-rich domain is know to form poly-proline type II helical structures, which seem to be important for stabilization of the polyglutamine structure and is discussed to play a critical role for the polyglutamine mediated aggregation [164], [165]. Generally, a two-step aggregation mechanism for the huntingtin was observed analogously to the Ataxin-3 aggregation system discussed above (Figure 2.1) [164]. One step is the formation of reversible α-helix-rich tetramers of huntingtin domains, which can further assemble into higher-order oligomers. The other step is the nucleation of oligomers to amyloid structures, where a β-sheet structure of the polyglutamine segment is formed [164], [180], [181]. The elongation of both steps occurs simply by monomer addition, as illustrated in Figure 2.3.

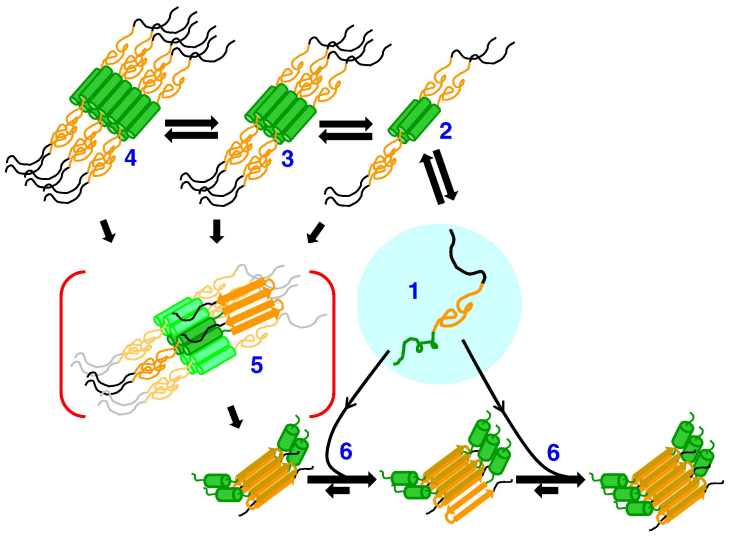
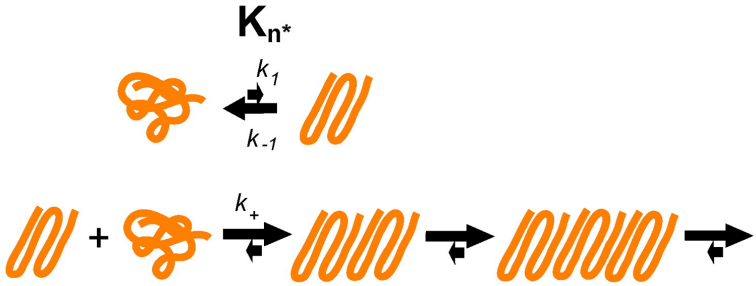


Figure 2.3: Two-step aggregation mechanism of huntingtin N-terminus modified from Wetzel *et. al.* [164]. One step is initialized by the α-helix bundle stabilized tetramer and elongated to high-order oligomers. For the formation of amyloid fibrils any oligomer can undergo the nucleation, which depict a β-sheet structure for the polyglutamine tract. In both reaction pathways of aggregation elongation take place by the addition of monomers. Huntingtin is illustrated in green, proline-rich region is shown in black and the polyglutamine tract is depicted in orange.

The intrinsic polyglutamine sequence seems to be the most appropriate approach to study the disease-related aggregation. The general aggregation of polyglutamine chains were described by a classic nucleation mechanism [164], [182] (Figure 2.4).

In this classic thermodynamic model for nucleated growth polymerization the critical nucleus n^* is described as the thermodynamically unfavorable multimer, which becomes stabilized by further addition of monomers [183] (B in Figure 2.4). In previous work, for a polyglutamine length of about 26 repeats a critical nucleus number of $n^* = 1$ was found. This is not consistent to the classic *nucleated growth* model, where the definition implies that all species on the assembly pathway must be oligomeric [183]. A newly defined *monomeric nucleus* model for polyglutamine does not change the mathematical formalism, but defines the possibility of $n^* = 1$, where the critical nucleus is also thermodynamically unfavorable and represents a very small percentage of total number of monomers [164], [183]. In comparison to the *nucleated growth* model, which is also known as *multimeric nuclei* model, in the new *monomeric nuclei* model the thermodynamic barrier of the formation of the monomeric nucleus depends strongly on the conformational dynamics of the monomer [183].

A



B

Figure 2.4: Models of nucleation mechanism of polyglutamine [164]. **A** Nucleation and elongation mechanism of polyglutamine, where k_1 and k_{-1} are the forward and reverse rate constant of nucleation, K_{n^*} is the equilibrium constant of nucleation and k_+ represents the second-order rate constant for the elongation step, where a β-sheet structure was found. **B** Classic *nucleus* model, where n* is the energetically unfavorable critical nucleus.

Regardless of the critical nucleus, monomeric polyglutamine chains exhibit a predominantly disordered structure [184]–[186].

FCS, atomic force stretching experiments as well as SAXS experiments indicated a compact and more collapse polypeptide chain in water, which means that no fully extended polyglutamine is expected [180], [187], [188]. Consequently, the monomeric state of polyglutamine becomes more and more in the focus of research. In particular, the dynamics and dimensions of the monomeric state of polyglutamine may provide further information to the disease-related behavior of polyglutamine chains.

3. Aim of Research

The disease-related significance of proteins containing polyglutamine is well known for a number of different neurodegenerative diseases. However, the molecular origin of the accompanied aggregation and fibrillation of these polyglutamine-containing proteins and peptides is poorly understood. The aim of this research is to characterize biophysical properties as a possible disease-related effect of polyglutamine at the molecular level. Using the time-resolved FRET method, fundamental physical parameters of chain dimensions and dynamics at the monomeric level of the polypeptide chains should be monitored. Especially, by comparison to non-amyloidic polypeptide chains a better molecular understanding of the monomeric polyglutamine may be achieved and differences to non-amyloidic polypeptide chains should be identified.

A fundamental aspect in protein and especially in IDP research is the influence of solvent viscosity on chain dynamics and structure, which can be investigated in detail by the trFRET method used here. Not only the viscosity itself, which is mediated by the use of viscogenic co-solutes, but also the size of the viscogenic co-solute can strongly influence the dynamics and dimension of unstructured polypeptide chains. Therefore, two viscogenic co-solutes with different size can be used for a detailed investigation. Glycerol is suitable as a well-established and frequently used small viscogenic co-solute. PEG_{6000} can be used as a representative of large co-solutes. Already established model peptides like glycine-serine chains or natural sequences like the EF- and the DE-loop of β-carp parvalbumin can be used for this investigation. Thus, the general influence of viscosity on the dynamics and dimension of IDPs can be further elucidated, but also taken into account for subsequent experiments. With respect to polyglutamine chains, a first important step is to investigate the dynamics and structural properties of these chains in aqueous milieu in order to compare them with non-amyloidic model peptides and to identify possible peculiarities of polyglutamine chains. With respect to the general aggregation properties of polyglutamine chains, it is important to examine the monomeric state of the polyglutamine chains in an aqueous milieu. Methods such as CD, especially light scattering experiments and analytical ultracentrifugation are suitable for this purpose.

Another important aspect is the influence of temperature on dynamics and dimension of IDPs. Physical parameters, such as the activation energy according to *Arrhenius* or conformational energy barriers according to *Zwanzig*, were previously used to describe the dynamic change of unstructured peptide chains. In earlier work on IDPs, the use of denaturants, such as GdmCl, has been established for another important characterization. Besides the influence of GdmCl on hydrophobic properties, the influence on hydrogen bonds is also discussed. Especially with respect to fibril formation, hydrogen bonds play a decisive role. The structural influence of GdmCl on IDPs can be described in more detail with the help of a weak-binding model includes the equilibrium exchange constant of water molecules with GdmCl molecules. Of course, the influence of viscosity on the dynamics and dimension of polyglutamine chains should be known, not only to identify possible disease-related behavior of polyglutamine chains but also to be able to consider the viscosity in the temperature and GdmCl experiments mentioned above. Although some non-amyloidic model peptides have been investigated in previous work, the results obtained here can be used to introduce new model polypeptide chains that represent previously disregarded properties such as enhanced hydrogen bonding. For this purpose, *e.g.* polyserine chain seems to be a suitable candidate. Another goal is to investigate a broad spectrum of different IDP types for comparison. For this purpose, the DE-loop of parvalbumin, as a representative of natural IDP sequences, can be used for completion. This generates a wealth of information, which, by comparing both artificial and natural non-amyloidic IDPs with the amyloidic polyglutamine chains, provides valuable clues for elucidating the biophysical properties as a possible disease-related effect of polyglutamine chains.

4. Results and Discussion

4.1. The Effect of viscogenic co-solutes on the dynamics and dimensions of polypeptide chains

4.1.1 The influence of the viscosity on chain dynamics and dimensions

A fundamental aspect of protein folding is that the unfolded polypeptide chains have to collapse to form a functional compact and globular structure. But the kinetic barriers, which limit the folding rate, are not fully understood [189], [190]. The mechanisms of protein folding depend not only on the intrinsic sequence properties but also on solvent conditions like pH, salt and viscosity [191]. A physical limit on the protein folding speed is given by the rate of diffusional motion of the unfolded polypeptide chain in the solvent [17], [190]–[196]. However, the effect of solvent viscosity is directly correlated to the rate of folding of polypeptide chains [190], [193], [194], [197]–[201]. The Kramers theory provides the relationship of the diffusion-controlled reaction rates and the solvent viscosity [202]. In Kamers' theory the chemical reaction rate of an unimolecular reaction is influenced by the viscosity of the medium. Generally, reactants and products are separated by an energy barrier, which can be crossed in a diffusion-controlled process. Therefore, the particles are subject to Brownian motion [199], [202], [203]. According to Kramers' theory with increasing the macroscopic viscosity η, the rate k decreases with $k \propto 1/\eta$. A number of simulations and experiments show that the diffusion rate and protein folding rate do not always scale with the macroscopic viscosity of the solvent. Especially for fast-folding structures like small proteins, loops, helices etc., the reaction rate is not strictly proportional to the solvent viscosity [117], [193], [204]–[208]. The deviation from the Kramers theory implies that other interactions exist, which are not related to the macroscopic solvent viscosity. Such interactions are summarized in the term of the so-called internal friction or microviscosity [87], [193], [203], [204], [208]–[212]. The origin of internal friction can be potential energy barriers to backbone rotations, long-range interresidue interactions, accessibility of free volume in a non-continuum solvent as well as hydrogen bonds [193], [213].

There exist two empirical models, which include the internal friction in addition to the macroscopic solvent friction. For the first one, a simple additive relationship of the internal and solvent friction is found (the here so-called additive model) [211], [214]:

$$f = f_{solv}\left(\frac{\eta}{\eta_0}\right) + f_{\text{int}} \qquad \text{(Eq. 4.1)}$$

Where f_{solv} and f_{int} denote the coefficient of the solvent and internal friction. The viscosity is represented by η and the reference viscosity is η_0. The other empirical relationship for the internal and solvent viscosity is given as a power-law equation:

$$f = f_{solv}\left(\frac{\eta}{\eta_0}\right)^{\alpha} \qquad \text{(Eq. 4.2)}$$

The viscosity-sensitive exponent α includes the influence of the internal friction and ranges between 0 and 1. Note that with α = 1 the above discussed Kramers' relationship of $k \propto 1/\eta$ is reproduced. However, the power-law equation (equation 4.2) was often showed to be more accurate than the additive model (equation 4.1) in describing the kinetics for viscosity dependence, especially of unfolded polypeptide chains [159], [203], [207], [209], [215]–[218].

Generally, the friction is related to the diffusion coefficient according to the Einstein-Smoluchowski equation of Brownian motion [219]–[221]:

$$D = \frac{k_B T}{f} \qquad \text{(Eq. 4.3)}$$

Where D is the diffusion coefficient, f the friction coefficient, k_B is the Boltzmann constant and T the absolute temperature. Therefore, the intrachain diffusion coefficient can be converted directly to the friction coefficient.

Assuming the validity of the power-law equation 4.2 the combination with the Einstein-Smoluchowski equation 4.3 yields the empirical equation 4.4 [103], [159], [218]:

$$D = D_0 \cdot \left(\frac{\eta}{\eta_0} \right)^{-\alpha} \qquad \text{(Eq. 4.4)}$$

Where η_0 is the reference viscosity and D_0 is the intrachain diffusion coefficient at the lowest viscosity of water under buffer conditions (10 mM potassium phosphate, pH 7.0). The scaling factor α represents a dimensionless exponent, which includes a possible deviation from the relationship of 1/η predicted by Kramers' theory, and is comparable to the α-value yields in equation 4.2.

4.1.2 The influence of the size of viscogenic co-solutes on chain dynamics and dimensions

Not only the macro- and the microviscosity of the medium itself influence the dynamics and dimensions of unfolded polypeptide chains, also the size of viscogenic co-solutes is another important factor in this context [222]–[226]. For instance, protein-protein interaction was shown to be strongly influenced by the size of co-solutes [227], [228]. Additionally, it was observed that the aggregation and fibrillation of Aβ-peptides can be prevented by viscogenic co-solutes of large size [229]. Also the movement of the motor protein kinesin-1 was strongly dependent on the size of the added viscogenic co-solutes leading to a complete suppression of motor function [230].

4.1.3 trFRET measurements using glycerol and PEG_{6000} as viscogenic co-solutes

In order to test the influence of viscosity and the size of the viscogenic co-solutes on both the dynamics and the dimensions of model polypeptide chains, two different viscogenic co-solutes were investigated.
Glycerol as one of the smallest viscogenic co-solutes and PEG_{6000} (polyethylene glycol with an average molecular mass of 6000 Da) as a member of the larger viscogenic co-solutes were investigated at different co-solute concentrations.

A non-natural repetitive sequence (glycine-serine)$_8$ polypeptide chain ((GS)$_8$) and two fragments of 14 aa and 16 aa length from the natural IDP β-carp parvalbumin, called DE-loop and EF-loop were investigated as model polypeptide chains. Generally, the dynamics and dimensions were measured by trFRET using fluorescence donor an acceptor attached to the N-terminus and to an amino acid side chain at the C-terminal end of the polypeptide chain (detailed information on the model polypeptide chains are given in chapter 4.2 and 4.6).

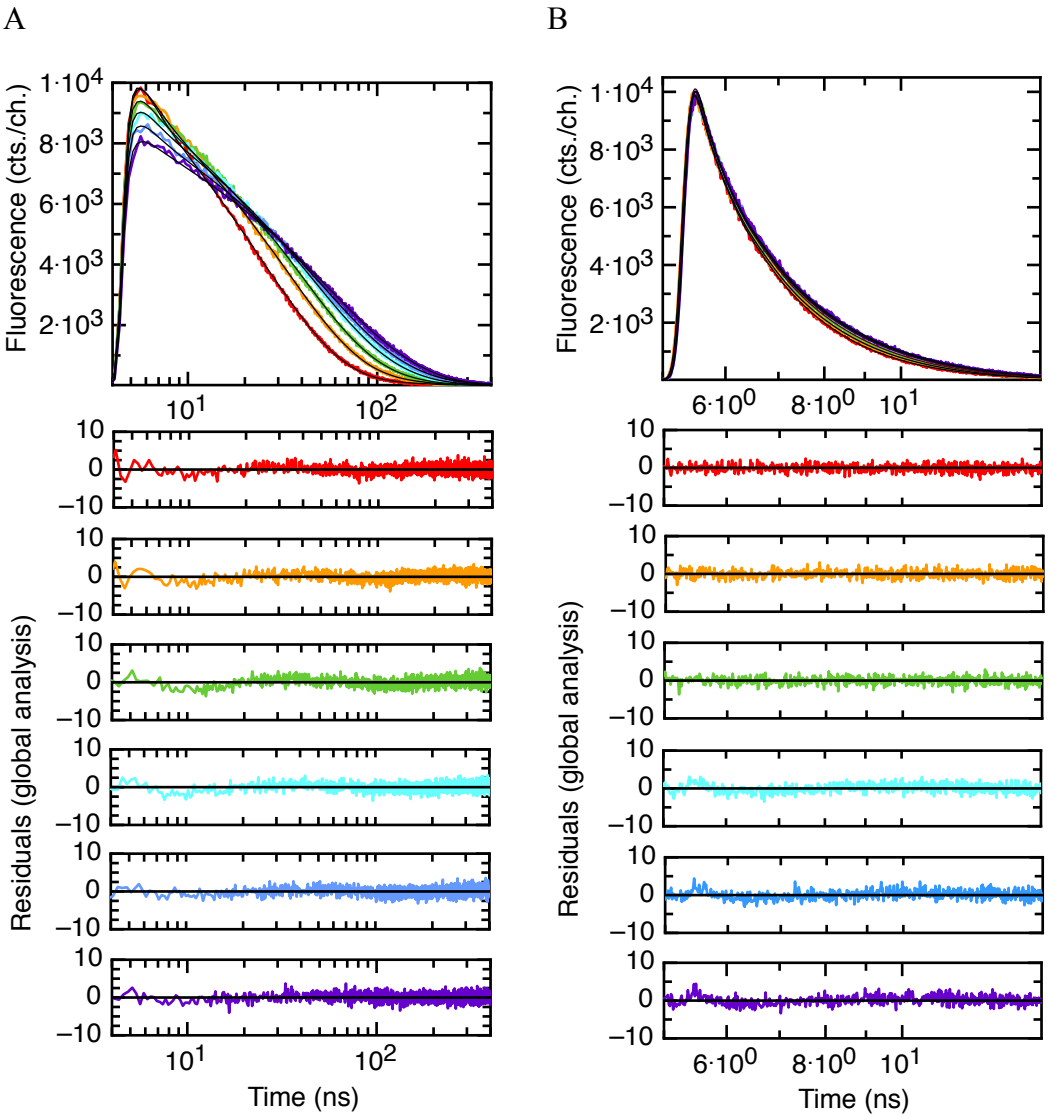


Figure 4.1.1: Glycerol dependence of time-resolved fluorescence decays curves of the β-carp parvalbumin fragment DE-loop with the donor-acceptor pairs **A** dansyl-*pyrene* and **B** DNB-AMCA in 10 mM potassium phosphate pH 7.0. Concentrations of glycerol ranges from 0% (red), 16% (orange), 28% (green), 40% (turquoise), 48% (blue) to 56% (purple) (w/w). The residuals of the global analysis of each concentration point are shown in different colors corresponding to the different concentrations of glycerol for both FRET-pairs.

Fluorescence decay curves of the DE-loop using glycerol ranging from 0% to 56% (w/w) (Figure 4.1.1) and of the $(GS)_8$ (Figure 4.1.4), the EF-loop (Figure 4.1.5) and the DE-loop (Figure 4.1.6) using PEG_{6000} ranging from 1.8% to 21.1% (w/w) of both FRET pairs were measured in 10 mM potassium phosphate buffer at pH 7.0 and 22.5 °C.

The intrachain diffusion coefficients of the DE-loop fragment decreased by a factor of 9.6 from water to 56% (w/w) glycerol. The decrease of the intrachain diffusion coefficient with increasing concentration of the viscogenic co-solute can be described with the empirical equation 4.4 discussed above. According to equation 4.4 a α-value of 1.14 ± 0.1 was found for the DE-loop fragment of the natural IDP β-carp parvalbumin. In previous time-resolved FRET measurements both the $(GS)_8$ and the EF-loop fragment were investigated in the presence of different glycerol concentrations [159]. The analysis using equation 4.4 revealed an α-value of 1.09 ± 0.1 for $(GS)_8$ and of 1.03 ± 0.1 for the EF-loop [159].

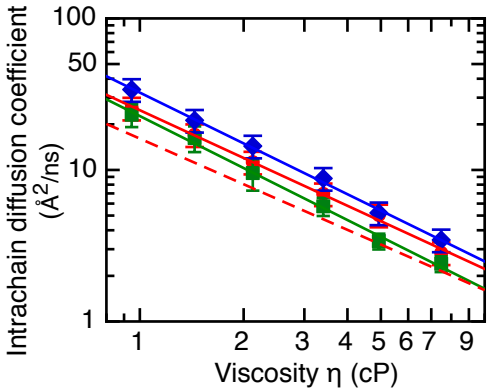


Figure 4.1.2: Double-logarithmic illustration of the intrachain diffusion coefficient in dependence of different glycerol concentrations (0% - 56% (w/w)). The intrachain diffusion coefficient of DE-loop is shown in green squares, of EF-loop is shown in red circles and of $(GS)_8$ is shown in blue diamonds. The results of the analysis using the empirical equation 4.4 are depicted in the colors mentioned above. The red dashed line corresponds to 1/η according to the Kramers' theory.

For all three model polypeptide chains an α-value greater than 1 was found, which indicated that the intrachain diffusion coefficient decreased stronger as excepted for free diffusion from the Kramers' theory with $D \propto 1/\eta$ (Figure 4.1.2). In addition to the intrachain diffusion, the parameters b and σ provide information on the dimension

of the polypeptide chains in terms of the root mean square end-to-end distance (chapter 1.4). The fitted parameter *b* showed an increasing inaccuracy with increasing glycerol concentration, while relative low errors for the fitted parameter σ are found. Nevertheless, the root mean square end-to-end distance using *b* and σ (equation 1.12) showed a high accuracy for all glycerol concentrations. For the root mean square end-to-end distance of the DE-loop a slight increase of about 4% with increasing glycerol concentration was observed (B in Figure 4.1.3). This was also reflected in the end-to-end distribution (C in Figure 4.1.3). In previous work, a comparable slight increase of the root mean square end-to-end distance for the $(GS)_8$ and EF-loop fragment were shown [159]. The alteration of the dimension with increasing glycerol concentration may explain the deviation from the theoretically expected α-value of 1.

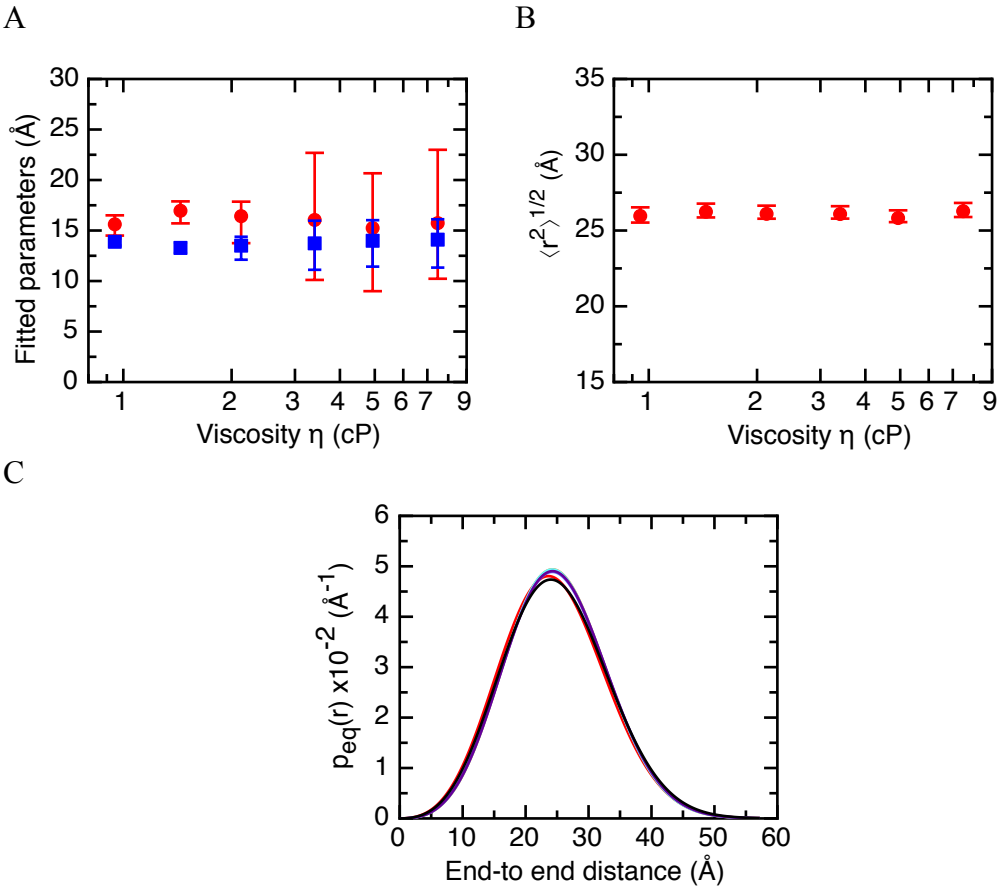


Figure 4.1.3: Glycerol dependence of dynamics and dimensions of the DE-loop of the natural IDP β-carp parvalbumin. **A** Semi-logarithmic illustration of the fitting parameters *b* (red circles) and σ (blue squares); **B** the root-mean square end-to-end distance was calculated using equation 1.12 and **C** linear plot of the end-to-end distribution $p_{eq}(r)$ at different glycerol concentrations ranges from 0% (red), 16% (orange), 28% (green), 40% (turquoise), 48% (blue) to 56% (purple) (w/w).

However, the direct relation of the intrachain diffusion coefficient to the root mean square end-to-end distance is still unclear and a matter of future studies.

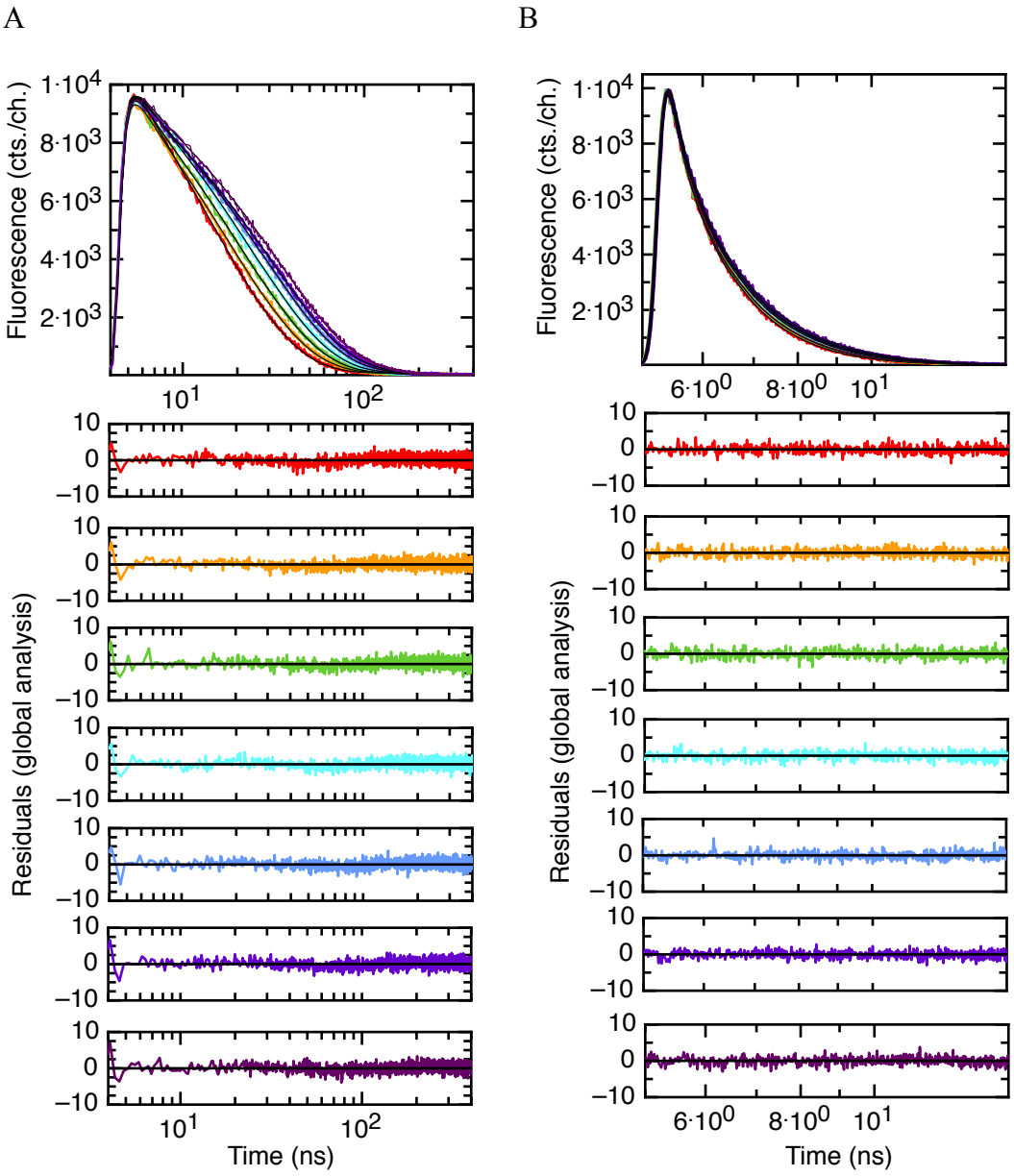


Figure 4.1.4: Time-resolved fluorescence decays curves of $(GS)_8$ in dependence of PEG_{6000} using donor-acceptor pair **A** dansyl-*pyrene* and **B** DNB-AMCA in 10 mM potassium phosphate pH 7.0 at different concentration ranges from 1.8% (red), 4.5% (orange), 7.3% (green), 10.5% (turquoise), 14.5% (blue), 17.7% (purple) to 21.1% (dark purple) (w/w). The residuals of the global analysis of each concentration point are shown in different colors corresponding to the different concentrations of PEG_{6000} for both FRET-pairs.

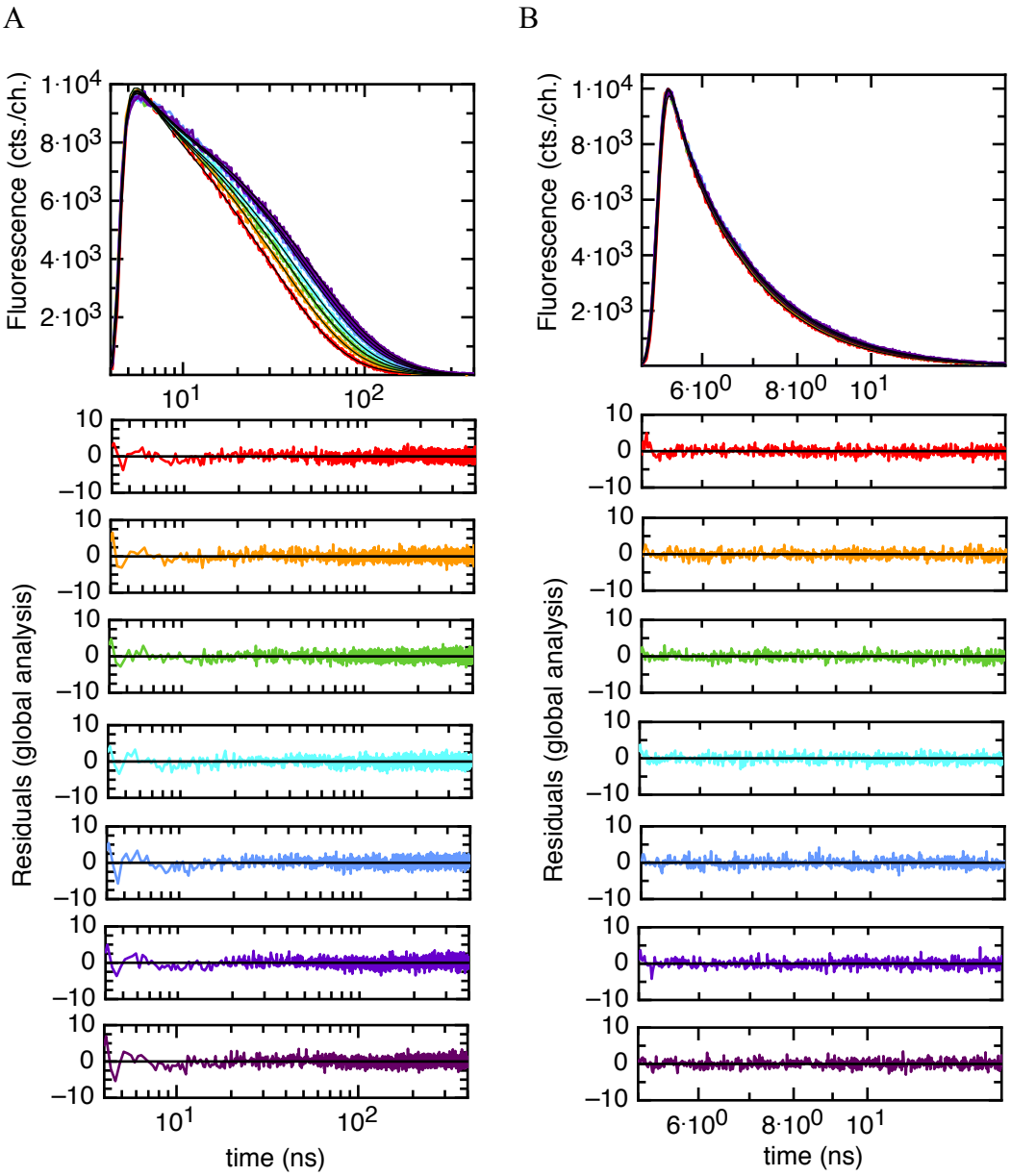


Figure 4.1.5: Time-resolved fluorescence decays curves of β-carp parvalbumin fragment EF-loop in dependence of PEG_{6000} using donor-acceptor pair **A** dansyl-*pyrene* and **B** DNB-AMCA in 10 mM potassium phosphate pH 7.0 at varies concentration ranges from 1.8% (red), 4.5% (orange), 7.3% (green), 10.5% (turquoise), 14.5% (blue), 17.7% (purple) to 21.1% (dark purple) (w/w). The residuals of the global analysis of each concentration point are shown in different colors corresponding to the different concentrations of PEG_{6000} for both FRET-pairs.

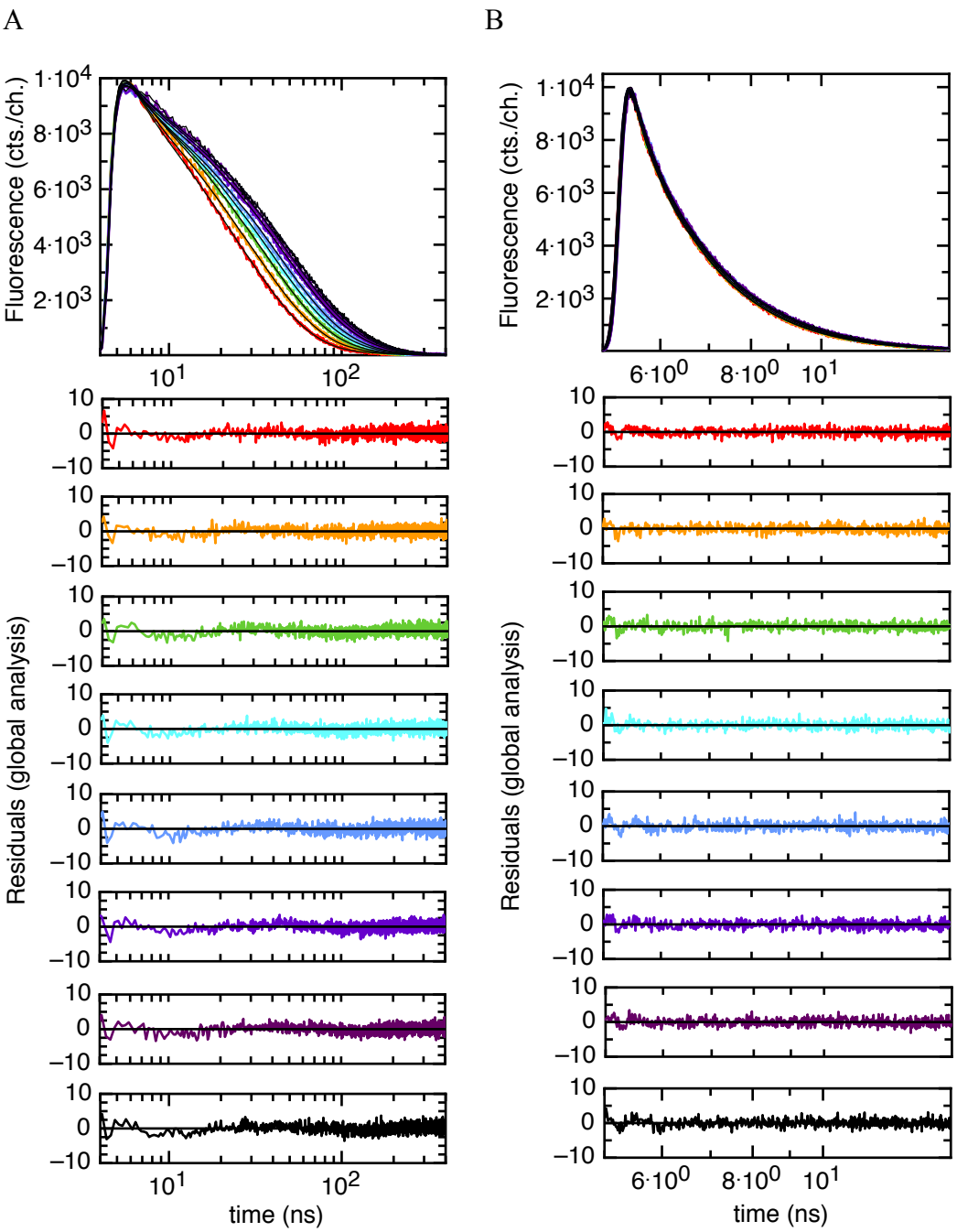


Figure 4.1.6: Time-resolved fluorescence decays curves of β-carp parvalbumin fragment DE-loop in dependence of PEG_{6000} using donor-acceptor pair **A** dansyl-*pyrene* and **B** DNB-AMCA in 10 mM potassium phosphate pH 7.0 at varies concentration ranges from 0% (red), 1.8% (orange), 4.5% (green), 7.3% (turquoise), 10.5% (blue), 14.5% (purple), 17.7% (dark purple) to 21.1% (black) (w/w). The residuals of the global analysis of each concentration point are shown in different colors corresponding to the different concentrations of PEG_{6000} for both FRET-pairs.

Nevertheless, the intrachain diffusion coefficient dependence on the glycerol concentration can be well described by the empirical power-law equation (equation 4.4), which was also shown for the rate constant of loop formation from TTET measurements for different lengths of $(GS)_8$ and the DE- and EF-loop fragment [218]. Compared to the small viscogenic co-solute glycerol, the larger PEG_{6000} exhibited a different quantitative effect on the intrachain diffusion. According to the equation 4.4 α-values of 0.4 ± 0.01 for the DE-loop and 0.375 ± 0.01 for the EF-loop were obtained. For the repetitive sequence $(GS)_8$ an α-value of 0.47 ± 0.01 was found (Figure 4.1.7).

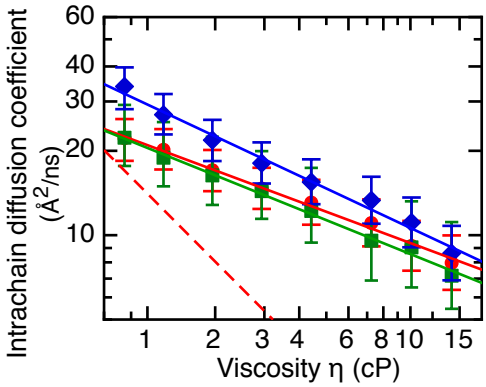


Figure 4.1.7: Double-logarithmic illustration of the intrachain diffusion coefficient in dependence of different PEG_{6000} concentrations (0% - 21.1% (w/w)). The intrachain diffusion coefficient of DE-loop is shown in green squares, of EF-loop is shown in red circles and of $(GS)_8$ is shown in blue diamonds. The results of the analysis using the empirical equation 4.4 are depicted in the colors mentioned above. The red dashed line corresponds to 1/η according to the Kramers' theory.

As observed in TTET experiments with viscogenic co-solutes of larger size a significant lower α-value was found independent of length and sequence of the model polypeptide chains [218]. The value of $\alpha<1$ indicated a strong deviation from Kramers' theory and the existence of internal friction. For all three polypeptide chains a slight increase of the root mean square end-to-end distance with increasing PEG_{6000} concentration was observed. In detail, for the DE-loop fragment an increase of 2%, for the EF-loop fragment of 4% and for $(GS)_8$ of 3% in 21.1% (w/w) PEG_{6000} were obtained in comparison to water (C in Figure 4.1.8).

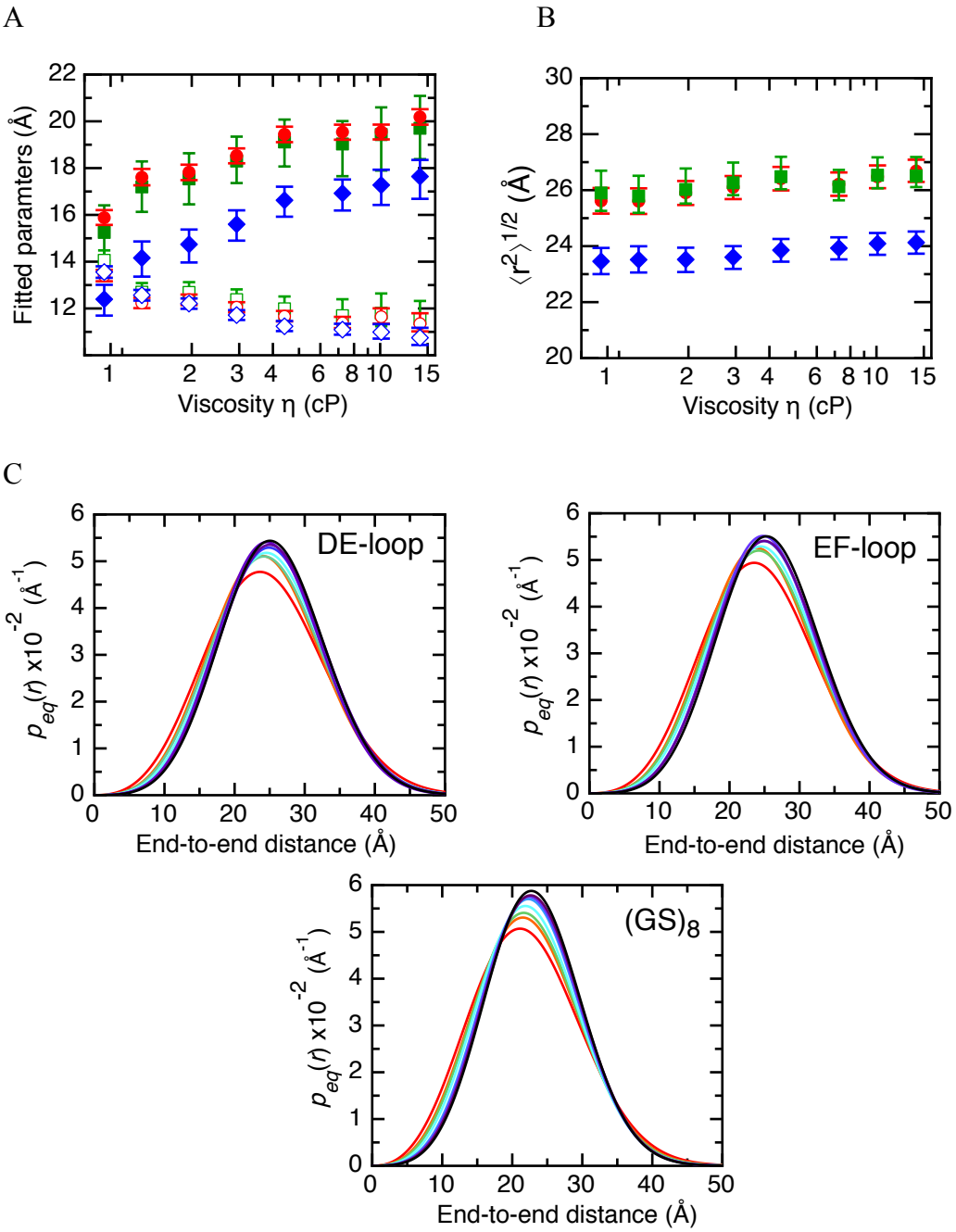


Figure 4.1.8: PEG_{6000} dependence of dynamics and dimensions of the DE- and EF-loop of the natural IDP β-carp parvalbumin as well as $(GS)_8$. **A** Semi-logarithmic illustration of the fitting parameters *b* (filled) and *σ* (open), **B** The root-mean square end-to-end distance was calculated using equation 1.12. The DE-loop is shown in green squares, the EF-loop is shown in red circles and the $(GS)_8$ is shown in blue diamonds and **C** linear illustration of the end-to-end distribution $p_{eq}(r)$ at different PEG_{6000} concentrations ranges from 0% (red), 1.8% (orange), 4.5% (green), 7.3% (turquoise), 10.5% (blue), 14.5% (purple), 17.7% (dark purple) to 21.1% (black) (w/w).

With the assumption of the Einstein-Smoluchowski equation, the intrachain diffusion coefficient can be directly converted into the friction coefficient (equation 4.3). The resulting dependence of the friction coefficient on the viscosity of the two different sizes of the viscogenic co-solutes was analyzed with equation 4.5. The equation 4.5 represents a combination of both above discussed models, the additive (equation 4.1) and the power-law model (equation 4.2). With a theoretical $\alpha = 1$ for equation 4.5, the above discussed simple additive model 4.1 is generated. On the other hand, with introducing the experimental determined α-value ($\alpha \neq 1$), the empirical found power-law model with an additional term of the internal friction is similar to equation 4.2. This additive term of the internal friction depicts the influence of internal friction in the case of $\alpha \neq 1$.

$$f = f_{solv}\left(\frac{\eta}{\eta_0}\right)^{\alpha} + f_{\text{int}} \qquad \text{(Eq. 4.5)}$$

The dependence of the friction coefficient on the viscosity can be well described with equation 4.5 for both viscogenic co-solutes glycerol and PEG_{6000}. If the α-value was set to the theoretical value of one, which is consistent with the additive model (equation 4.1), a high inaccuracy is observed (A and B in Figure 4.1.9). Moreover, with $\alpha = 1$, for the glycerol dependence a negative internal friction for all polypeptide chains are observed (Figure 4.1.9 right plot in A, and Table 2). Instead, with the experimentally determined α, the additional internal friction term yields a value of around zero for all polypeptide chains. However, in case of the PEG_{6000} dependence and with $\alpha = 1$, the internal friction term becomes positive. For the experimentally determined α-value of the PEG_{6000} dependence an internal friction term of around zero was shown as also observed for the glycerol dependence (Table 2).

A

B

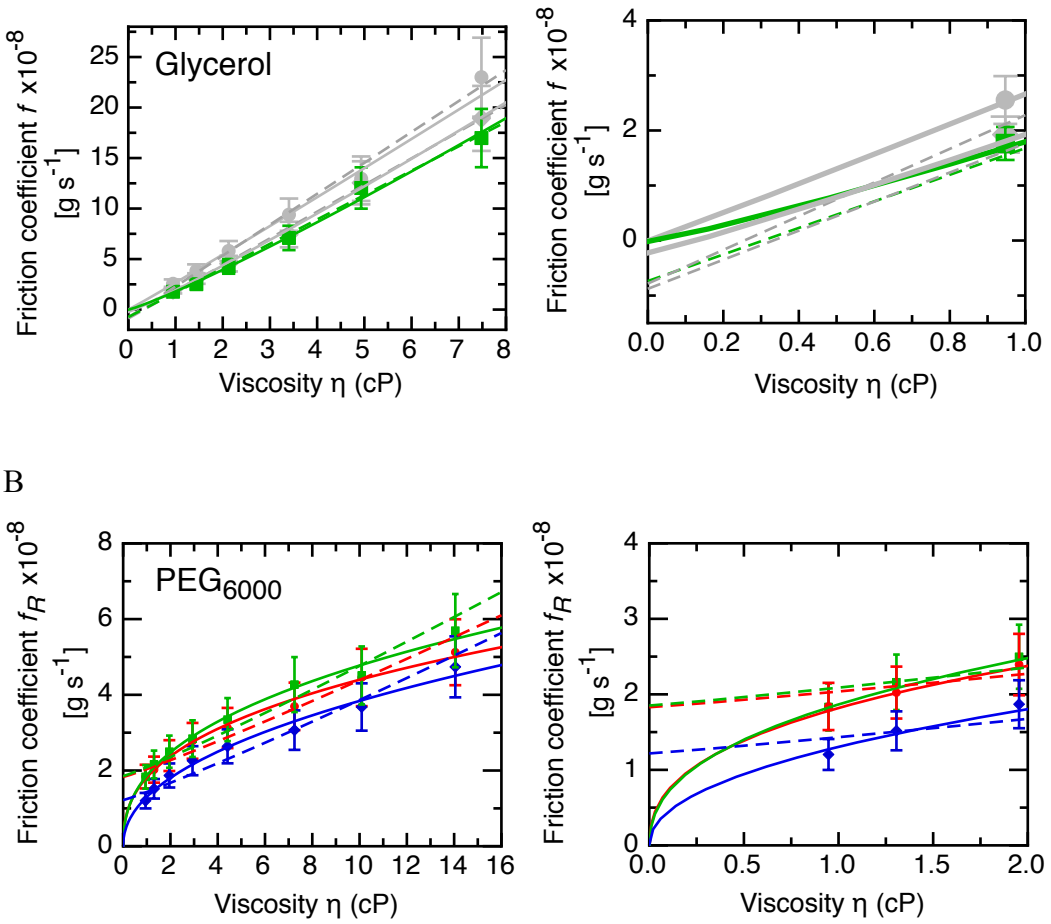


Figure 4.1.9: Viscosity dependence of the friction coefficient. **A** The glycerol dependence for the DE-loop fragment is shown in green, data of the EF-loop fragment as well as $(GS)_8$ are taken from [159] and is shown in grey. **B** The PEG_{6000} dependence for the DE-loop is shown in green, for the EF-loop fragment is shown in red and for the $(GS)_8$ chains is shown in blue. The friction coefficient was calculated from the intrachain diffusion coefficient according to equation 4.3. The friction coefficient was analyzed using equation 4.5 and the corresponding α-value determined from equation 4.4 (solid lines). With α = 1 the equation was reduced to the additive relationship of the solvent and internal friction (equation 4.1, dashed lines). The right plots in **A** and **B** (zoom) give a detailed picture at the lowest viscosity for glycerol and PEG_{6000} dependence.

The solvent friction was comparable for both viscogenic co-solutes and generally yields a lower value for the flexible $(GS)_8$ chain (Table 2). At α = 1 for the glycerol dependence, the solvent friction becomes higher in comparison to that fitted with the experimentally determined α-value (Table 2). The opposite behavior was observed for the PEG_{6000} dependence.

Table 2: Viscosity dependence of glycerol and PEG_{6000} on polypeptide chains. The friction coefficient was calculated from the intrachain diffusion coefficient using equation 4.3. The resulting viscosity dependence on the friction coefficient was analyzed using equation 4.5. If α = 1 the additive model of the solvent and the internal friction (equation 4.1) is valid. If the experimentally determined α-value (α = exp.) was applied, the power-law model is valid with an additional term of the internal friction (equation 4.2) to yields some information on the internal friction in dependence of α ≠ 1. The unit for the solvent and for the internal friction coefficient of every value is given as $x10^{-8}$ g/s. The data of the glycerol dependence of the EF-loop and $(GS)_8$ are taken from previous work [159] (grey).

	Glycerol				**PEG_{6000}**			
	α = exp.		*α=1*		*α = exp.*		*α=1*	
	f_{solv}	f_{int}	f_{solv}	f_{int}	f_{solv}	f_{int}	f_{solv}	f_{int}
DE-loop	1.70 ±0.2	-0.17 ±0.4	2.2 ±0.3	-0.6 ±0.4	1.8 ±0.3	0.02 ±0.5	0.29 ±0.05	1.8 ±0.2
EF-loop	1.48 ±0.2	0.07 ±0.5	1.6 ±0.3	-0.03 ±0.5	1.8 ±0.3	0.03 ±0.5	0.25 ±0.05	1.8 ±0.2
$(GS)_8$	1.15 ±0.2	0.03 ±0.4	1.4 ±0.2	-0.2 ±0.4	1.3 ±0.2	0.01 ±0.3	0.26 ±0.04	1.1 ±0.2

4.1.4 Discussion and summary

It can be shown that the viscosity dependence on the intrachain diffusion coefficient is well described by the empirical equation 4.4 independent of the size of the viscogenic co-solutes. Furthermore, the intrachain diffusion coefficient is converted into the friction coefficient by the Einstein-Smoluchowski equation and the resulting dependence of the friction coefficient on the viscosity can be well fitted according to the power-law model. Thereby, the additive term of the internal friction in equation 4.5 will become zero for both types of co-solutes and for all investigated polypeptide chains, if the experimentally determined α-value is included. Thus, the effect of the internal friction is represented by the viscosity-sensitive exponent α of the power-law equation and cannot be described by an additive internal friction term. Furthermore, it can be shown that the α-value depends strongly on the size of the used viscogenic co-solutes. Hence, with small co-solutes the predicted Kramers-like behavior $D \propto 1/\eta$ is found and for larger co-solutes a deviation from this theory with α<1 is observed. Detailed investigation of the dependence of the rate constants of loop formation of unfolded polypeptide chains on different sizes of the co-solute revealed the same influence of the co-solute's size on chain dynamics [218].

In other words for the small co-solute glycerol the macroviscosity matches the microviscosity, which means the chain dynamics feel the full solvent viscosity. However, for co-solutes with larger size, like PEG_{6000}, the microviscosity does not match the macroviscosity. This fact is related to the global or large scale chain motions and the local chain motions, which show different correlation to the co-solute. Local chain motions were shown to be less sensitive to the solvent resulting from local torsion angle barriers [212], [231]–[234]. The deviation from the Kramers' theory was discussed to be mainly caused by bond rotations [235], [236]. Also hydrogen bonds were proposed to be one reason for reduced sensitivity of local chain motions to a bulky solvent [216], [237]. Another reason for the deviation of local chain motion from the simple $1/\eta$ relationship was given by the hydration-shell, controlling local and fast fluctuations instead of the large-scale motions, which are coupled directly to a bulky solvent [238], [239]. Additionally, the so-called solvent memory effects is also discussed to be an origin of the internal friction and $\alpha<1$, respectively [234], [240], [241]. In general, the classic viscosity is valid under the assumption of a homogenous and isotropic bulky solvent, where the Brownian theory is applicable. But for co-solutes, which are larger as the polypeptide chains, such as the investigated polypeptide chains in PEG_{6000}, the solution in which the polypeptide chains is dissolved is no longer considered as a continuous and homogenous solution. Large co-solutes may interact only locally with the solute. Thus, the solvent is inhomogeneous for the peptide on a microscopic scale and alters the viscosity sensitivity of the dynamics, which results in $\alpha<1$ [218], [222], [242], [243]. Additionally, it should be taken into account that larger viscogenic co-solutes like the commonly used PEG show a polymer length distribution. DLS measurements of different concentration of PEG_{6000} showed an increase of the apparent hydrodynamic radius with a decreasing co-solute concentration in a nonlinear manner, which illustrated the change of physical properties with changing concentration of the co-solute (Appendix, Figure A.1). In summary, the viscosity dependence of the intrachain diffusion coefficient can be described using the empirically found equation 4.4, where a viscosity-sensitive factor α can be determined. It was shown that the exponent α includes the so-called internal friction and that the additive term of the internal friction cannot describe the experimental data. The co-solute glycerol seems to be suitable for the investigation of chain dynamics' dependence on viscosity

without additional interferences like the discussed inhomogeneity found for larger co-solutes such as PEG_{6000}.

The knowledge about the viscosity effect on the intrachain diffusion coefficient enables the investigation of other important effects *e.g.* caused by denaturants or changed temperature.

4.2. Dynamics and dimensions of Q_{14} in aqueous solution

As discussed in the introduction, the polyglutamine chains represent interesting members of the unstructured IDPs, which are responsible to cause different neurogenic diseases like Chorea Huntington. In order to identify a possible intrinsic disease-related characteristic with respect to the dynamics and dimensions of the polyglutamine chains, a polyglutamine chain of 14 amino acids (Q_{14}) was chosen and compared to previously measured non-amyloidic polypeptide chains of comparable length (Figure 4.2), such as the EF-loop of the natural IDP β-carp parvalbumin and the $(GS)_8$ chains. This specific length was chosen because longer peptides tend to aggregate and complicate data interpretation. Walters and Murphy showed with dynamic light scattering (DLS) and transmission electron microscopy (TEM) experiments that aggregation occurred already at a length of 16 amino acids, while shorter ones did not aggregate [244]. Nevertheless, the peptides under investigation should be as long as possible with regard to the disease-inducing threshold value for huntingtin (chapter 2) in order to better illustrate disease-related behavior of polyglutamine at the monomeric level. Generally, little is known about the real population and dynamics of the unstructured polyglutamines in solution. In previous work, a compact structure in solution was observed for polyglutamine monomers. Simulations yielded also a relatively compact structure shown for polyglutamine chains of different length [245]. Atomic force stretching and SAXS experiments on polyglutamine chains were also consistent with a compact structure of polyglutamine monomers in water [187], [188]. Nevertheless, the previous methods and results failed to give detailed information of the real conformer distribution of unfolded polypeptide chains. The time-resolved FRET method is an appropriate method to get detailed information about the population of the conformer ensemble as well as

intrinsic chain dynamics of the polyglutamine chain monomers in aqueous solution. Especially, the description of the intrachain dynamics should contribute to a biophysical better understanding of the disease-related behavior of polyglutamines compared to non-amyloidic polypeptide chains.

4.2.1 Preparation and validation of monomeric Q_{14} in aqueous solution

In order to get valid results on the dimensions and dynamics of the investigated polypeptide chains, a homogenous solution of monomeric polyglutamine is indispensable. Wetzel and colleagues developed a disassembly protocol to yield homogenous solutions of polyglutamine monomers also in sufficient concentration [183], [246] (Material and Methods 6.3). For the trFRET experiments a $(glutamine)_{14}$ polypeptide chain (Q_{14}) with a solubility-tail of serine-arginine-serine-arginine-glycine was coupled at the C-terminus of the polyglutamine chains (Figure 4.2).

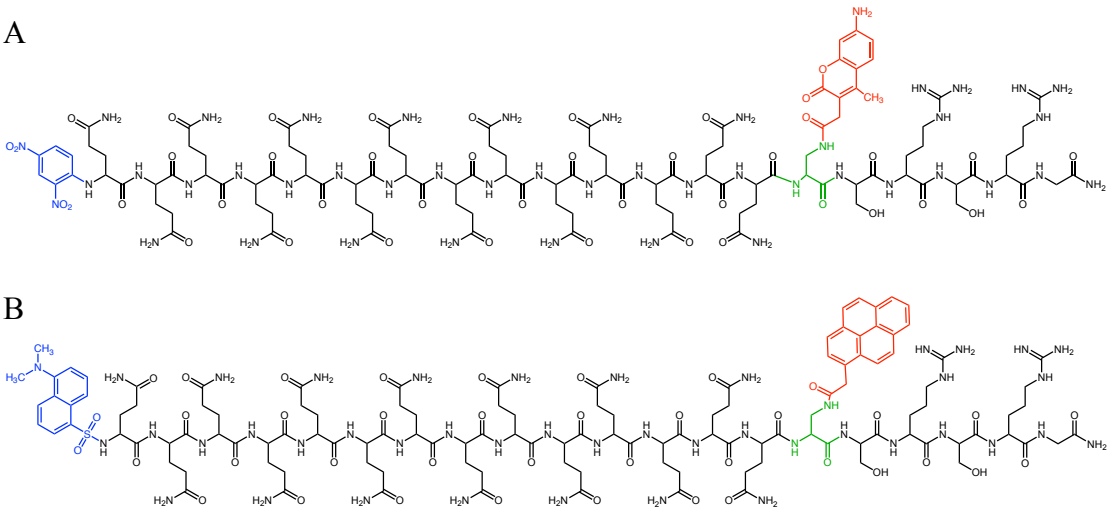


Figure 4.2: Scheme of the $(glutamine)_{14}$ polypeptide chains (Q_{14}) for the time-resolved FRET measurements. The FRET-acceptor (**A** DNB or **B** dansyl, colored in blue) was coupled at the N-terminus, while the FRET-donor (**A** AMCA or **B** *pyrene*, colored in red) was coupled at the NH_2-side chain group of a non-canonic amino acid called Dpr (green). To improve solubility a solubility tail of serine-arginine-serine-arginine-glycine was coupled at the C-terminus. The C-terminus was protected with a NH_2-group to avoid a charge end.

In order to prove that a monomeric and homogenous polyglutamine chains with both FRET chromophores and the solubility tail can be prepared according to the disassembly protocol established by Wetzel and colleagues, three different methods with different physical background were used. At first, simple scattering experiments were performed to test the prepared solution for possible larger aggregates or higher order species using a conventional fluorimeter. The second method, analytical ultracentrifugation (aUZ) is based on absorption or fluorescence instead of light scattering. It represents a method that offers higher accuracy for the discrimination of higher order species (aggregates) compared to simple scattering experiments. The third method is far UV circular dichroism, which is used to gather information of aggregates and also of general structure.

4.2.1.1 Scattering experiments

For the scattering experiments a conventional fluorimeter was used. In order to get an adequate scattering signal and to exclude any absorbance of the coupled chromophores, the monochromators were set to 450 nm for dansyl-Q_{14}-pyrene and 550 nm for DNB-Q_{14}-AMCA (Figure 4.2.1).

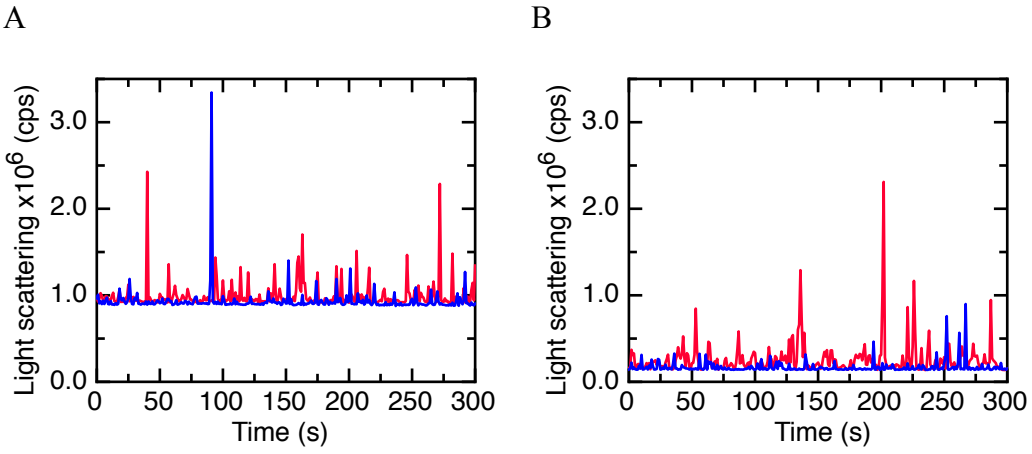


Figure 4.2.1: Light scattering profile of different Q_{14}. **A** Profile of dansyl-Q_{14}-pyrene (red) and 10 mM potassium phosphate buffer, pH 7.0 (blue). **B** Profile of DNB-Q_{14}-AMCA (red) and 10 mM potassium phosphate buffer, pH 7.0 (blue). The monochromators were set to 450 nm for dansyl-Q_{14}-pyrene and 550 nm for DNB-Q_{14}-AMCA.

The observed basic signal of the peptide signal for both peptides were not significantly higher than the buffer signal and ranges within the method error (Figure 4.2.1). These results allowed the first conclusion that no aggregated polypeptide chains were present in the solution. Nevertheless, for a quantitative comparison, an average value over the whole measurement time was calculated. The average value for the DNB-Q_{14}-AMCA polypeptide chain was around 34 % higher as the buffer average value. For dansyl-Q_{14}-pyrene polypeptide chain only a 6 % higher average value was observed in comparison to the buffer signal. Of course the average value was strongly influenced by the counted spikes, which may artificially increase the average value. The observed spikes represent undefined light-scatter centers. It cannot be completely ruled out that the spikes represent single aggregated particles, but more likely they are air bubbles or incorporated dust particles. It should be noted that all peptide solutions were prepared according to the published disassembly protocol by Wetzel and colleagues without filtration after preparation compared to the buffer that was filtrated. Therefore, the solution may contain more dust particles in comparison to the filtrated buffer. Of course, the simple scatter experiment using conventional fluorimeter represents a relative crude method for analysis of aggregation processes or oligomers. Therefore, other methods should also be used to verify the previous results. A powerful method to analyze the samples for aggregates or oligomers is the analytical ultracentrifugation (aUZ).

4.2.1.2 Analytical ultracentrifugation

By analytical ultracentrifugation the investigated sample is strongly centrifuged at high speed and particle sedimentation is observed by a specified absorption signal. Using analytical ultracentrifugation oligomers or aggregates can be determined by a sedimentation velocity experiments (A and B in Figure 4.2.2). A detection limit for oligomers or aggregates is to be estimated by around 5 %. In addition to the sedimentation velocity experiment, the molecular mass can be determined by the sedimentation equilibrium experiment (C and D in Figure 4.2.2).

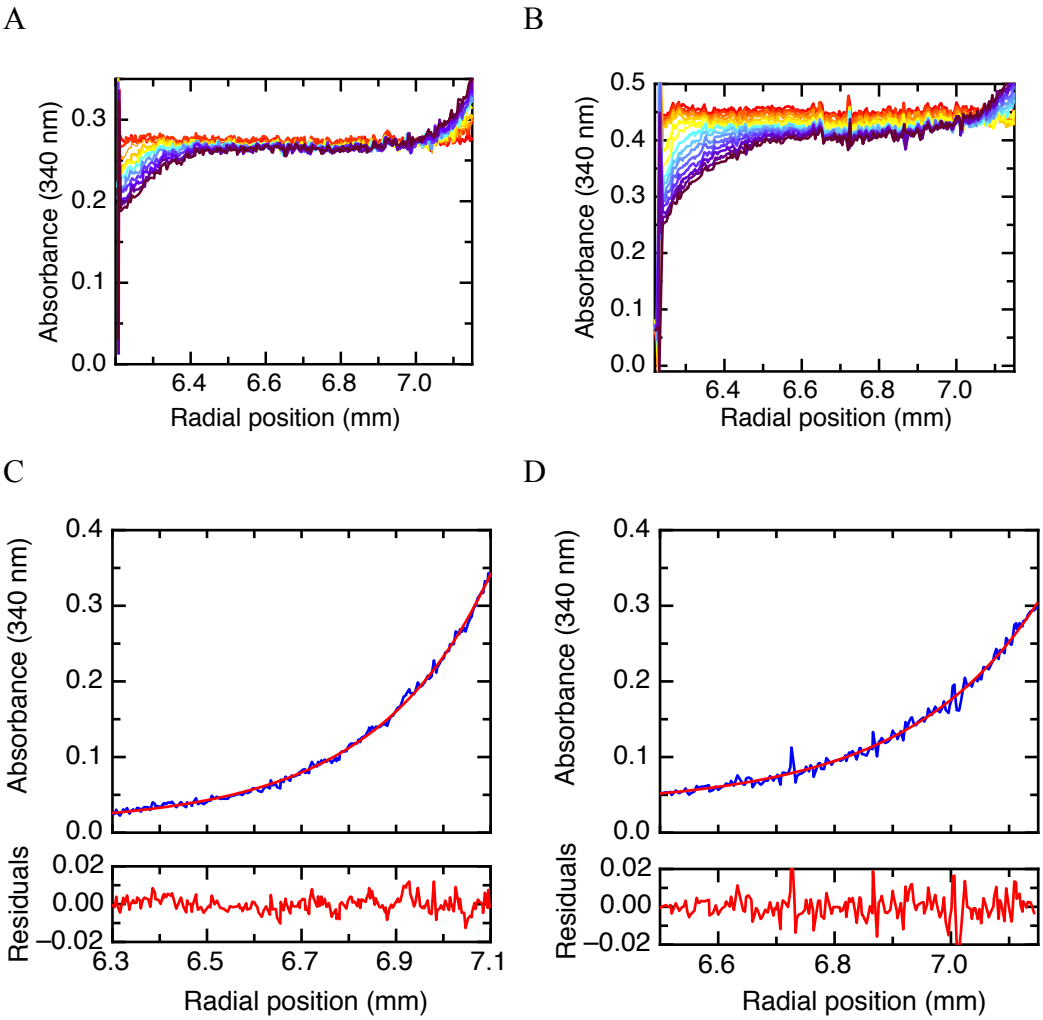


Figure 4.2.2: A-B: Sedimentation velocity experiments of **A** dansyl-Q_{14}-pyrene and **B** DNB-Q_{14}-AMCA. Time scan every 10 min (colors in rainbow from red to purple). **C-D:** Sedimentation equilibrium experiment. Analysis of the molecular weight (red lines, analysis provided using the company software) of **C** dansyl-Q_{14}-pyrene and **D** DNB-Q_{14}-AMCA with the resulting residuals at the bottom. All samples were measured at 22.5 °C, pH 7.0 in 10 mM potassium phosphate. The analytical absorption wavelength of both polypeptide chains was set at 340 nm, which corresponded to the absorbance wavelength of both coupled chromophores. The measurements were performed and analyzed by Prof Dr. Lilie (Martin-Luther-University Halle Wittenberg, Institute of Biochemistry and Biotechnology).

The analysis by aUZ showed no oligomers or aggregates for dansyl-Q_{14}-pyrene and for DNB-Q_{14}-AMCA polypeptide chain as depicted in Figure 4.2.2. The molecular mass for dansyl-Q_{14}-pyrene was 2890 ± 60 Da and for DNB-Q_{14}-AMCA
2880 ±160 Da, which corresponded to the monomeric state for both polypeptide chains compared to theoretically calculated molecular mass (C and D in Figure 4.2.2).

4.2.1.3 Circular dichroism

Circular dichroism (CD) is primarily used to investigate the structure of proteins or polypeptide chains. CD can provide information about aggregates or oligomers, because most aggregation is caused by unspecific secondary structure interactions, which changes the observed CD signal dramatically. A minimum of ellipticity at around 200 nm was published for polyglutamine chains at the monomeric state [164], [244]. On aggregation of polyglutamine the minimum at 200 nm disappeared. Here, CD was performed to detect aggregates and to investigate the general structure of Q_{14}. Besides the both FRET-pairs also the donor- as well as the acceptor-only variants of Q_{14} was measured, because these are used for the determination of the Förster distances (Figure 4.2.3). The measured CD spectra were not converted into the mean residual ellipticity, because there are no accurate extinction coefficients published for the AMCA- and *pyrene*-chromophores. However, the ellipticity yields adequate information for the investigation of random structures and the state of a polypeptide chain. For all measured Q_{14} a minimum of the ellipticity at around 200 nm was found. The observed minimum closed to 200 nm is consistent with the published data of polyglutamines of different length and can be interpreted as a property of their unordered structure [164], [185], [244]. This features is also consistent with published theoretical and experimental data from another unfolded proteins or polypeptide chains [247]–[255]. Interestingly, the observed additional weak negative band at around 220 nm is often measured for unfolded polypeptide chains, but not further interpreted [185], [244], [249]–[252], [255], [256]. Furthermore, there were no indications of aggregation for all measured Q_{14} variants, which were expected to yield a positive signal at around 200 nm [257].

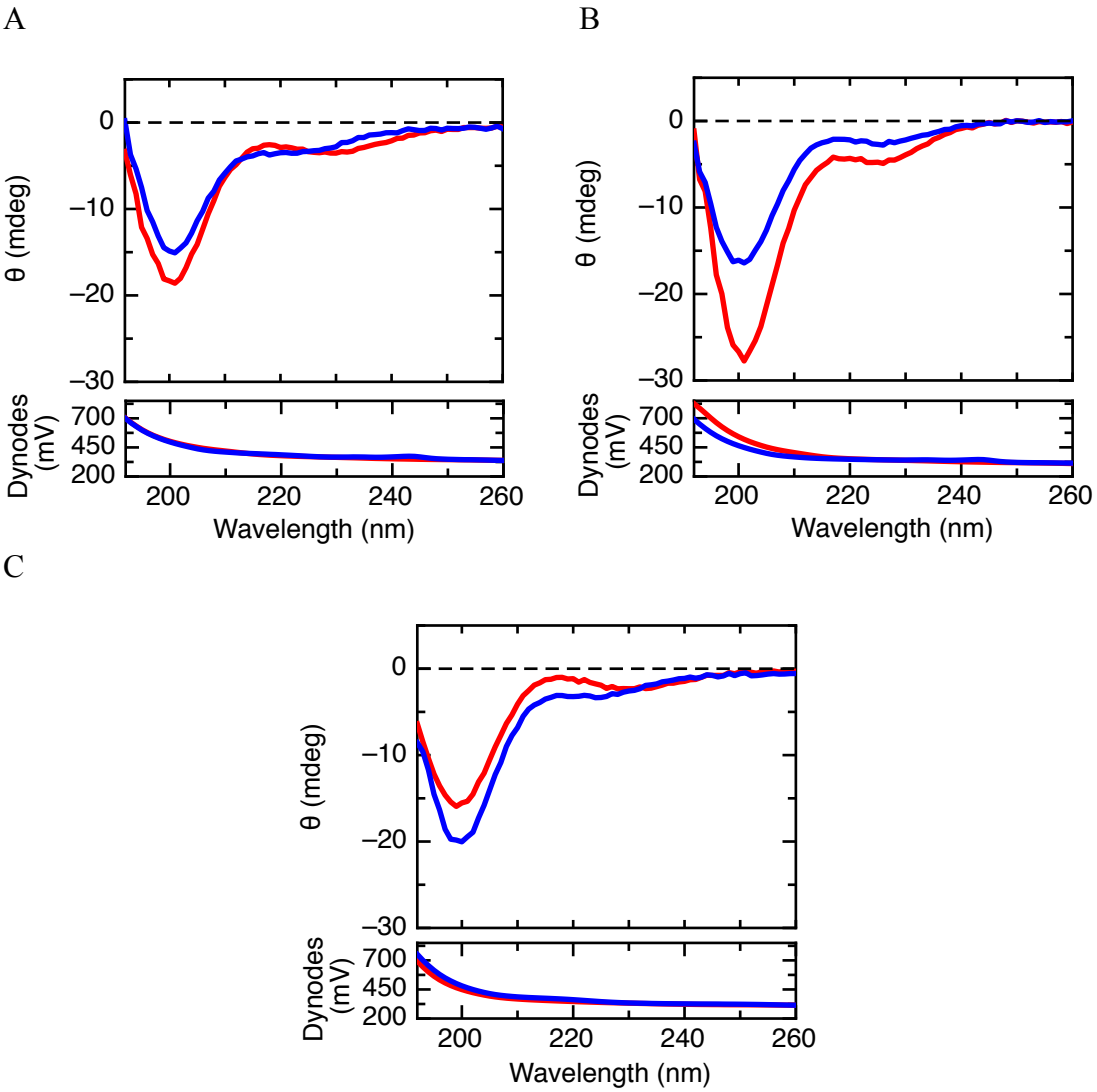


Figure 4.2.3: CD spectra of Q_{14}. **A** Q_{14} coupled with both FRET pair chromophores, dansyl-Q_{14}-pyrene (blue line) and DNB-Q_{14}-AMCA (red line). **B** Donor-only Q_{14} coupled with pyrene (blue line) and AMCA (red line). **C** Acceptor-only Q_{14} coupled with dansyl (blue line) and DNB (red line). The dynodes voltage of each measurement is plotted in the corresponding colors at the bottom.

4.2.1.4 Summary and discussion

The preparation based on the disassembly protocol published by Wetzel and colleagues [183], [246] represent a necessary method to obtain a sufficiently high concentration of polyglutamine chains at monomeric level for the trFRET experiments. Generally, no aggregation of all used variants of Q_{14} were found. Neither the scattering experiments, nor the CD spectra gave evidence for aggregation of the Q_{14} variants.

Furthermore, the analytical ultracentrifugation indicated a homogenous monomeric solution of polypeptide chains. All three methods for the investigation of aggregates were performed with comparable or slightly higher polypeptide chain concentrations that were used for the trFRET experiments. Also important is the fact that a random structure of all variants of Q_{14} was observed by CD. Hence, the distance distribution of random polymers described by the EDWARDS model should be valid (chapter 1.3).

4.2.2 Investigation of Q_{14} in aqueous solution with trFRET

In order to calculate the partial differential equation (equation 2.1, chapter 1.5) given by Haas *et al.* [99], the determination of constants such as the Förster distances and the donor lifetime are essential.

4.2.2.1 Donor-lifetime measurements of both Ac-Q_{14}-pyrene and Ac-Q_{14}-AMCA

Time-resolved lifetime experiments of Ac-Q_{14}-AMCA and Ac-Q_{14}-pyrene in 10 mM potassium phosphate buffer pH 7.0 at 22.5 °C were performed to determine the donor lifetime of both donors (Figure 4.2.4). A reconvolution fit of the fluorescence decay curves was used to calculate the donor-lifetime (Material and Methods 6.5). In previous works, for donor-only polypeptide chains only one lifetime was found. Therefore, a single-exponential reconvolution fit was applied. The donor lifetime of the AMCA-only polypeptide chain could be well described by a single-exponential reconvolution fit, deduced from the analysis error χ^2 and resulting residuals, respectively (B in Figure 4.2.4). However, for the *pyrene*-only polypeptide chains no well-defined analysis were obtained by a single-exponential reconvolution fit (A in Figure 4.2.4 red line and corresponding residuals).

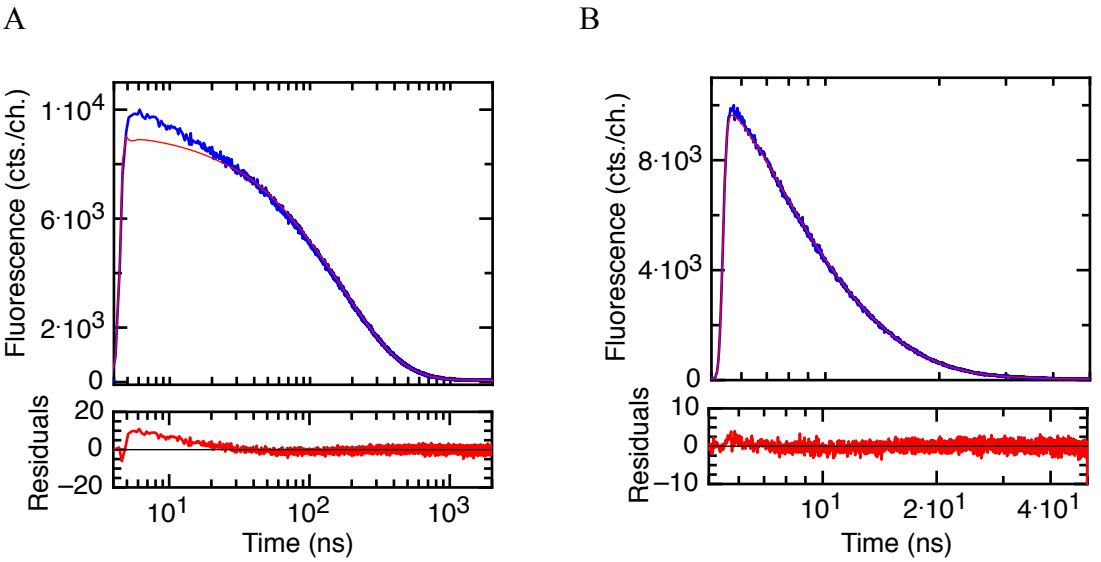


Figure 4.2.4: Fluorescence decay curves of the donor-only polypeptide chain of Q_{14}variants. **A** Ac-Q_{14}-pyrene and **B** Ac-Q_{14}-AMCA (blue lines) analyzed using the reconvolution fit (red lines, Material and Methods 6.5). The resulting residuals are shown in red at the bottom.

Therefore, the donor-only fluorescence decay curve of the *pyrene*-only variant was fitted using double-exponential reconvolution fit and displayed a high accuracy with the original data (Figure 4.2.5). The main lifetime of the double-exponential fit of 168 ± 1 ns was in agreement to the lifetime of a single-exponential fit and showed fractional amplitudes of about 87%, whereas the additional second lifetime of 14 ± 2 ns had a fractional amplitude of about 13%.

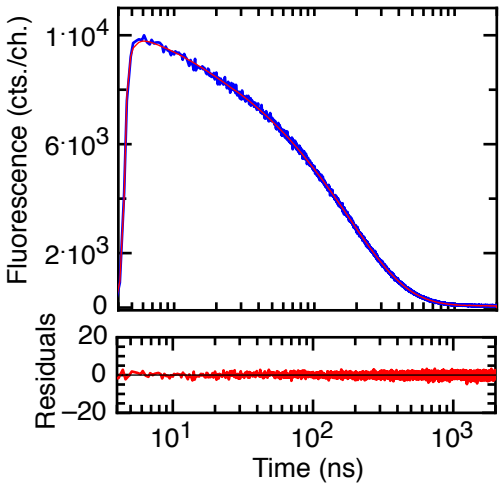


Figure 4.2.5: Fluorescence decay curves of Ac-Q_{14}-pyrene (blue line) analyzed by a double-exponential re-convolution fit (red lines). The resulting residuals are shown in red at the bottom.

To verify whether the smaller fractional amplitude of the secondary lifetime was real-valued, a so-called *Support Plane Analysis* (*SPA*) (Material and Method 6.6.3), which has often been used for lifetime measurements, was calculated for both lifetimes (Figure 4.2.6). The shape and the width of the resulting *SPA* give evidence for the quality of the fit.

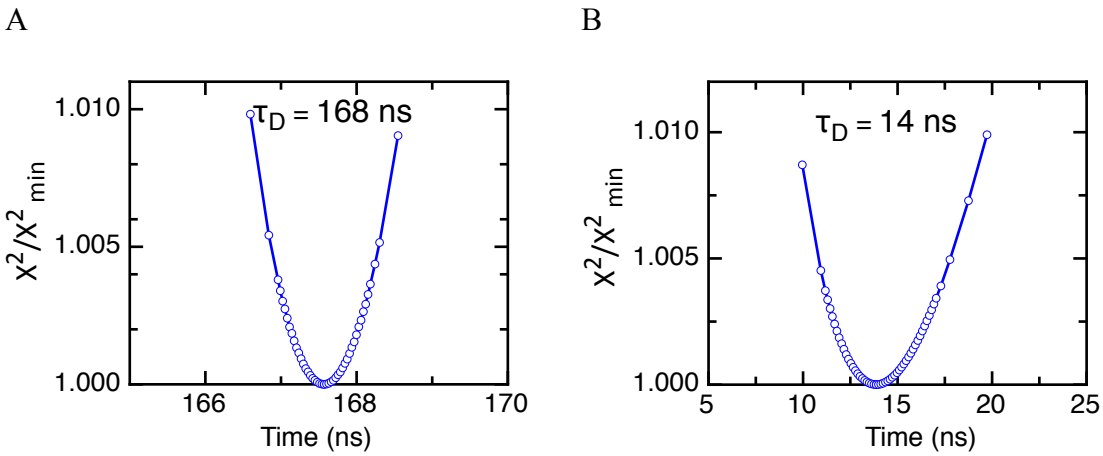


Figure 4.2.6: *Support Plane Analysis* of the fluorescence decay curve of Ac-Q_{14}-pyrene for **A** the main lifetime of 168 ns and **B** the secondary lifetime of 14 ns.

The *SPA* implied that the main (biggest amplitude) lifetime of 168 ns as well as the secondary lifetime of 14 ns were well defined with a small error. However, an important question remained about the origin of the secondary lifetime found for the *pyrene*-only Q_{14} variant.

4.2.2.2 Investigation of the second lifetime with Ac-Q_3-pyrene

In order to answer the question of the origin of the secondary lifetime found, we analyzed whether the unexpected secondary lifetime of around 14 ns was a possible result of the disassembly procedure of the polyglutamine chains. Therefore, a shorter polypeptide chain with higher solubility and less (expected) aggregation tendency was used. Thus, a solution of peptide with sufficient concentration could be prepared without the disassembly protocol. For this, a *pyrene*-donor-only polypeptide chain of three glutamine repeats was synthesized.

In order to avoid any aggregation during the measurements, the denaturant guanidine hydrochloride (GdmCl) was used as additive. Fluorescence decay curves of Ac-Q_3-pyrene peptide chain with and without the disassembly procedure were prepared at 1 M and 8 M GdmCl to test the influence of the denaturant on the secondary lifetime.

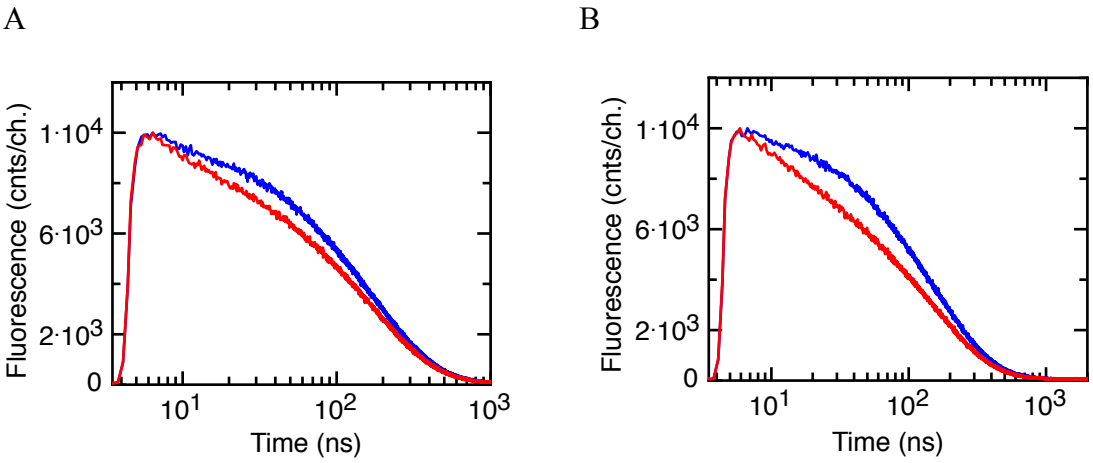


Figure 4.2.7: Fluorescence decay curves of Ac-Q_3-pyrene prepared with (red line) and without (blue line) the disassembly procedure and measured in **A** 1 M GdmCl and **B** 8 M GdmCl buffer in 10 mM potassium phosphate buffer at 22,5 °C, pH 7.0.

As shown in Figure 4.2.7 A for 1 M GdmCl, a secondary lifetime was observed for the sample that was prepared with the disassembly procedure. Fluorescence decay curve of Ac-Q_3-pyrene that was prepared without the disassembly protocol gave no evidence for a secondary lifetime of 14 ns. In order to exclude any possibility of aggregation during the measurements, fluorescence decay curves of Ac-Q_3-pyrene at 8 M GdmCl were monitored. As observed at 1 M GdmCl, the secondary lifetime was only detected on samples that were prepared with the disassembly procedure. The second lifetime amounted to 11 ± 2 ns at 1 M GdmCl and to 14 ± 2 ns at 8 M GdmCl. The secondary lifetime slightly differed between the GdmCl concentrations, but was within the range of the fitting error and comparable with the secondary lifetime found for Q_{14}. A slight increase of the fractional amplitude of the secondary lifetime from 18% at 1 M GdmCl to 24% to 8 M GdmCl was observed. Aggregation during the experiment could be completely excluded, because it is generally accepted that no aggregation is possible at such a high concentration of 8 M GdmCl. Additionally, the fractional amplitude of the secondary lifetime is significantly higher at 8 M GdmCl

compared to 1 M GdmCl. In conclusion, the procedure of the disassembly protocol itself seemed to be the origin of the secondary lifetime. An important step for the disassembly protocol is the incubation of trifluoroacetic acid (TFA) overnight. Chemical alterations of the polypeptide chain were conceivable during this procedure. Liquid-chromatography mass spectrometry (LC-MS) was used to investigate chemical alterations of the polypeptide chains during the disassembly protocol. No mass change of the polypeptide chains with accuracy of 1 Da was detected for the samples, which were prepared with the disassembly protocol and without. Although both trifluoroacetic acid (TFA) and hexafluoro-2-propanol (HFIP) should be completely removed during the disassembly procedure (Materials and Methods 6.3), the influence of the two compounds as an additive in solution was tested. Fluorescence decay curves of Ac-Q_3-pyrene in 10 mM potassium phosphate buffer were measured in the presence of TFA, HFIP and both chemicals together. Neither the addition of TFA-solution of up to 1% (v/v) or HFIP up to 1% (v/v), nor the addition of both TFA and HFIP to the solution resulted in a change of fluorescence decay curves or any lifetimes. The origin of the secondary lifetime is in the disassembly procedure itself. A chemical change or influence of incomplete removed chemicals does not seem to be the reason for the secondary lifetime.

Time-dependent fluorescence anisotropy experiments of *Atto*-dyes labeled Aβ peptides revealed an additional slow anisotropy decay component, which was caused by chromophore-peptide interactions [258]. Such an interaction cannot be completely ruled out for the time-resolved fluorescence experiments, but seems unlikely due to the shorter linker and the smaller sizes of chromophores. To date, the molecular origin of the secondary lifetime is not finally enlightened. However, the secondary lifetime was also observed in the FRET lifetime experiments, where an acceptor was present. Subsequent experiments showed that the secondary lifetime did not change by the alteration of temperature, viscosity and denaturant concentration, while the main lifetime depended strongly on these factors, as expected. Thus, the secondary lifetime seem to be not affected by the acceptor or other quenching processes. Additionally, Lakowicz *et al.* discuss the possibility of multi-exponential donor-only decays and suggest a well accepted assumption that additional found lifetimes show the same R_0 [158]. Steiner *et al*. found also multi-exponential donor-only decays for *dansylaziridine* linked to troponin C as a donor [259]. They calculated for all donor-only lifetimes a corresponding R_0 and compared these results to the results computed

with the assumption of similar R_0 for all found donor-lifetimes. Just a slightly broader distribution was observed with multiple R_0 with a slightly increase of χ^2 [259]. Therefore, no influence of the secondary lifetime on the FRET efficiency as well as on the global analysis is expected.

It should be noted that the addition of a second population of end-to-end distances as a result of energy transfer of the secondary donor lifetime to the acceptor cannot result in well defined global analysis, which is justified by an extremely high number of fitting parameters and the small fractional amplitude of the secondary lifetime. In the following work, the secondary lifetime is considered as donor-only decay without any energy transfer to the acceptor and without influence on the FRET efficiency for the global analysis of fluorescence decay curves of polyglutamine chains (equation 4.6). The time-resolved fluorescence intensity decay $F(t)$ can be described as an addition of the time-resolved fluorescence intensity decay $F_{slow}(t)$ as discussed in chapter 2 and the additional time-resolved fluorescence intensity decay $F_{fast}(t)$ resulting from the fast *pyrene*-donor-only decay τ_{fast}, which was measured in separate lifetime experiments of *pyrene*-donor-only polypeptide chains:

$$F(t) = F_{slow}(t) + F_{fast}(t) \cdot \exp\left(-\frac{t}{\tau_{fast}}\right) \qquad \text{(Eq. 4.6)}$$

Of course, a careful evaluation of the global analysis results is essential, which will be done with a fitting error of the global analysis χ^2 is less than 1.3 and the *SPA*.

4.2.2.3 trFRET measurements of Q_{14} in aqueous solution

Fluorescence decay curves of Q_{14} of both FRET-pairs in 10 mM potassium phosphate buffer at pH 7.0 and 22.5 °C were measured. All samples were prepared with the modified disassembly protocol from Wetzel and colleagues. [183], [246]. In Figure 4.2.8 the fluorescence decay curves of the Q_{14} labeled with FRET chromophores as well as the corresponding donor-only labeled polypeptide chains are shown. The

fluorescence decay curves were globally analyzed using equations 2.1 and 1.12 including the additional fluorescence decay of the secondary lifetime of the *pyrene*-chromophore, which was measured in separate lifetime experiment (equation 4.5). The resulting residuals of the global analysis exhibited a well-defined global analysis on both FRET-pairs-labeled Q_{14}. The resulting global χ^2 for both fluorescence decay curves was 0.99. However, for the dansyl-Q_{14}-pyrene fluorescence decay curves a significant higher deviation in the global analysis at early time points was observed (see residuals, A in Figure 4.2.8). This deviation at earlier times of the decay curves is due to an inaccurately defined found starting point of the fit in the global analysis. However, resulting fitting parameters were determined beginning from the starting point of the decay curve at the maximum fluorescence intensity of 10.000 counts per channel (cnts./ch.). Thus, the deviations at earlier times observed from the residuals of the global analysis do not seem to be relevant for globally analyzed results.

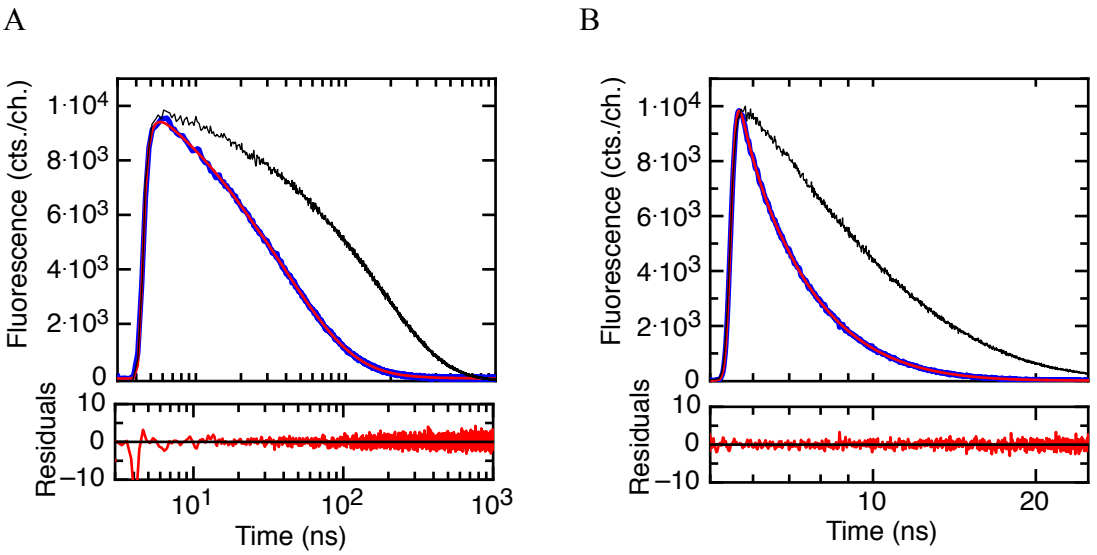


Figure 4.2.8: Fluorescence decay curves of Q_{14}. Fluorescence decay curves (blue lines) for both **A** dansyl-Q_{14}-pyrene and **B** DNB-Q_{14}-AMCA were fitted globally using equations 1.12 and 2.1 considering the secondary donor-only decay of *pyrene* as a single-exponential decay without any contribution in FRET (red lines). The corresponding fit residuals are shown at the bottom. Donor-only decay curves are shown in black lines for **A** Ac-Q_{14}-pyrene and **B** Ac-Q_{14}-AMCA. The large residuals at very short times in A are due to inaccuracies in the definition of the starting time.

The resulting fitting parameters of the global analysis were b = 17.5 Å and σ = 12.6 Å. With the EDWARDS model the population of end-to-end distances were calculated and shown in Figure 4.2.10 (equation 1.12). The root mean square donor-

acceptor distances of Q_{14} in water was 25.9 Å. An intrachain diffusion coefficient of 9.7 Å^2/ns was found.

However, a global analysis without the secondary lifetime for *pyrene* as a simple decay revealed a χ^2-value of 1.14 and a significant higher deviation of the resulting residuals was observed for the dansyl-Q_{14}-pyrene fluorescence decay curve (Figure 4.2.9).

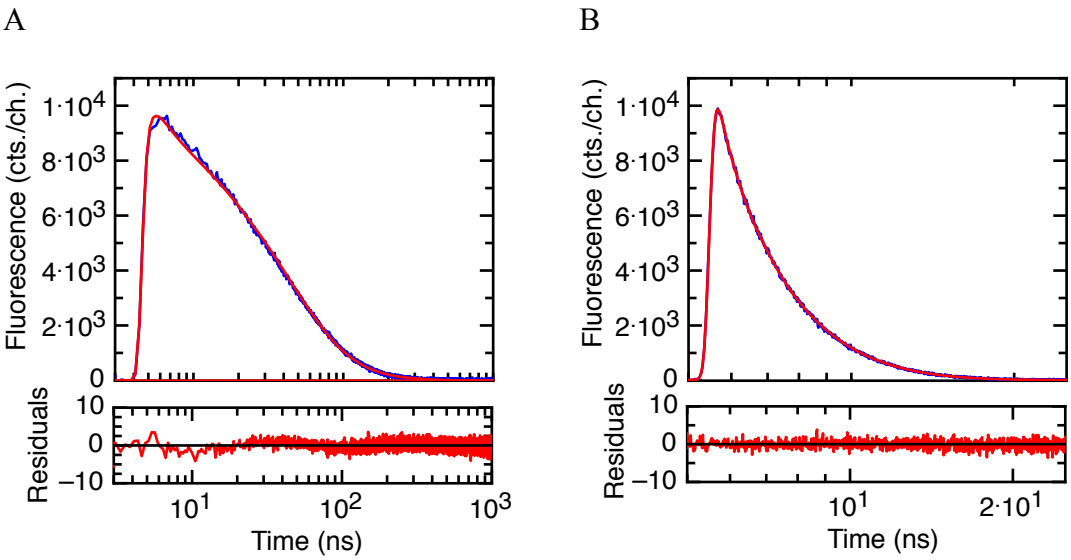


Figure 4.2.9: Fluorescence decay curves of Q_{14}. FRET decay curves (blue lines) for both **A** dansyl-Q_{14}-pyrene and **B** DNB-Q_{14}-AMCA were fitted globally using equation 1.12 and 2.1 considering the secondary donor-only decay of *pyrene* as a single-exponential decay without any contribution in FRET (red lines). The corresponding fit residuals are shown at the bottom.

The fitting parameters without the secondary lifetime were slightly different compared to the results considering the secondary lifetime of *pyrene* as a single-exponential decay without any contribution in FRET. For *b* a value of 16.3 Å, for σ = 13.3 Å and for the root mean square donor-acceptor distance a value of 25.9 Å were found. The intrachain diffusion coefficient was 10.3 Å^2/ns. The population of end-to-end distances was calculated using the EDWARDS model (equation 1.12) and the fitted parameters *b* and σ. No significant deviation of the population of end-to-end distances was found between the global analysis procedure that considers or neglects the single-exponential decay of the secondary lifetime (Figure 4.2.10).

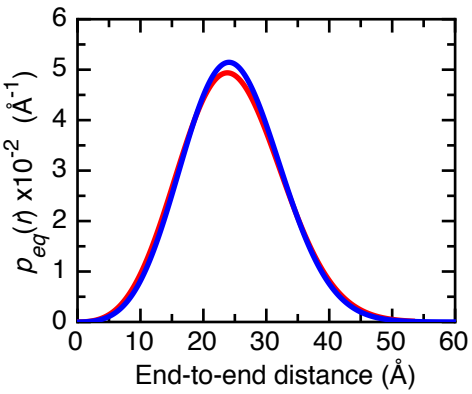


Figure 4.2.10: Population of end-to-end distances of Q_{14} in water. The population of end-to-end distances was calculated using the EDWARDS model (equation 1.12) and the parameters *b* and *σ* of the global analysis procedure that considers (blue line, equation 4.6) or neglects (red line, equations 2.1 and 1.12) the single-exponential decay of the secondary lifetime.

4.2.2.4 Analysis of accuracy of Q_{14} results with the *SPA*

In order to evaluate the fitting quality of the global analysis a so-called *Support Plane Analysis* (*SPA*) regarding all three fitting parameters *b*, *σ* and the intrachain diffusion coefficient was established. The calculated results of the global analysis were used for the *SPA*. Here, one of the parameters is set to a fixed value, whereas the others are freely fitted. A stepwise variation of the fixed parameter at lower or higher values should result in a higher χ^2 and the characteristic parabolic shape of a *SPA* (Materials and Methods 6.6.3). With the *SPA* the accuracy of all three fitting parameter *b*, *σ* and the intrachain diffusion coefficient was determined (Figure 4.2.11). Therefore, an F-statistic value of 68.3% was set, which is comparable to the well-known standard deviation of a *Gaussian* deviation (Materials and Methods 6.6.3) (Figure 4.2.11 dashed red lines). The following values with corresponding errors were determined of b = 17.5 ± 1.1 Å and σ = 12.6 ± 0.6 Å for Q_{14} in water (A and B in Figure 4.2.11). The intrachain diffusion coefficient yielded a value of 9.7 ± 2.1 Å^2/ns in water (C in Figure 4.2.11).

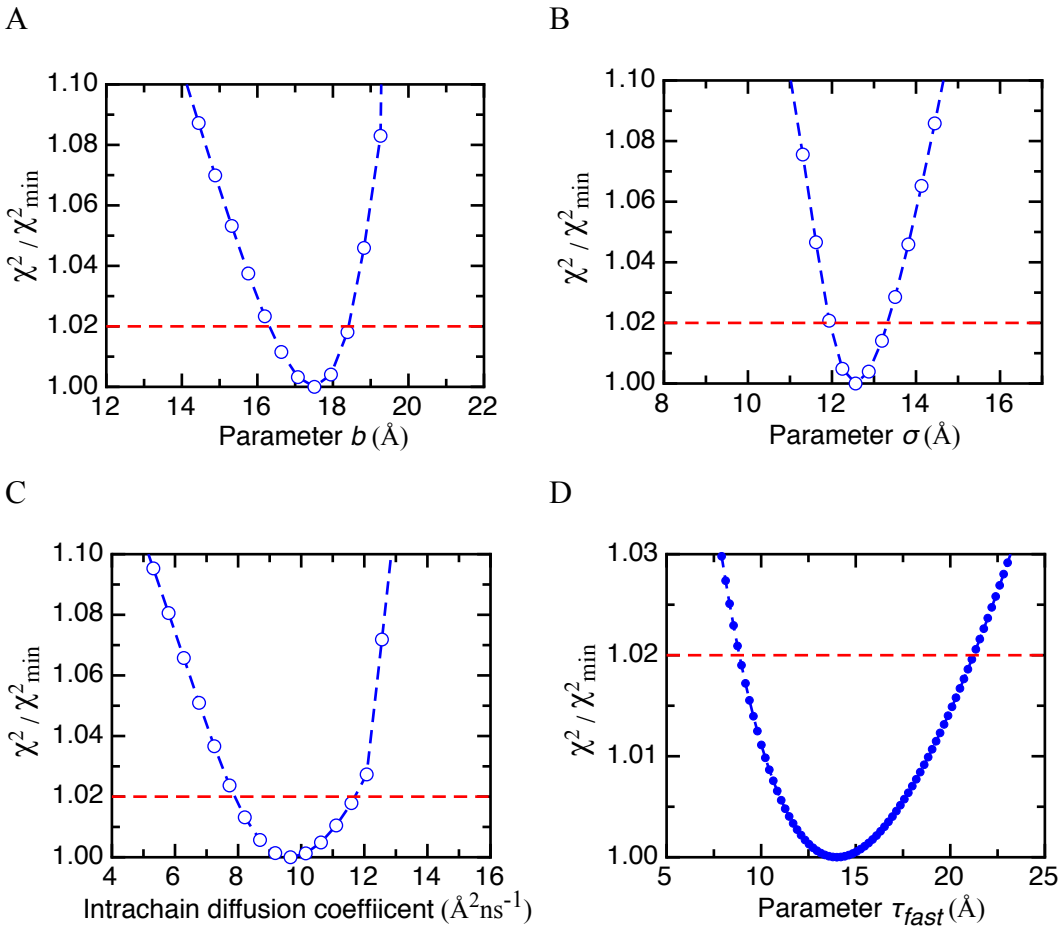


Figure 4.2.11: *Support Plane Analysis* of the fitted parameters *b*, *σ* and intrachain diffusion coefficient resulting from global analysis of Q_{14} in water. The best freely analyzed parameters altered towards lower and higher values for **A** *b*, **B** *σ*, **C** intrachain diffusion coefficient and **D** secondary lifetime of *pyrene*. An F-statistic value of 68.3% was used for calculation of the fitting error of the *Support Plan analysis* (dashed red line).

Additionally, the secondary lifetime of *pyrene* was investigated with the *SPA*. It should be taken into account that the fitting results of the free global analysis did not change when the secondary donor-only lifetime of *pyrene* was used as a fixed or as a flexible fitting parameter in the global analysis. The secondary donor-only lifetime of *pyrene* τ_{fast} showed a higher inaccuracy of 14 ± 60 ns, which can be explained by the small proportion of the amplitude of the secondary donor-only lifetime of about 16% compared to the entire amplitude in the time-resolved FRET experiment (D in Figure 4.2.11). It should to be noted that in experiments with denaturants and at different temperatures the amplitude of the secondary lifetime increased to 40%, where the error was mostly found at about 14 ± 2.1 ns.

The calculation of the error of the root mean square end-to-end distances became more difficult, because the normal routine is working with the three fitting parameters *b*, σ and the intrachain diffusion coefficient and the root mean square end-to-end distance has to be implemented instead of one of the three fitting parameters. For a direct implementation of the root mean square end-to-end distance in the global analysis equation, a mathematical approximation was used (Materials and Methods 6.6.3, equation 6.7). Therefore, the fitting parameter *b* was replaced by the root mean square end-to-end distance. This approximation can be done, if $0.5 < b/\sigma < 1.5$ is given. With this assumption a direct *SPA* with the root mean square end-to-end distance became possible. In Figure 4.2.12 the resulting *SPA* of the root mean square end-to-end distance with a small error of 25.9 ± 0.3 Å is presented.

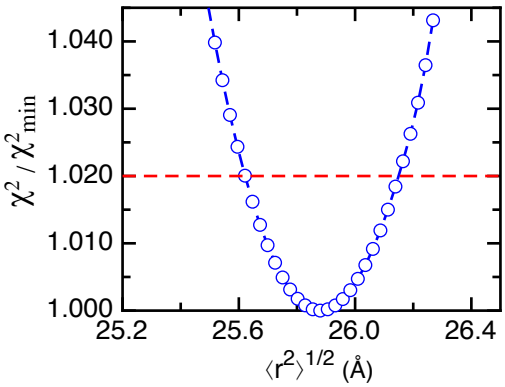


Figure 4.2.12: *Support Plane Analysis* of the root mean square end-to-end distance of Q_{14} in water. For calculation of the error of the root mean square end-to-end distance a mathematical approximation was done, where *b* replaced by the root mean square end-to-end distance (Materials and Methods). An F-statistic value of 68.3% was used for calculation of the fitting error of the *Support Plan analysis* (dashed red line).

Both the intrachain diffusion coefficient as well as the root mean square end-to-end distance represent suitable parameters to characterize the chain dynamics and dimensions of unstructured polypeptide chains, whereas the errors calculated from the *SPA* reveal a reasonable certainty for a quantitative interpretation of data.

4.2.3 Determination of the R_H and R_G of Q_{14} in aqueous solution

As discussed in the introduction, the hydrodynamic radius R_H and the radius of gyration R_G represent common used parameters to describe the size of polypeptides. For small and short polypeptide chains, such as the Q_{14}, the peptide concentration often limits a high-quality analysis. For the determination of the R_G, a more powerful synchrotron small-angle X-ray scattering measurements for DNB-Q_{14}-AMCA were performed (SAXS at EMBL Hamburg). The resulting data were fitted under the assumption of the Debye formula and yielded an R_G = 0.81 ± 0.1 nm (The measurements were performed by Philipp Skorupa and PD Dr. König (Martin-Luther University Halle-Wittenberg, Institute of Biochemistry and Biotechnology). The data was analyzed by Prof. Dr. Saalwächter (Martin-Luther University Halle-Wittenberg, Institute of Physics), for detailed information see Appendix, Figure A.2). It should be noted that the data and analysis from the SAXS experiment are not completely reliable and require further measurements and evaluations. With the dynamic light scattering (DLS) experiments the R_H can be determined. It should be considered that the estimation of the hydrodynamic radius R_H is correct using the self-diffusion coefficient only in the limit range of low concentrations, as is the case in this experiment, because in this limit the hydrodynamic coil-coil interactions do not influence the measured diffusion coefficient. Only for dansyl-Q_{14}-pyrene a sufficiently high concentration was achieved and an apparent R_H of 1.1 ± 0.5 nm was yielded. Both radii show a similar size of the Q_{14}. All peptides were measured in pure water at pH 7 to achieve a higher concentrated solution. Unfortunately, the extremely high concentrated samples tended to aggregate in pure water and impaired the results, thus the resulting R_G and R_H should be considered as estimations only. Despite the upper limit of the peptide concentration, the contribution of amplitude is extremely low. Thus, a higher uncertainty of the data was found. However, a direct relation to the root mean square end-to-end distances, determined using the time-resolved FRET method, became challenging and requires other polymer model and assumptions as discussed in the introduction. Additionally, no analytic expression exists to calculate the R_H from the distance distribution of the EDWARDS model [159].

4.2.4 Summary and Discussion

It was shown that the modified disassembly protocol from Wetzel and colleagues is a reliable tool to generate homogenous solutions of monomeric Q_{14} in aqueous solvent. Different methods such as CD, light scattering and analytical ultracentrifugation confirmed a stable monomeric solution of Q_{14} over hours. First measurements of the donor-only Q_{14} showed an unexpected secondary donor-only lifetime of 14 ± 2.1 ns for the donor *pyrene*, which was also detected in the FRET experiments.

This additional secondary lifetime was included in the global analysis as donor-lifetime decay without energy transfer to the acceptor. This approach was justified by analysis of the fitting results using the *SPA*. The global analysis of Q_{14} in water (10 mM potassium phosphate buffer at 22.5 °C) revealed a root mean square distance of 25.9 ± 0.3 Å. This found root mean square end-to-end distance was in good agreement with the results of FRET experiments by Finke *et al.*, who determined a value of about 30 Å for the two glutamines longer Q_{16} [260]. Finke *et al.* used two different fluorescence methods based on the simple relation of Förster distance and FRET efficiency (equation 2.6). The FRET efficiency can be determined from fluorescence lifetime or fluorescence intensity by comparison of the donor-acceptor signal and the donor-only signal (equation 2.5). This time-resolved FRET and the steady-state FRET experiments revealed similar root mean square end-to-end distances of 30.3 ± 1.1 Å and 30.4 ± 2.4 Å, respectively. Nevertheless, in this simple calculation, the influence of the intrachain diffusion coefficient was ignored as well as the end-to-end distribution of unfolded polypeptide chains. Walters *et al.* used the same simple relationship to calculate the average end-to-end distance *R,* determined with steady-state FRET measurements. They neglected the intrachain diffusion in the way as indicated by Lakowicz *et. al.* [158]. In this assumption the influence of the intramolecular diffusion can be ignored, under condition of $(D \cdot \tau / s^2) \ll 1$, where τ is the donor lifetime, *D* the expected intrachain diffusion and *s* the estimated end-to-end distance. Additionally, they used the *freely jointed chain* model with a Gaussian distribution of these distances to calculate the average end-to-end distance R_{FJ}. For the investigated Q_{12} an R = 21.2 ± 0.1 Å and R_{FJ} = 15.2 Å was found. The longer Q_{16} revealed an R = 22.5 ± 0.6 Å and R_{FJ} = 16.6 Å. Both average end-to-end distances were about 15% smaller compared to the Q_{14} root mean square distance of 25.9 ± 0.3 Å. It should be noted that in this publication the polyglutamine chains are flanked

with two charged lysine residues on both chains end. The resulting repulsive effects of the lysine at pH 7.4 were critically discussed in the same publication. The repulsive effect has a significant influence on the average end-to-end distance, which increases with increasing length of the polypeptide chain. Additionally, for the short lifetime of the donor tryptophan of about 3 ns the intrachain diffusion coefficient was assumed to be negligible. Simulated fluorescence decay curves of the DNB-Q_{14}-AMCA peptide chain with an AMCA lifetime of 5 ns depicted a slightly, but nevertheless significant influence of the intrachain diffusion on the fluorescence decay curve (chapter 2.2, A in Figure 2.1). The estimation of the end-to end distance distribution using the *freely jointed chain* model, which is only valid for ideal chains, additionally complicated the comparison of the obtained results and might explain the 15% difference in the end-to end distances. Interestingly, if the same assumption was used for the time-resolved fluorescence decay curves of the shorter lifetime of AMCA, a value of 26.5 Å for the end-to-end distance was found. But, in the time-resolved fluorescence decay curve of the DNB-Q_{14}-AMCA chain two lifetimes were found, regarding the reconvolution fit. The main fractional amplitude of about 54% was used for the calculation of the 26.5 Å end-to-end distance. The other lifetime with a fractional amplitude of 46% yielded an average end-to-end distance of 22.0 Å. These results reveal that the simple assumption according to equations 2.5 and 2.6, discussed in the literature, can follow in expected results but should be carefully applied. However, the average or root mean square end-to-end distance will not reveal, whether the polyglutamine chains are classified as more collapsed or extended chain. The FLORY theory, based on the physics of polymer solutions, is frequently used to classify the conformational state of polymers and links the conformational state with the properties of the solvent (chapter 1.4). In the FLORY assumption, the chain-chain interactions are compared to chain-solvent interactions. A more preferential chain-chain interaction depicts a more collapsed state, whereas a stronger chain-solvent interaction results in an expansion of the polymer. In order to describe the "goodness" of a solvent, a simple equation $R \approx N^{\nu}$ links the polymer sizes to the excluded volume exponent ν, where N is the number of chain segments (chapter 1.4, equation 1.9). For a more expanded polymer a $\nu = 0.588$ is predicted in good solvent. A so-called "theta-solvent" describes the balance between chain-chain and chain-solvent interactions with $\nu = ½$, where the polymer behaves like an ideal linear chain. A ν of 0.33 reflects a poor solvent, where the polymers collapse. By fluorescence correlation spectroscopy a ν of 0.32 ± 0.02 was

determined for glutamine repeats from 5 to 55 in water and it was concluded that water is a more poor solvent for polyglutamine chains in general [180].

A more detailed investigation of polyglutamine with different length were published by Walter *et al.*. The length dependence steady-state FRET experiments indicated an extended polypeptide chain for less than 12 glutamine repeats. The polypeptide chains containing 20 glutamine repeats or longer could be described as collapsed chains in poor solvent, while for Q_{16} the water is a theta solvent [244].

In summary, for Q_{14} water seems to be classified as a solvent between poor and theta solvent, or in other words Q_{14} are in the range of a collapsed and a theta state. However, the description of the dimension of flexible polypeptide chains and the interpretation may be difficult, just from the fact that the ensemble of conformations causes in a broad end-to-end distance distribution. Besides the dimension another important parameter is the dynamics of an unstructured polypeptide chain. In this work, the calculated intrachain diffusion coefficient described the dynamics of unstructured polypeptide chains in high accuracy. The determined intrachain diffusion coefficient of 9.7 ± 2.1 $Å^2$/ns was two times lower compared to the natural polypeptide chains EF-loop of the IDP β-carp parvalbumin, with 22.2 $Å^2$/ns [159]. In comparison to a highly flexible $(GS)_8$ a more than three times lower intrachain diffusion coefficient of 33 $Å^2$/ns compared to the Q_{14} was found [159]. In tryptophan/cysteine triplet quenching experiments of different length of polyglutamine chains the rate of contact formation was determined [115]. The intrachain diffusion coefficient was calculated from these rates with the help of the *Szabo-Schulten-Schulten* theory [261]. It was found that the intrachain diffusion coefficient slowed down with increasing polyglutamine chain length. Unfortunately no polyglutamine with 14 repeats was measured in the work mentioned above. The Q_{13} was the compound closely related to the Q_{14} with an intrachain diffusion coefficient of 6.6 x 10^{-7} cm^2/s or 6.6 $Å^2$/ns, which is 32% lower compared to that found value using time-resolved FRET measurements. The longer polyglutamine of 16 repeats revealed a still lower diffusion coefficient of 1.5 x 10^{-7} cm^2/s or 1.5 $Å^2$/ns, so that the theoretical intrachain diffusion coefficient of Q_{14} was expected between 1.5 and 6.6 $Å^2$/ns for the tryptophan/cysteine quenching experiment.

In comparison to the 9.7 ± 2.1 $Å^2$/ns found by the here used time-resolved FRET method, an approximately two times slower intrachain diffusion coefficient in comparison to the tryptophan/cysteine triplet quenching experiment was determined.

Nevertheless, this tryptophan triplet quenching method is known to be not fully diffusion-limited, which restricts the interpretation of chain dynamics on an absolute time-scale [262]. Yeh and Hummer *et al.* found significantly faster rate constants for loop formation in penta-peptide simulations, leading to faster diffusion coefficient compared to the rate constant determined from tryptophan triplet quenching experiments [116]. Additionally, the Szabo-Schulten-Schulten theory (SSS-theory) requires two important assumptions. First, the theory assumes an idealized chain with a *Gaussian* distribution of the conformations, which was shown in this work to fail to describe exactly the realistic polypeptide chain distribution (chapter 2.2, Figure 2.2). Second, the SSS-theory presupposed a limit contact distance, which described the closed contact of the donor-acceptor for the energy transfer. The limit contact distance is not exactly known and reveals a high impact in the calculation of the diffusion coefficient.

Furthermore, in the tryptophan/cysteine triplet quenching experiment sucrose as an additive was used. Due to the sucrose as a viscogenic co-solute the viscosity is altered, which decreases the apparent intrachain diffusion coefficient (chapter 4.1) as well as the contact rate constant of loop formation [218]. Also the included two lysines residues on both ends of the model polypeptide chain for the tryptophan/cysteine triplet quenching method is known to show an extreme pH dependence and alters the dynamics of polypeptide chains [244]. Generally, the found intrachain diffusion coefficient was in the same range of the published intrachain diffusion coefficient for comparable polyglutamine chains. It seems to be difficult to compare in detail the intrachain diffusion coefficient, *e.g.* to the tryptophan/cysteine quenching. However, for Q_{14} an extremely slow chain dynamics was found in comparison to artificial and natural model polypeptide chains. The intrachain diffusion coefficient is known to be influenced by a couple of different factors, like ionic or steric effects as well as H-bonds, which slows down the dynamics of chains. Detailed investigation of the chain dynamics and dimensions of Q_{14} may illustrate the slow chain dynamics of Q_{14}. One of them is the activation energy, which can be determined by measuring the temperature dependence of chain parameters. The comparison to non-amyloidic model polypeptide chains can give further evidence for the disease-related behavior of the polyglutamine chains. Furthermore, the impact of a denaturant on the chain dynamics and dimensions can be investigated. As discussed in chapter 4.1, the influence of the solvent viscosity on chain dynamics and dimensions

has to be known to correct its influence on the intrachain diffusion coefficient, also in dependence on temperature and denaturants. At first, Q_{14} were measured in dependence of different glycerol concentrations.

4.3 The effect of the viscogenic co-solute glycerol on chain dynamics and dimensions of Q_{14}

Further investigation of dynamics and dimensions of Q_{14} require detailed information about their viscosity dependence. As shown and discussed in chapter 4.1, glycerol is suitable as a small viscogenic co-solutes in this concern. Effects like inhomogeneity, as shown for larger viscogenic co-solutes, can be neglected by using small viscogenic co-solutes such as glycerol.

4.3.1 trFRET measurements of the glycerol dependence on Q_{14}

Time-resolved FRET measurements on Q_{14} were performed in 10 mM potassium phosphate buffer pH 7.0 at 22.5 °C containing glycerol ranging from 8% to 56% (w/w). As shown in Figure 4.3.1, a well-defined global analysis was found for the dansyl-*pyrene* fluorescence decay curves at all glycerol concentrations. Slight deviations from an optimal residual distribution were found for the DNB-AMCA fluorescence decay curves, but the global χ^2 under 1.2 generally showed a good accuracy.

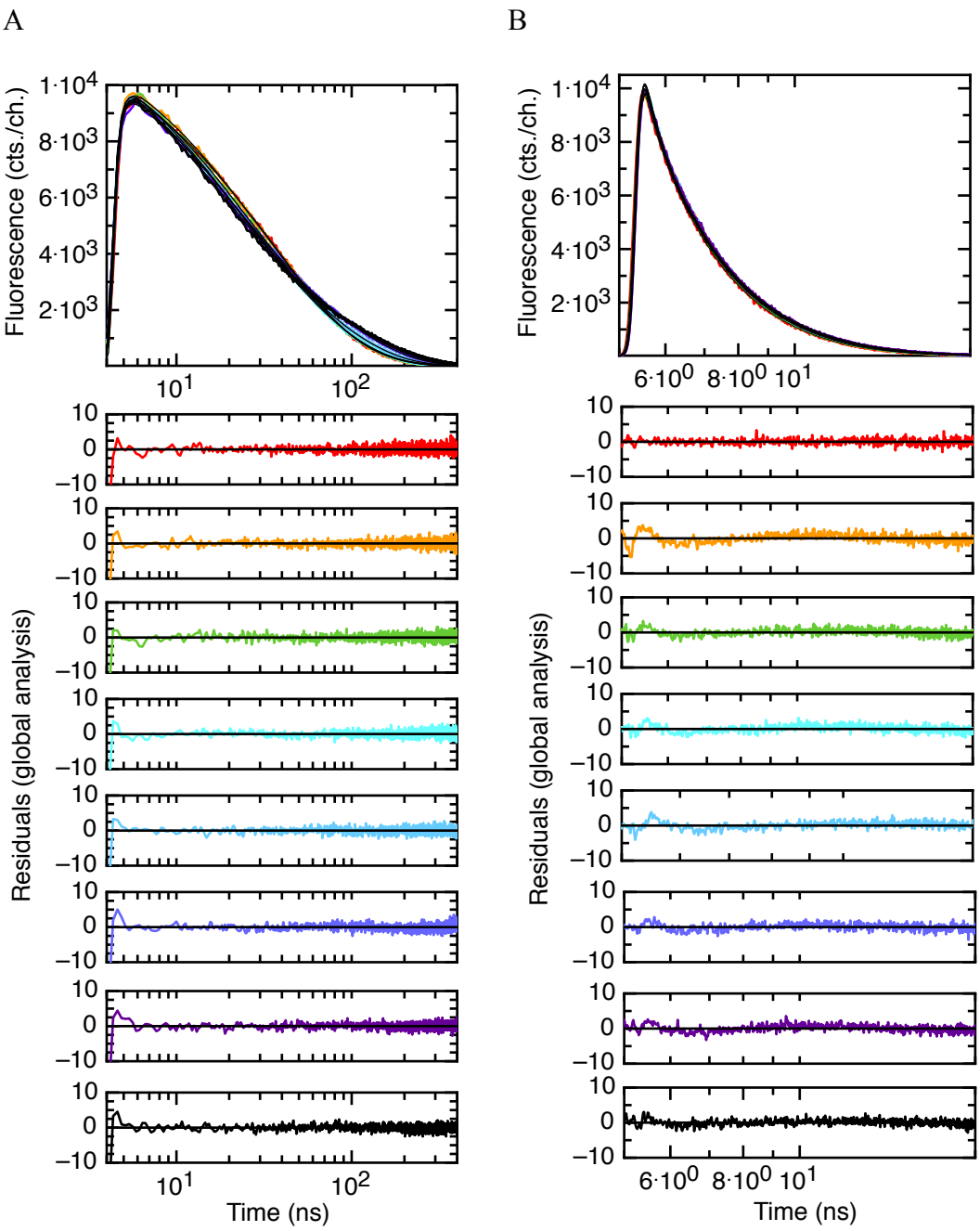


Figure 4.3.1: Time-resolved fluorescence decay curves of Q_{14} in dependence of glycerol using donor-acceptor pair **A** dansyl-*pyrene* and **B** DNB-AMCA in 10 mM potassium phosphate pH 7.0 at different concentration ranges from 0% (red), 8% (orange), 12% (green), 16% (turquoise), 28% (blue), 40% (purple), 48% (dark purple) to 56% (black) (w/w). The residuals of the global analysis of each concentration point are shown in different colors corresponding to the different concentrations of glycerol for both FRET-pairs.

The fitted parameters b and σ are shown in Figure 4.3.2 **A** and depicted a slight change with increasing viscosity at higher concentrations of glycerol. The root mean square end-to-end distances calculated using equation 1.12 showed a slight increase of about 3% ranging from water to the highest glycerol concentration of 56% (w/w) (B in Figure 4.3.2). The population of end-to-end distances was slightly decreased at increasing glycerol concentrations, which corresponded to the fitting parameter σ (C in Figure 4.3.2).

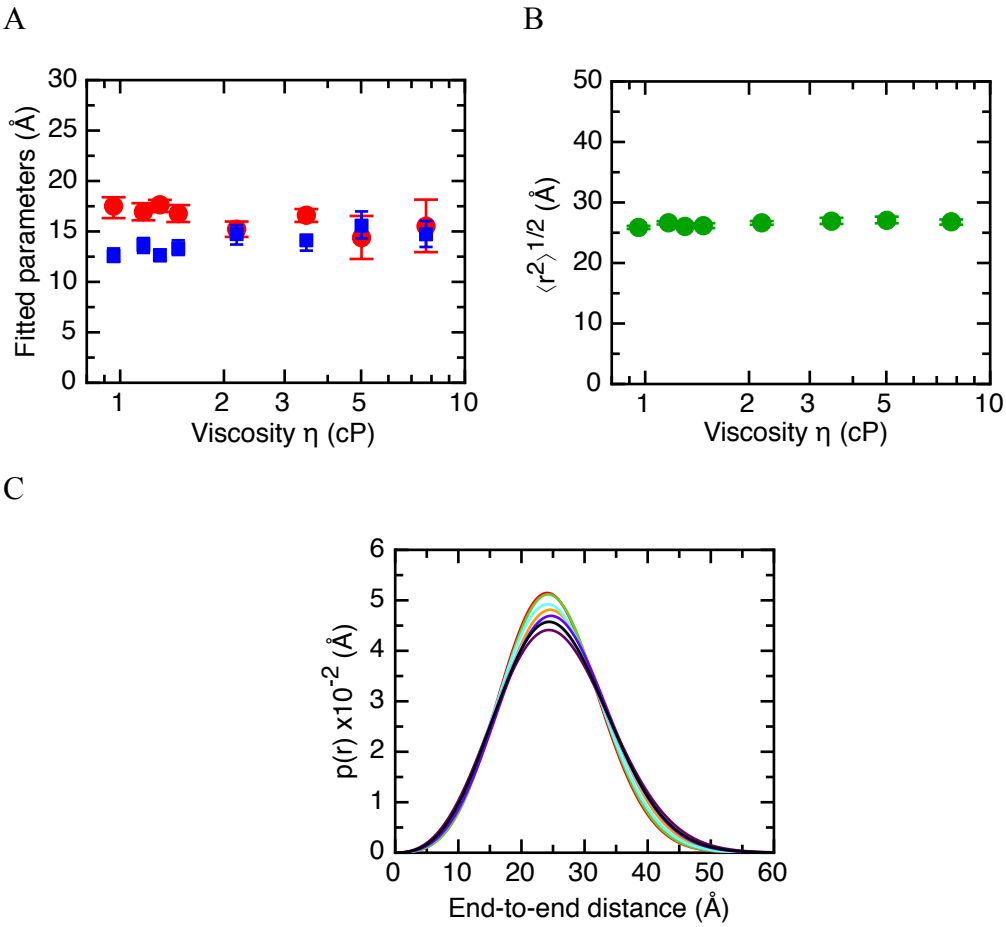


Figure 4.3.2: Glycerol dependence of dynamics and dimensions of Q_{14}. **A** Semi-logarithmic illustration of the fitting parameters b (red circles) and σ (blue squares); **B** the root-mean square end-to-end distance was calculated with equation 1.12 **C** linear plot of the end-to-end distribution $p_{eq}(r)$ at different glycerol concentrations ranges from 0% (red), 8% (orange), 12% (green), 16% (turquoise), 28% (blue), 40% (purple), 48% (dark purple) to 56% (black) (w/w).

In Figure 4.3.3 the dependence of the intrachain diffusion coefficient on the viscosity of glycerol is shown. A fit using the empirically determined equation 4.4 (chapter 4.1) yields the viscosity-sensitive exponent α.

$$D = D_0 \cdot \left(\frac{\eta}{\eta_0} \right)^{-\alpha} \qquad \text{(Eq. 4.4)}$$

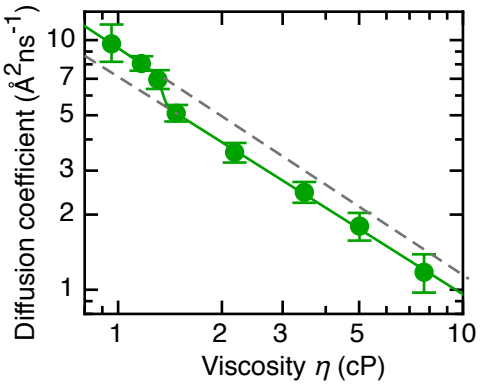


Figure 4.3.3: Glycerol dependence of the intrachain diffusion coefficient of Q_{14}. The double-logarithmic illustration of the intrachain diffusion coefficient against the viscosity of glycerol was analyzed with equation 4.4 to yield the viscosity-sensitive exponent α. Two separate parts were observed and fitted using a combined equation 4.4 (Martials and Methods 6.13). The dashed gray lines describe the trend resulting from a α-value of 0.90 (upper line) and from α-value of 0.85 (lower line) fitted with equation 4.4.

As shown in Figure 4.3.3 two separated ranges were observed. The first part starts from the lowest viscosity 0.96 cP (0% (w/w) glycerol) and ends at 1.30 cP (12% (w/w) glycerol). The second main part was found between 1.47 cP (16% (w/w) glycerol) and 7.71 cP (56% (w/w) glycerol). Both parts were fitted with a combined equation 4.4 (Materials and Methods 6.13) and yields an α-value of 0.90 ± 0.75 for the first part and a α-value of 0.85 ± 0.13 for the main part. The extremely high error found for the first part resulted from the low number of data points in combination with the high error of the first data point (lowest viscosity). Considering the accuracy a comparable α-value for both parts was found.

4.3.2 Summary and Discussion

For the glycerol dependence of the intrachain diffusion coefficient of the Q_{14} two separate parts with intrinsic α-values determined using the equation 4.4. The first one displays less precision, which results from the low number of data points. In contrast, the second part, ranges from 16% to 56% (w/w) of glycerol, reveals a high precision. In accordance to these results, the α-value of the second part is used for the viscosity correction in further experiments such as temperature and denaturant dependences. TTET as well as previous trFRET experiments yielded single dependences of artificial amino acid sequences like glycine-serine or natural sequences like the fragment EF- and DE-loop from the IDP β-carp parvalbumin [159], [218]. In these experiments, the data could be described with a single α-value (chapter 4.1.3). Generally, a physical change of glycerol with increasing concentration was not described in the literature. The anomalous dependence of intrachain diffusion coefficient of the Q_{14} on the glycerol viscosity could be hypothesized by the formation of the transient sub-structures originated from H-bonds, which are disturbed by adding glycerol. However, the root mean square end-to-end distance shows no significant change at increasing glycerol concentrations and is comparable with previously measured EF-loop and DE-loop at 4% (w/w) glycerol ([159] and chapter 4.1). The interpretation of the viscosity dependence mentioned above is considered as a first hypothesis. Nevertheless, the found α-value of 0.85 ± 0.13 is significantly less than expected for the Kramers theory of α = 1 (chapter 4.1). This may be a further hint for transient sub-structures that do not "feel" the full solvent viscosity. However, the determined α-value of 0.85 ± 0.13 is useful for the viscosity correction of temperature or denaturant dependences of the polypeptide chains (equation 4.7):

$$D_{corr} = D \cdot \left(\frac{\eta}{\eta_0} \right)^{\alpha} \qquad \text{(Eq. 4.7)}$$

Where the viscosity corrected intrachain diffusion coefficient D_{corr} is calculated using the determined α -value and the experimental intrachain diffusion coefficient D.
The η is the viscosity of the solvent and the η_0 is the viscosity of the reference solution (10 mM potassium phosphate buffer at 22.5 °C and pH 7.0).

4.4 The effect of temperature on the dynamics and dimensions of Q_{14}

The significantly low intrachain diffusion coefficient for the Q_{14} in water, as previously determined, may be one reason for the disease-related behavior of monomeric polyglutamine chains. The slow dynamics found for Q_{14} may caused by higher conformational energy barriers between unstructured conformers. In order to test this, a temperature-dependent time-resolved FRET experiments were performed. In literature different theoretical models are discussed to determine the conformational energy barriers of unfolded polypeptide chains. In a first approach, the *Arrhenius* equation and the so-called *Zwanzig* model can be applied for the calculation of energetic barriers of the conformational changes of unfolded polypeptide chains in dependence on temperature [105], [123], [159], [218], [263]–[274].

4.4.1 Arrhenius formula applied for unfolded polypeptide chains

The well-known *Arrhenius* equation connects the temperature-dependent reaction rate constant with the activation energy. Thereby, the activation energy generally describes the energy barrier that separates the folded and the unfolded state (transition state theory) (equation 4.8).

$$k = A \cdot e^{\frac{-E_A}{RT}} \qquad \text{(Eq. 4.8)}$$

Where k is the temperature-dependent rate constant of a reaction, A is the pre-exponential factor, R is the gas constant and E_A represents the activation energy. When the *Arrhenius* equation is compared to the *Eyring* equation (equation 4.9), it becomes clear that the pre-exponential factor A includes the activation entropy ΔS^* of the transition state with a maximum rate constant (or frequency factor) k^0 (equation 4.10):

$$k = \frac{k_B \cdot T}{h} \cdot e^{-\frac{\Delta H^*}{RT}} \cdot e^{\frac{\Delta S^*}{R}} \qquad \text{(Eq. 4.9)}$$

$$A = k^0 \cdot e^{\frac{\Delta S^*}{R}} \text{ and } k^0 = \frac{k_B \cdot T}{h} \qquad \text{(Eq. 4.10)}$$

Where ΔH^* is the activation enthalpy of the transition state, h is the Planck constant and k_B is the Boltzmann constant. The activation enthalpy ΔH^* is equal to the activation energy E_A ($E_A = \Delta H^* + RT$) with the approximation that the term RT is negligible at room temperature. Unlike for a folded protein, the *Arrhenius* activation energy for unfolded polypeptide chains can be understood as an energy barrier between different conformations within the whole conformational space. The *Arrhenius* equation for the rate constant of a reaction can be adopted for the intrachain diffusion coefficient D_{corr} of unfolded polypeptide chains (equation 4.11) [159], [269], [275]. Notice, the intrachain diffusion coefficient has to be viscosity-corrected to neglect the influence of the changing viscosity at different temperatures (chapter 4.3).

$$D_{corr} = A \cdot e^{\frac{-E_A}{RT}} \qquad \text{(Eq. 4.11)}$$

4.4.2 *Zwanzig* model of a rough energy landscape

The *Zwanzig* model describes the interconversion of conformational states of an unfolded polypeptide chain with a free energy surface as a rough potential, where the energetically sub-states or microstates are separated by energy barriers [276], [277]. Therefore, the sub-states can only exist when the energy barriers separating the sub-states are as large as the thermal energy k_BT. The dynamics of the polypeptide chains can be represented by the diffusion.

The *Zwanzig* model deals with the diffusion in a rough one-dimensional potential. The interconversion between different conformers in a rough energy landscape crosses minor energy barriers, which can be easily overcome. The limitations for this model are 1.) the assumption of a one-dimensional diffusion to cross the minor energy barriers, despite a multi-dimensional diffusion occurs for unfolded polypeptide chains and 2.) the assumption of a square height of energy barriers, which mathematically sum up the single barriers without discrimination between different energy barriers.

Nevertheless, in previous work the *Zwanzig* model was successfully used to describe the dynamics of unfolded polypeptide chains [263]–[268], [273]:

$$D = D^{*} \exp\left(\frac{\varepsilon^{2}}{\left(k_{B}T\right)^{2}} \right) \qquad \text{(Eq. 4.12)}$$

In this equation, D is the intrachain diffusion coefficient, which has to be viscosity-corrected as discussed in chapter 4.3 (equation 4.7). D^{*} reflects the intrachain diffusion coefficient without minor energy barriers, k_B is the Boltzmann constant and T is the temperature in Kelvin. ε represents the root mean square height of the minor energy barrier.

4.4.3 trFRET measurements of Q_{14} at different temperatures

The time-resolved FRET experiments were carried out in 10 mM potassium phosphate buffer, pH 7.0. The temperature ranges from 2.5 °C (275.65 K) to 42.5 °C (315.65 K). The data were globally analyzed using equation 1.12 and 2.1 (Figure 4.4.1). Figure 4.4.1 shows the individual global analysis of the fluorescence decay curves, which exhibit a good quality of the fit, shown by the residuals at the bottom of each particular temperature.

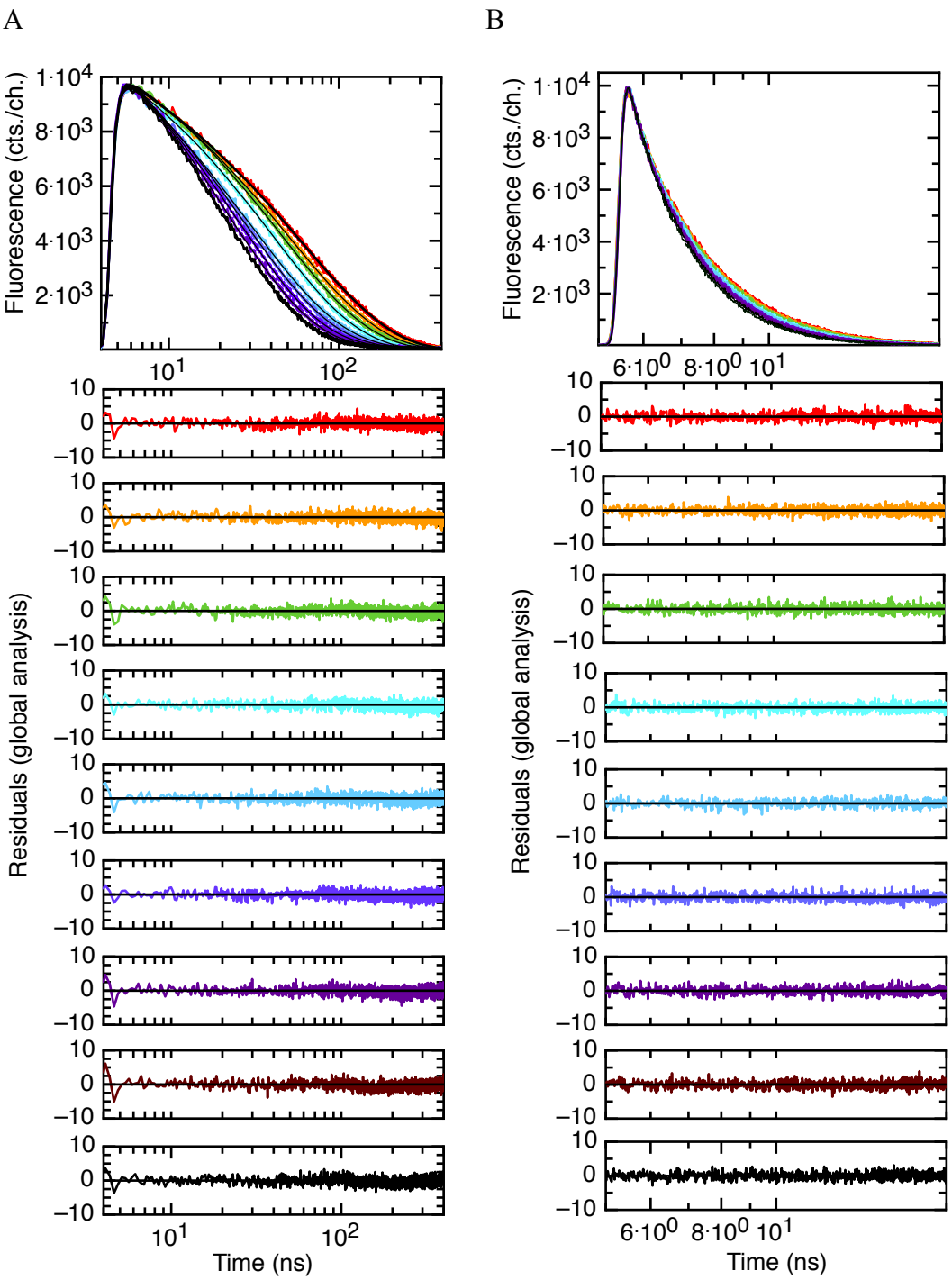


Figure 4.4.1: Time-resolved fluorescence decay curves of Q_{14} in dependence of temperature using donor-acceptor pair **A** dansyl-*pyrene* and **B** DNB-AMCA in 10 mM potassium phosphate pH 7.0 at different temperature ranges from 2.5 °C (red), 7.5 °C (orange), 12.5 °C (green), 17.5 °C (turquoise), 22.5 °C (blue), 27.5 °C (purple), 32.5 °C (dark purple), 37.5 °C (brown) to 42.5 °C (black). The residuals of the global analysis of each temperature point are shown in different colors corresponding to the different temperature for both FRET-pairs.

The intrachain diffusion coefficient revealed a large temperature dependence and increased with increasing temperature (A in Figure 4.4.2). As expected, the viscosity-corrected intrachain diffusion coefficient increased less strongly with increasing temperature compared to the non-viscosity corrected diffusion coefficient (A and B in Figure 4.4.2 A).

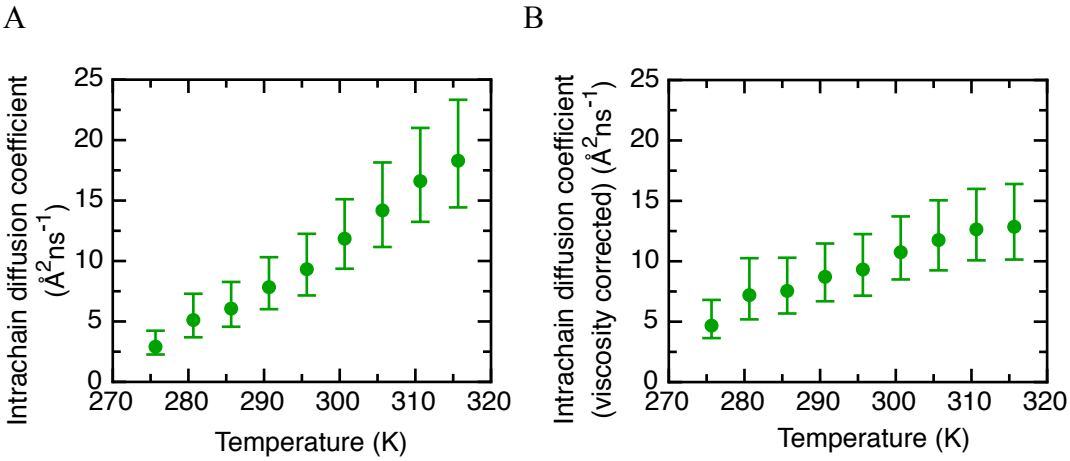


Figure 4.4.2: Temperature dependence of the intrachain diffusion coefficient of Q_{14}. The intrachain diffusion coefficient ranges from 2.5 °C (275.65 K) to 42.5 °C (315.65 K) are obtained from the global analysis **A** without viscosity correction and **B** with viscosity correction (chapter 4.3, equation 4.7).

No significant increase at increasing temperature in the global analysis parameter σ was observed as well as in the parameter b (Figure 4.4.3 A)). Due to the steady parameters σ and b at increasing temperature, a constant root mean square end-to-end distance was observed in the range of calculated error (Figure 4.4.4). The plot of the end-to-end distance distribution illustrates a slight change of the conformational space at particular temperature, corresponding to the fitted parameter σ, whereas a slight overall shift towards lower end-to-end distances is observed at increasing temperature (B in Figure 4.4.3).

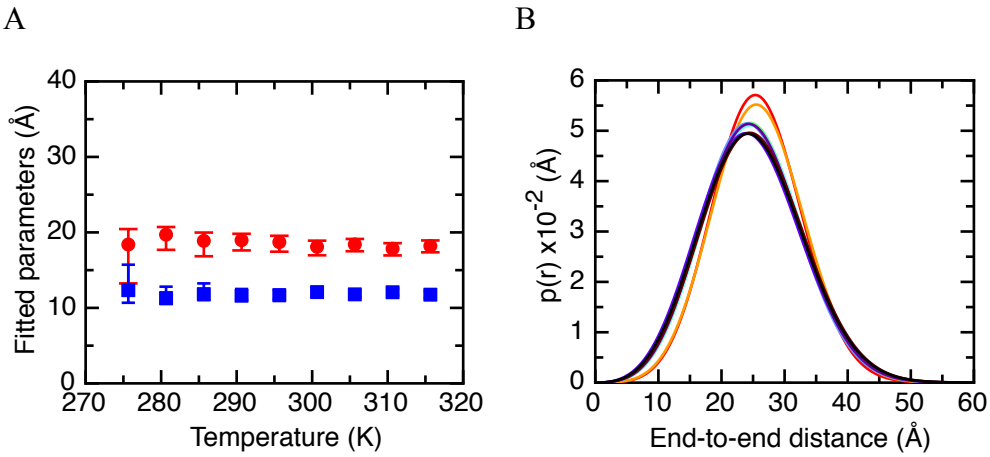


Figure 4.4.3: Global analysis of Q_{14} in dependence of temperature. **A** Fitted parameter *b* (red circles) and σ (blue squares) and **B** the end-to-end distance distribution $p_{eq}(r)$ at different temperature ranges from 2.5 °C (red), 7.5 °C (orange), 12.5 °C (green), 17.5 °C (turquoise), 22.5 °C (blue), 27.5 °C (purple), 32.5 °C (dark purple), 37.5 °C (brown) to 42.5 °C (black).

However, the dimension represented by the root mean square end-to-end distance does not change significantly with changing temperature, taking into account the error (Figure 4.4.4).

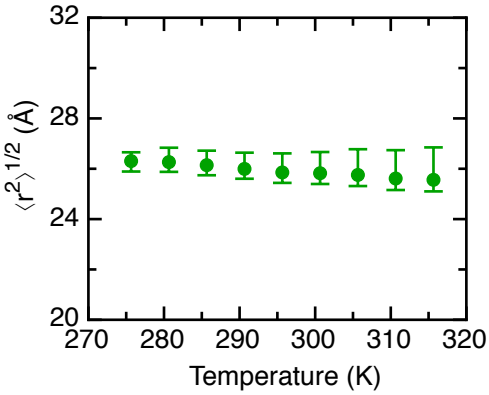


Figure 4.4.4: Root mean square end-to-end distances of Q_{14} in dependence of temperature ranges from 2.5 °C (275.65 K) to 42.5 °C (315.65 K).

The *Arrhenius* activation energy E_A can be calculated by plotting the viscosity-corrected $\ln D_{corr}$ against $1/T$ as shown in Figure 4.4.5 A (equation 4.11). The semi-logarithmic illustration of D_{corr} against $1/T^2$ enables the calculation of the root mean square energy barriers ε from the slope of the linear regression according to the *Zwanzig* model (equation 4.12).

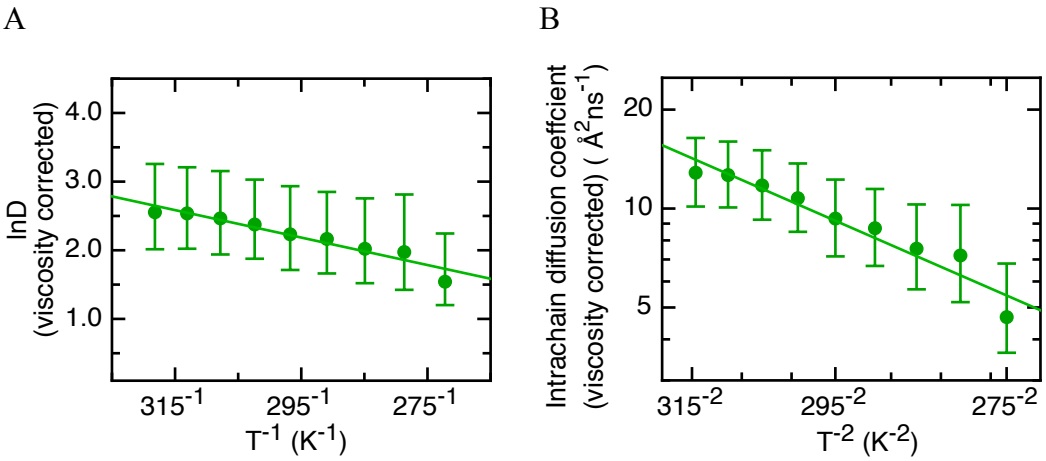


Figure 4.4.5: Temperature dependence of the intrachain diffusion coefficient of Q_{14}. **A** For the calculation of the *Arrhenius* activation energy the natural logarithm of the viscosity-corrected intrachain diffusion coefficient $\ln D_{corr}$ is plotted against *1/T*. **B** For calculation of the root mean square height of the energy barrier according to the *Zwanzig* model the viscosity-corrected intrachain diffusion coefficient is plotted against $1/T^2$.

As can be derived from the linearity of both plots (A and B in Figure 4.4.5) the *Arrhenius* as well as the *Zwanzig* model described the dependence of the intrachain diffusion coefficient on the temperature with comparable accuracy. Considering the data quality and fitting accuracy based on the measured data set, it is not possible to distinguish between the two models. The intrachain diffusion coefficient could be described by both the *Arrhenius*-like behavior of the diffusion comprising entropic and enthalpic contributions and by a one-dimensional diffusion in a rough energy potential given by the *Zwanzig* model. An *Arrhenius* activation energy of 16.6 ± 5.2 kJ/mol was found. The pre-exponential factor *A* of 7987 ± 17050 exhibited a very large fitting error, which is related to the extrapolation to *1/T = 0*, while the data only covers a small temperature range. In accordance to the *Zwanzig* model the root mean square height of the minor energy barrier ε is 4.6 ± 1.3 kJ/mol. The intrachain diffusion coefficient in the absence of energy barriers D^* adopted a value of 298 ± 303 $Å^2$/ns and is subject to a similarly high error resulting from the large extrapolation to high temperatures. However, the calculated D^* was 30 times higher as the intrachain diffusion coefficient at 22.5 °C.

4.4.4 Discussion and Summary

In summary, the *Arrhenius* activation energy of 16.6 ± 5.2 kJ/mol was significantly higher in comparison to the model sequence of $(GS)_8$ of 9.8 ± 3.2 kJ/mol and the EF-loop fragment from the IDP β-carp parvalbumin of 12.7 ± 3.2 kJ/mol measured in previous work [159]. The root mean square height of the energy barrier ε of 1.55 ± 0.04 kcal/mol (6.47 ± 0.17 kJ/mol) for Q_{14} revealed a significantly higher value compared to the $(GS)_8$ with an ε of 0.86 ± 0.03 kcal/mol (3.6 ± 0.2 kJ/mol) and for the EF-loop with an ε of 0.94 ± 0.03 kcal/mol (3.9 ± 0.1 kJ/mol) [159]. The root mean square end-to-end distance does not change at changing temperature. However, the intrachain diffusion coefficient increases more than 7 fold from 2.5 °C to 42.5 °C. These results are in agreement with the simple explanation that the significantly lower intrachain diffusion coefficient of Q_{14} in water can be explained by a higher activation energy to overcome conformational changes due to the comparably less flexible conformers. The resulting higher activation energy or energy barrier of Q_{14}, and thereby emerge low intrachain diffusion coefficient, may be originated from a steric effect of the relatively large side chain group of glutamine, which increases the energy barrier of side chain rotations and the overall dynamics of the polypeptide chains in comparison to the repetitive model peptide poly-(glycine-serine). On the other hand, this explanation seems to be unlikely relating to the EF-loop fragment of the IDP natural β-carp parvalbumin, which comprises a number of amino acids with large side chains like aspartic acid, lysine and isoleucine. However, the occurrence of thermodynamic sub-structures may be another reason for the high activation energy found. The change of the heat capacity Δc_p can be calculated using the *Eyring* plot, where *lnD* is plotted against *1/T*. For Q_{14} a Δc_p = -0.62 ± 0.92 was found. Regarding to the high inaccuracy of Δc_p, it can be concluded that $\Delta c_p \approx 0$ and, therefore, no significant evidence for thermodynamically stable sub-structures was found (Appendix, Figure A.3). Another argument for higher activation energy of Q_{14} might be the hydrogen bonds between side chains and main chains as well as side chains and side chains. In general, the energy of hydrogen bonds ranges from 4 kJ/mol in alpha-helices to 12 kJ/mol for buried bonds in water [8]. If the generally strong dependence of the energy value on the type of the hydrogen donor and acceptor and the surrounded environment are ignored, one more hydrogen bond in the Q_{14} can explain the resulting energy differences compared to the model polypeptide chains. However,

the energy barriers of bond rotation potentials are in the same energetic range [278], [279]. In previous work, GdmCl is discussed to break intramolecular hydrogen bonds upon binding to the peptide chain [280]. In order to study whether a low intrachain diffusion coefficient and a higher conformational energy barrier for Q_{14} can be caused by breaking intramolecular hydrogen bonds, a [GdmCl]-dependence time-resolved FRET experiments were performed.

4.5 The effect of denaturant on dynamics and dimensions of Q_{14}

Guanidine hydrochloride (GdmCl) is one of the strongest known denaturing agents and can be described as a positively charged, delocalized π-electron-system. The origin of the denaturation effect of GdmCl can be found in the interaction of the Gdm^+ ions with the peptide chain. For example, it is known that GdmCl increases the solubility by favoring the free transfer energy of a peptide backbone, which was shown for the model compound di-keto-piperazine (glycine-anhydride) [281]. Additionally, solubility experiments of Nozaki and Tanford *et. al.* determined an increased free energy of transfer for glutamine in GdmCl solution, which corresponds to a higher solubility of glutamine in GdmCl solution [282]. Different models of the mode of interaction of denaturants molecules have been proposed. In comparison to urea, which is known to form hydrogen bonds with the amide group of the peptide [25], the interaction mode of GdmCl is controversially discussed. Previous work predicts that guanidine ions form hydrogen bridges to the peptide carbonyl [247], [283]–[286]. The combination of MD simulations and neutron diffraction data shows that guanidine ions, in principle, can form hydrogen bonds with water [287]. Other studies argue that hydrophobic interaction of guanidine ions with hydrophobic amino acids is more dominant than hydrogen bonds [25], [285]. But such transient stacking interactions was also observed for neutral amine acid like serine, threonine as well as glutamine [285]. However, the hydrogen bridges build from side chain to main chain or from side chain to side chain have to be consequently interrupted by guanidine hydrochloride.

Therefore, in the following mainly the effects of hydrogen bonding are considered and hydrophobic effects will not be further taken into account. Interestingly, possible contributions from the chloride ions of guanidine hydrochloride seemed not to be significant [288]. Möglich *et. al.* showed with triplet-triplet energy transfer experiments of unfolded polypeptide chains, besides urea, that guanidine hydrochloride competes with water molecules for the binding site [280]. Consequently an interruption of transient hydrogen bonds between side chains-main chains of unfolded polypeptide chains took place. Moreover, a weak-binding model of Schellman *et. al.* was adopted to calculate an equilibrium exchange constant for replacing bound water molecules by denaturant molecules [289]. To conclude, GdmCl is a good co-solute to investigate the influence of hydrogen bonds on the dynamics and dimension of Q_{14}.

4.5.1 Effect of GdmCl on dynamics and dimensions of Q_{14}

The GdmCl dependent trFRET experiment, ranging from 1 M to 8 M GdmCl, were performed with 10 mM potassium phosphate buffer and pH 7.0 at 22.5 °C. Time-resolved fluorescence decay curves of both FRET-pairs were recorded, as shown in Figure 4.5.1. The residuals of the globally analyzed fluorescence decay curves of the DNB-Q_{14}-AMCA chain showed a significant deviation at early time points in high GdmCl concentrations (B in Figure 4.5.1).

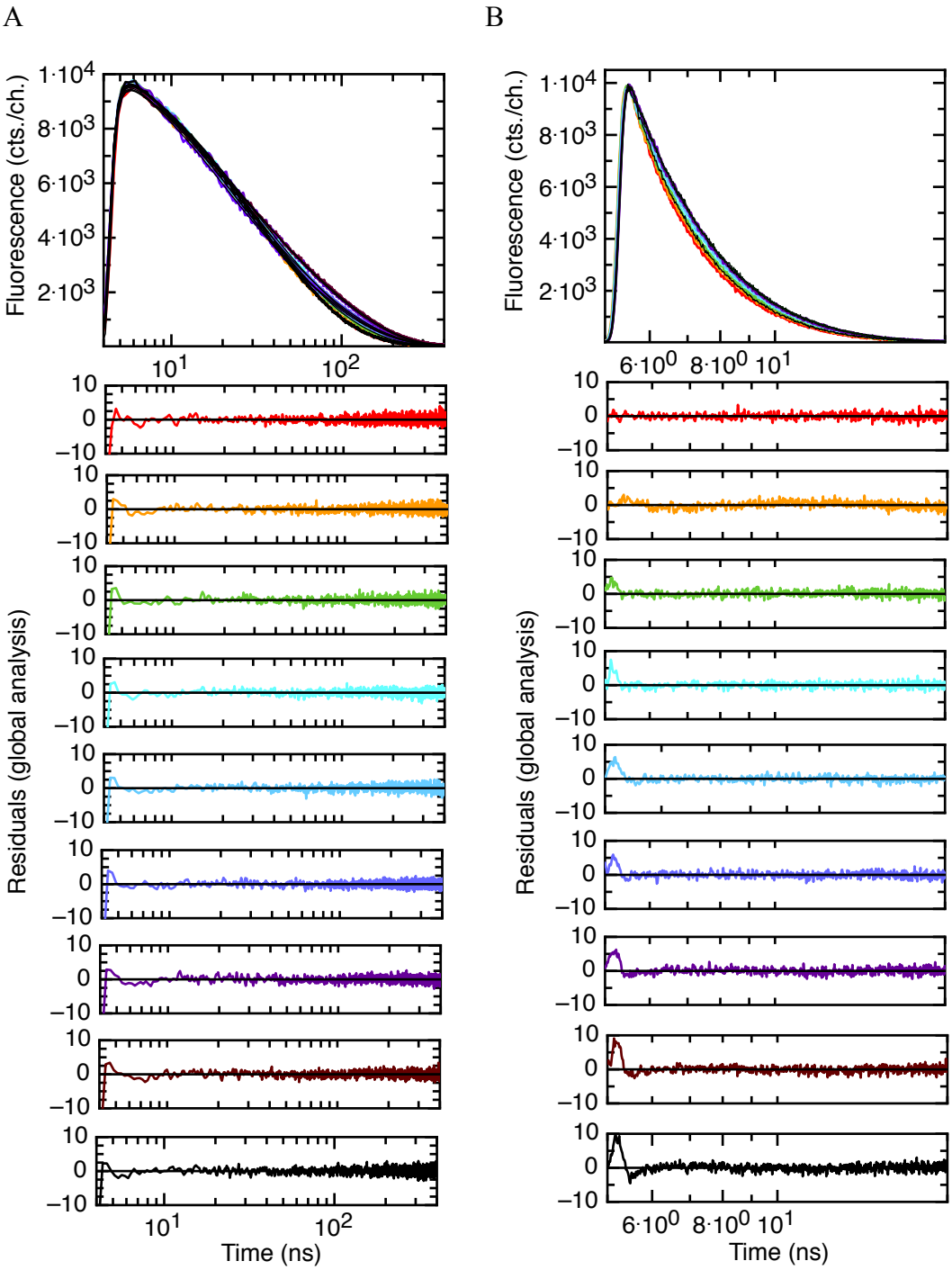


Figure 4.5.1: Time-resolved fluorescence decay curves of Q_{14} in different GdmCl concentrations using donor-acceptor pair **A** dansyl-*pyrene* and **B** DNB-AMCA in 10 mM potassium phosphate pH 7.0 at different denaturant concentrations ranges from 0 M (red), 1 M (orange), 2 M (green), 3 M (turquoise), 4 M (blue), 5 M (purple), 6 M (dark purple), 7 M (brown) to 8 M (black) are shown. The residuals of the global analysis of each concentration point are shown in different colors corresponding to the different concentrations of GdmCl for both FRET-pairs.

Nevertheless, the fluorescence decay, which starts at later times, showed acceptable residuals. Additionally, the global χ^2 depicted an expected accuracy and ranged for all concentration points under 1.3 (except for the 7 M and the 8 M points, where a higher χ^2 of 1.3 was found, which corresponds to the higher inaccuracy of the intrachain diffusion coefficient at high GdmCl concentrations, Figure 4.5.2). As depicted in Figure 4.5.2 the intrachain diffusion coefficient increased with increases GdmCl concentration. In Figure 4.5.2 B the viscosity-corrected intrachain diffusion coefficient calculated with the empirical equation 4.7 using experimentally determined α-value of 0.85 is shown. For the viscosity-corrected intrachain diffusion coefficient an increase of about 4 times from water to 8 M [GdmCl] was observed.

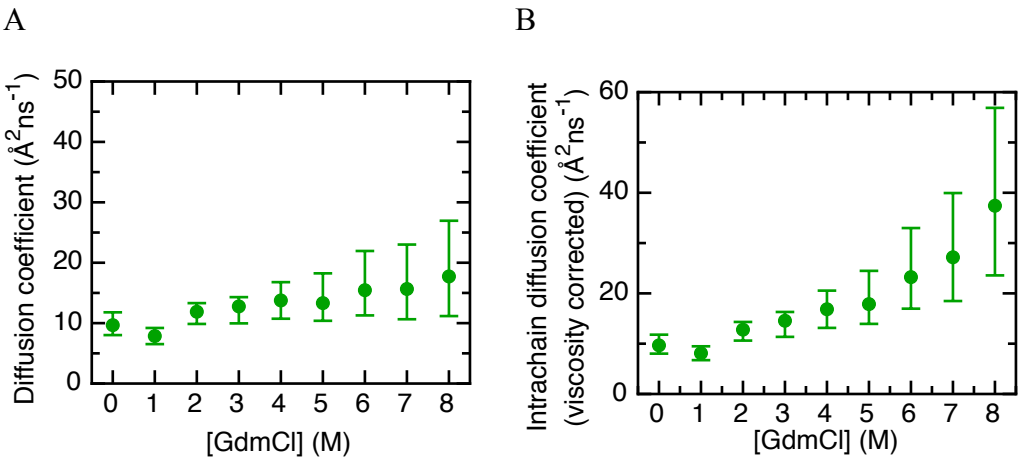


Figure 4.5.2: Intrachain diffusion coefficient of Q_{14} in dependence of GdmCl. From the global analysis calculated intrachain diffusion coefficient is illustrated ranges from 0 M to 8 M of [GdmCl] **A** without viscosity correction and **B** with viscosity correction (chapter 4.3, equation 4.7) using the experimental determined α-value of 0.85.

The parameter *b* increases with increasing denaturant concentration, however a slight decrease of the parameter *σ* was observed (Figure 4.5.2). The population of end-to-end distances calculated with equation 1.12 and parameters *b* and *σ* showed a significant shift to larger end-to-end distances with increasing denaturant concentration that can be correlated to the increase in the value of parameter *b* (B in Figure 4.5.3).

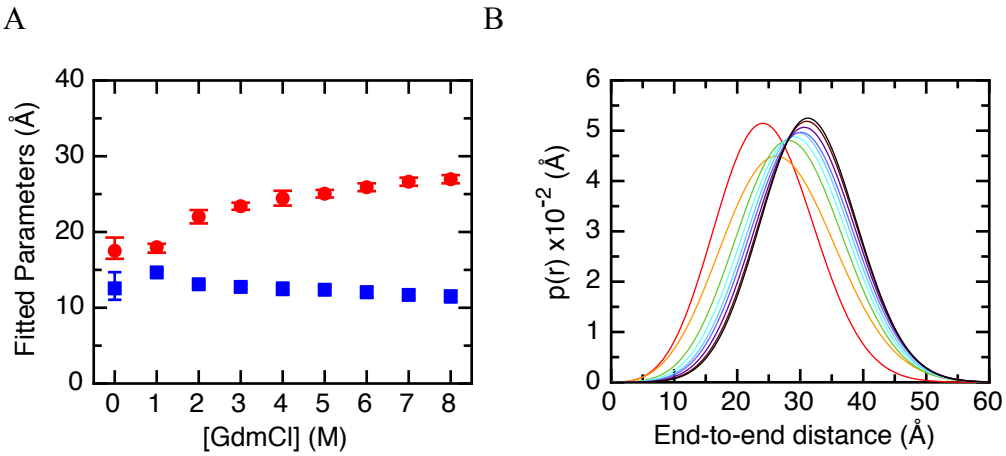


Figure 4.5.3: Global analysis of Q_{14} with increasing GdmCl concentration. **A** Fitted parameter b (red circles) and σ (blue squares) and **B** the end-to-end distance distribution $p_{eq}(r)$ at different denaturant concentrations ranges from 0 M (red), 1 M (orange), 2 M (green), 3 M (turquoise), 4 M (blue), 5 M (purple), 6 M (dark purple), 7 M (brown) to 8 M (black).

The root mean square end-to-end distance rises apparently asymptotically from 25.9 Å in water to 32.4 Å at 8 M [GdmCl]. The observed increase of the root mean square end-to-end distance from 0 M to 8 M [GdmCl] amounts to 20%. In comparison to non-amyloidic model polypeptide chain $(GS)_8$ with 15% and the EF-loop of β-carp parvalbumin with 17% (previous work [159]), a slightly higher increase for Q_{14} was observed.

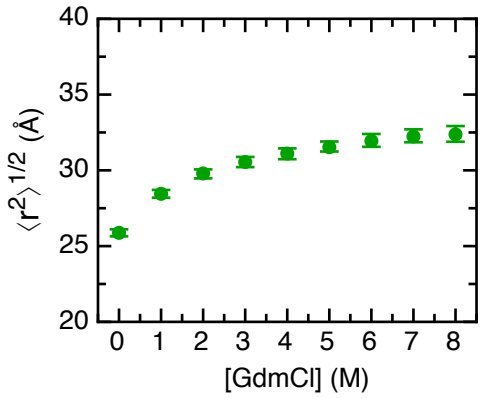


Figure 4.5.4: Dependence of root mean square end-to-end distances of Q_{14} on GdmCl concentration ranges from 0 M to 8 M.

The apparently asymptotically increase of the root mean square end-to-end distance can be analyzed with the weak-binding Schellman model to yields information such as the exchange constant between GdmCl and water molecules as shown in previous publications ([133], [280], [289]).

4.5.2 The weak-binding model of Schellman

Schellman's weak-binding model can be applied to describe the asymptotically increase of the root mean square end-to-end distances with increasing GdmCl concentration [36], [133], [280], [289]–[291]. The denaturant-specific effect on the dynamics as well as on the dimension of a polypeptide chains are described as competition of water molecules for the binding sites with the denaturant molecules. The equilibrium binding constant K_B is here an exchange constant K_{ex}, expressed through the quotient of the binding constant of the ligand (co-solute or denaturant) K_L and the water K_W.

$$K_B = \frac{K_L}{K_W} \equiv K_{ex} \qquad \text{(Eq. 4.13)}$$

In equation 4.14 the binding sites occupancy $\bar{\nu}$ depends on the exchange constant K_{ex}, where an independent binding sites on a polypeptide chains are assumed.

$$\bar{\nu} = \frac{\left(K_{ex} - 1\right) \cdot \mathrm{X}_D}{\left(K_{ex} - 1\right) \cdot \mathrm{X}_D + 1} \qquad \text{(Eq. 4.14)}$$

The denaturant concentration is expressed as the mole fraction of denaturant X_D.

K_{ex} is the exchange equilibrium constant for replacing water molecules with denaturant molecules. With the assumption that all denaturant binding sites on the polypeptide chains are identical and independent of each other, the effect of the denaturant binding on chain dimension is proportional to the binding sites occupancy $\bar{\nu}$ (equation 4.15).

$$\left\langle r^2 \right\rangle^{1/2} = \left\langle r_0^2 \right\rangle^{1/2} + \gamma \cdot \left\langle r_0^2 \right\rangle^{1/2} \cdot \bar{\nu} \qquad \text{(Eq. 4.15)}$$

Where γ reflects the sensitivity of the root mean square end-to-end distance $\langle r^2 \rangle^{1/2}$ to the denaturant. $\langle r_0^2 \rangle^{1/2}$ is the root mean square end-to-end distance in absence of the denaturant. The combination of equations 4.14 and 4.15 yields:

$$\langle r^2 \rangle^{1/2} = \langle r_0^2 \rangle^{1/2} \cdot \left(1 + \gamma \cdot \frac{(K_{ex} - 1) \cdot X_D}{(K_{ex} - 1) \cdot X_D + 1} \right) \qquad \text{(Eq. 4.16)}$$

The root mean square end-to-end distances is plotted against the mole fraction of GdmCl and was analyzed using equation 4.16 to yield the exchange equilibrium constant for replacing water molecules with denaturant molecules K_{ex} (Figure 4.5.5). It should be noted that there is no viscosity correction of the root mean square end-to-end distance necessary, because the glycerol viscosity dependence of Q_{14} as well as the other investigated model polypeptide chains (chapter 4.1) revealed no significant influence of the viscosity on the chain dimension.

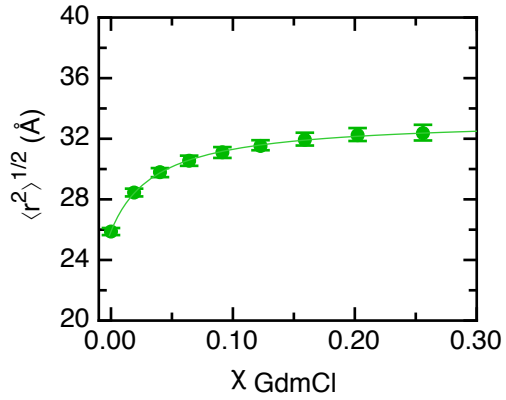


Figure 4.5.5: Root-mean square end-to-end distance of Q_{14} in dependence of GdmCl. The asymptotic increase of the increasing of the root mean square end-to-end distance was described with the weak-binding model of Schellman (equation 4.16).

The increase the root mean square end-to-end distances is well described by the weak-binding model of Schellman. An exchange equilibrium constant K_{ex} of 28.2 ± 6.2 was found. For $(GS)_8$ a K_{ex} of 11 ± 4.3 and for the EF-loop of β-carp parvalbumin a K_{ex} of 10.9 ± 4.3 was found (values taken from [159], with an assumed error of 1.2% of the root mean square end-to-end distance). Thus, a significantly high sensitivity to the

denaturant GdmCl for the Q_{14} in comparison to the previous measured non-amyloidic model polypeptide chains was observed (A in Figure 4.5.6) [159].

The intrachain diffusion coefficient was plotted on a logarithmic scale against the concentration of GdmCl to obtain a better comparison. For Q_{14} a strong increase of the intrachain diffusion coefficient was observed, which reached a comparable value as observed for the both non-amyloidic model polypeptide chains at high GdmCl concentration (8 M). Both non-amyloidic model polypeptide chains, which were measured in the previous work, showed no significant change of the intrachain diffusion coefficient with increasing GdmCl concentration (Figure 4.5.6 B)) [159].

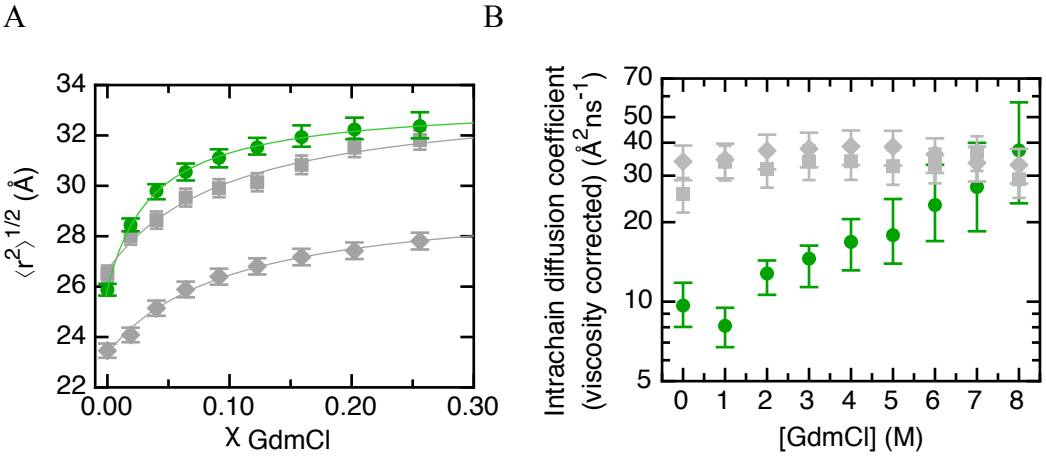


Figure 4.5.6: The effect of GdmCl on dynamics and dimensions. **A** The root mean square end-to-end distance was plotted against the mole fraction of GdmCl and **B** the viscosity-corrected diffusion coefficient is shown on illustrated logarithmic scale against the concentration of GdmCl. The Q_{14} is depicted in green. Values for the previous work for non-amyloidic polypeptide chains are shown in grey. The $(GS)_8$ is depicted in grey squares and the EF-loop of β-carp parvalbumin is shown in grey squares [159]. The asymptotically rise of the root mean square end-to-end distance was described with the weak-binding model of Schellman (equation 4.16). The error of the pervious data was predicted to be around 1.2% as was observed for Q_{14}.

4.5.3 Summary and discussion

Previous time-resolved fluorescence experiments investigated the influence of GdmCl on the intrachain diffusion coefficient and root mean square end-to-end distance of the non-amyloidic polypeptide chains EF-loop of β-carp parvalbumin and artificial $(GS)_8$ [159]. For both chains the root mean square end-to-end distance increased in an

apparently asymptotically manner with increasing GdmCl concentration and can be described by the weak-binding model of Schellman. In comparison to both model polypeptide chains, the amyloidic Q_{14} showed a significantly stronger increase of the root mean square end-to-end distance with increase GdmCl concentration. Similar to both model polypeptide chains, the increase of the root mean square end-to-end distance of Q_{14} is also describable with the weak-binding model and revealed a significantly higher equilibrium exchange constant.

The viscosity-corrected intrachain diffusion coefficient of Q_{14} increased strongly with increasing denaturants concentration. But for both model polypeptide chains no significant change of the intrachain diffusion coefficient with increasing GdmCl concentration was observed [159]. The amyloidic Q_{14} showed a significantly different behavior with respect to GdmCl in the chain dimension and especially in the chain dynamics represented by the intrachain diffusion coefficient. In general, the intrachain diffusion is mainly influenced by steric and ionic effects as well as transient sub-structures, which can be formed by hydrogen bonds. The ionic effect of the glutamine side chain can be excluded, because of absence of charged groups. Nevertheless, electrostatic effects like dipole-dipole interactions cannot be completely ruled out. The rather low observed diffusion coefficient of Q_{14} can be explained on one hand with a simple steric effect, resulting from the large side chain of the glutamine. But another explanation can be given in terms of transient hydrogen bond formation, which can also follow in higher energy barriers and therefore slow down the intrachain diffusion coefficient. Such an effect is *e.g.* described by Lapidus and Eaton for a cysteine-(alanine-glycine-glutamine)$_N$-trypthopane polypeptide chains [114]. Möglich *et. al*. discussed that intramolecular hydrogen bonds are more favorable as compared to the hydrogen bonds between the peptide chain and water, resulting into a slower internal chain dynamics [133], [280].

Additionally, the influence of denaturants *e.g.* guanidine hydrochloride and urea on the kinetics of poly-(glycine-serine) chains with the triplet-triplet-energy transfer experiment was investigated. They found that the denaturants interact with the polypeptide chains, which can by quantitatively described by the weak-binding model of Schellman [133], [280]. Interestingly, guanidine hydrochloride as salt shows the same influences on chain dynamics as urea, a neutral denaturant. Only the denaturation strength of both denaturants, corresponding to the so-called *m*-value, was different [133], [247]. Thus, it can be concluded that the guanidine hydrochloride

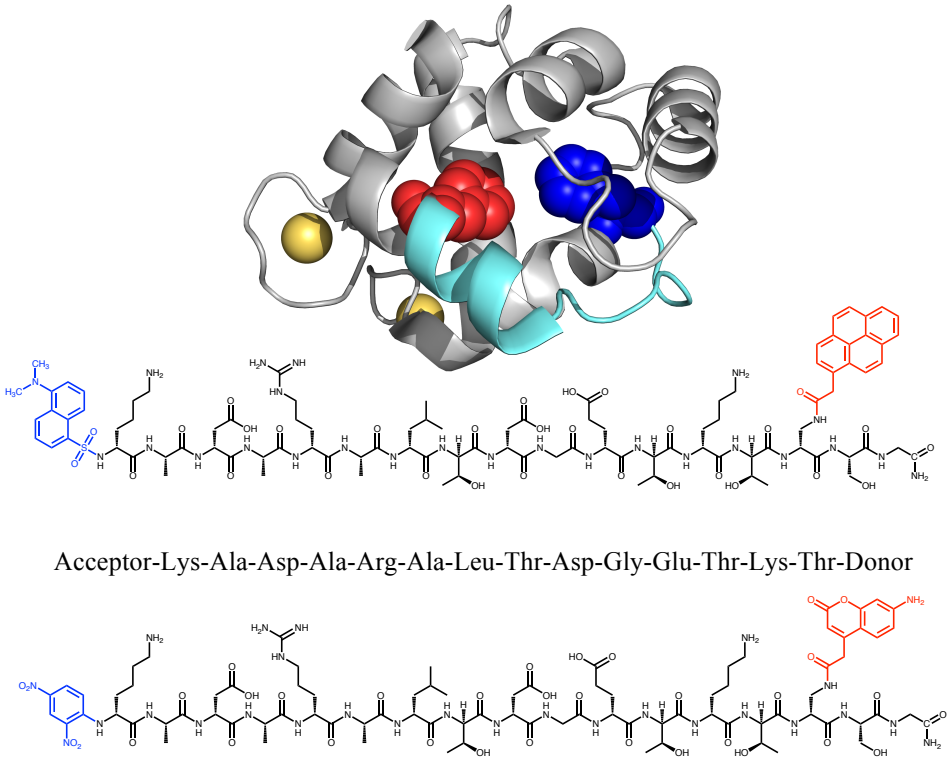
does not affect the dynamics or dimension by an ionic effect of the Gdm^+ ions or Cl^- ions. In conclusion, for Q_{14} a strong effect of denaturant GdmCl on the dynamics and dimension was shown, whereas for the artificial and the natural polypeptide chains no significant influence of GdmCl on the dynamics was observed [159]. The strong effect of GdmCl on the Q_{14} can be explained by interruption of intramolecular hydrogen bonds upon binding of Gdm^+ ions. Of course, the steric effect of the glutamine side chain itself could be another reason.

To get detailed information about whether the intramelocular hydrogen bond network or the steric effect or both together will explain the different behavior of Q_{14} on chain dimension and dynamics, additional model polypeptide chains should be investigated in the following. Polyserine seems to be a suitable model polypeptide chain, which should be stiffer in comparison to poly-(glycine-serine) chains and includes more hydrogen bond units. Additionally, the $(serine)_{14}$ polypeptide chain (S_{14}) represents a further non-amyloidic polypeptide. Furthermore, another model polypeptide chain, the DE-loop of IDP β-carb parvalbumin, was investigated. In comparison to the previously investigated EF-loop, the DE-loop includes only one glycine instead of four found in the EF-loop and consequently should represents stiffer chain. In addition, the investigated EF-loop consists of 16 amino acids between the labels, whereas the DE-loop has the same number of amino acids like the investigated Q_{14}. So that both polyserine, with a higher number of hydrogen bond forming units, and the DE-loop, with comparable length, seem to be suitable non-amyloidic model polypeptide chains for the investigation of an influence of the steric effect as well as the hydrogen bonds on chain dynamics and dimension in comparison to the amyloidic Q_{14}. This is addressed in the next sections.

4.6 Investigation of non-amyloidic model polypeptide chains S_{14} and DE-loop

As discussed, the low intrachain diffusion coefficient, the higher activation energy and the strong effect of GdmCl on the chain dimension and dynamics of Q_{14} can be explained by a steric effect of the side chains and/or intramolecular hydrogen bonds. For comparison, two additional model peptide are chosen, one that consists of the same number amino acids and hydrogen bonding capability with smaller side chains

as glutamine, and another, the DE-loop, which represents a natural IDP with comparable length and low amount of hydrogen bond forming units. The DE-loop as well as the above discussed EF-loop are fragments from the β-carp parvalbumin. β-carp parvalbumin is a member of the IDP's, which form the tertiary structure upon binding of two calcium ions (Figure 4.6.1).



Acceptor-Lys-Ala-Asp-Ala-Arg-Ala-Leu-Thr-Asp-Gly-Glu-Thr-Lys-Thr-Donor

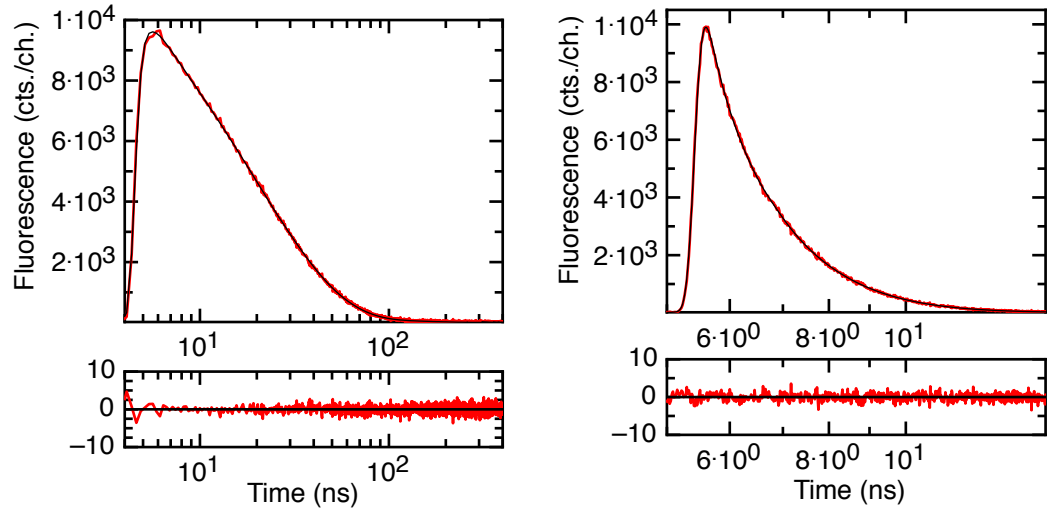
Figure 4.6.1. DE-loop of β-carp parvalbumin. At the top: Tertiary structure of β-carp parvalbumin. Upon binding of the both calcium ions (yellow spheres) the tertiary structure is formed. The fragment DE-loop is colored in cyan and flanked by two phenylalanine. For the time-resolved FRET measurements the both flanking phenylalanines were replaced by the FRET acceptor (blue) and donor (red). At the bottom: Scheme of the both investigated chains with the FRET pair dansyl-DE-loop-pyrene and DNB-DE-loop-AMCA. For all model polypeptide chain a solubility tail of serine-glycine-NH_2 was coupled on the C-termini. The figure was prepared using MacPyMOL and the PDB file 4CPV.

For both fragments, the EF- and the DE-loop, the phenylalanines on both ends are replaced by FRET chromophores. The replacement of these amino acids by FRET chromophores should minimize an effect of the coupled chromophores on chain dimension and dynamics (Figure 4.6.1).

4.6.1 trFRET measurements on S_{14} and DE-loop in aqueous solution

The trFRET measurements for the DE-loop was performed in 10 mM potassium phosphate buffer at pH 7.0 and 22.5 °C (Figure 4.6.1).

A



B

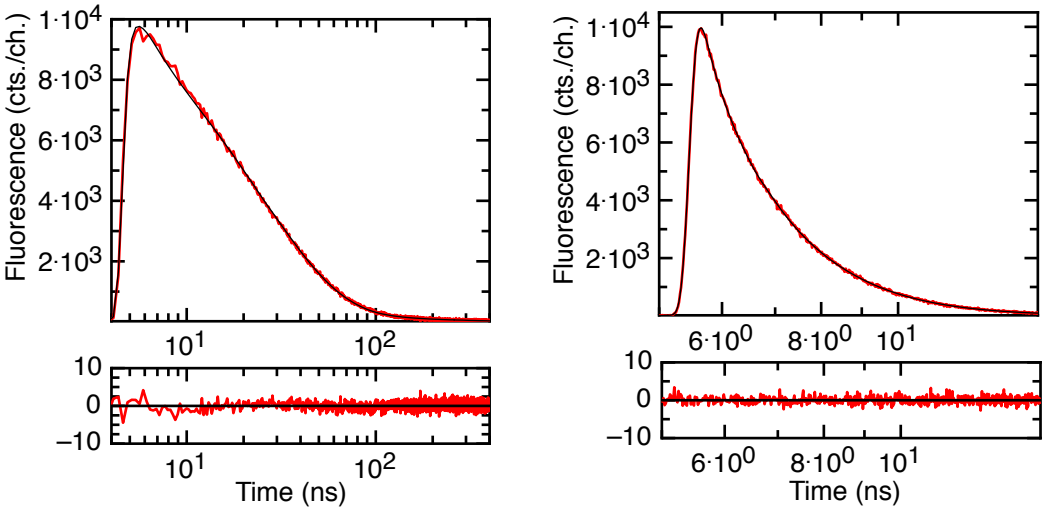


Figure 4.6.1: Time-resolved fluorescence decay curves of **A** S_{14} and **B** the DE-loop (blue lines). The global analysis using equation 1.12 and 2.1 are shown in black with the corresponding residuals at the bottom. On the left side: dansyl-*pyrene* and on the right side DNB-AMCA fluorescence decay curves are shown. The S_{14} was dissolved in pure water without potassium phosphate and the pH was adjusted to 7.0 ± 0.3 with sodium hydroxide.

For S_{14} potassium phosphate buffer could not be used, because of precipitation of the S_{14}. As a consequence, the polypeptide chains were dissolved in fresh pure water and the pH was adjusted to pH 7.0 with a high-diluted sodium hydroxide. The accuracy of the pH measurements under low salt solutions is limited to about pH ± 0.3. In order to exclude the effect of the pH on chain dimension and dynamics, also trFRET measurements as well as R_0 and τ_D determination were performed in pure water, which results in a pH of about 5 ± 0.3 (Appendix, Figure A.4). The R_0 yielded a shift of about 15% compared from pH 5 and pH 7, but trFRET measurements on both pH 5.0 and 7.0 showed no significant deviation of the resulting fitted parameters. Thus no influence of the pH range from pH 5 to pH 7 was observed and the uncertainty of the pH determination at the sample preparation was acceptable.

The global analysis of the fluorescence decay curves yielded intrachain diffusion coefficient of 17.9 ± 3.4 $Å^2$/ns and of 22.3 ± 4.5 $Å^2$/ns, for S_{14} and DE-loop respectively. The calculated root mean square end-to-end distance for S_{14} was 22.8 ± 0.5 Å and for the DE-loop 25.9 ± 0.6 Å. As expected, the S_{14} showed a lower intrachain diffusion coefficient compared to the two-fold higher value obtained for $(GS)_8$ measured in previous work. The root mean square end-to-end distance of S_{14} was 3% smaller in comparison to $(GS)_8$ and can be explained by the 2 amino acid longer $(GS)_8$. The DE-loop showed comparable values of both the intrachain diffusion coefficient and the root mean square end-to-end distance as observed for EF-loop in the previous measurements (Table 3).

As depicted in Figure 4.6.2 and Table 3, the end-to-end distance distributions are similar for the DE-loop and Q_{14}. The non-amyloidic S_{14} showed a larger deviation compared to the Q_{14}, which corresponds to the lower root mean square end-to-end distance.

Interestingly, the intrachain diffusion coefficient of Q_{14} was two times lower compared to the S_{14}, whereas the root mean square end-to-end distance of Q_{14} was significantly higher as for S_{14}.

Table 3: Comparison of the amyloidic Q_{14} with the non-amyloidic model polypeptide chains in aqueous solution. Both $(GS)_8$ and the EF-loop are taken from previous work [159] and are depicted in grey. All model polypeptide chains were measured in 10 mM potassium phosphate at 7.0 pH and 22.5 °C. The S_{14} was dissolved in pure water without potassium phosphate and the pH was adjusted to 7.0 ± 0.3 with sodium hydroxide.

model peptide	*number of amino acid*	*intrachain diffusion coefficient [Å²/ns]*	*root mean square end-to-end distance $\langle r^2 \rangle^{1/2}$ [Å]*
Q_{14}	14	9.7± 2.1	25.9 ± 0.3
S_{14}	14	17.9± 3.4	22.8 ± 0.5
DE-loop	14	22.3± 4.5	25.9 ± 0.6
EF-loop	16	22.2± 3.8	26.5
$(GS)_8$	16	33.9± 5.8	23.5

Also for the Q_{14} a similar root mean square end-to-end distance was observed compared to the DE-loop, with the same number of amino acids. This demonstrates that the Q_{14} forms a more expanded chain and that the rather slow chain dynamics are not caused by chain compression.

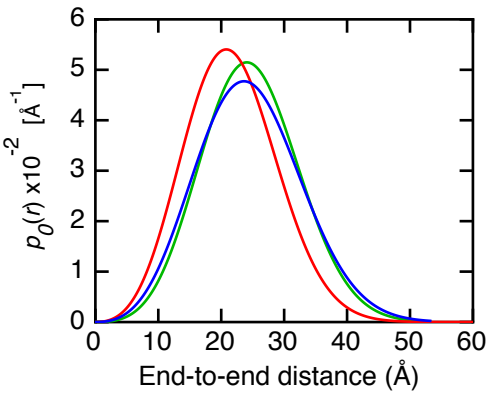


Figure 4.6.2: Comparison of the end-to-end distance distribution of Q_{14} (green line), S_{14} (red line) and DE-loop (blue line).

4.6.2 The excluded volume and segment length

The EDWARDS model for polypeptide chains includes the parameter b and σ to describe the end-to-end distance distribution (chapter 1.4). But b and σ are simplified terms, where b is related to the excluded volume v, segment length l and segment number n, while σ is related to the segment length l and number n (chapter 1.4, equation 1.12 a and b). Assuming constant contour length $L = n \cdot l$, the excluded volume v and the segment length l can be calculated from b and σ (Table 4). For the contour length, the distances of 3.8 Å for one amino acid times the number of amino acids of the model polypeptide chain is assumed. To consider the additional distance of the coupled chromophores, two amino acids are included [159].

Table 4: Comparison of non-amyloidic model polypeptide chains and amyloidic Q_{14}. The values in grey are taken from previous work [159]. The shown parameters are derived from equation 1.12 a and b.

peptide	*parameter b (Å)*	*parameter σ (Å)*	*contour length L (Å)*	*segment length l (Å)*	*excluded volume v*
Q_{14}	17.51	12.56	60.8	3.5	52.2
S_{14}	14.46	12.32	60.8	3.3	19.3
DE-loop	14.30	14.65	60.8	4.8	25.8
EF-loop	15.3	14.6	68.4	4.2	22.4
$(GS)_8$	12.4	13.6	68.4	3.6	6.8

As shown in Table 4, S_{14} has the smallest segment length, whereas the DE-loop has the largest one. Interestingly, the intrachain diffusion coefficient for Q_{14} was found to be rather low, however its segment length is comparable to the more flexible chains such as S_{14} or $(GS)_8$. Nevertheless, the segment lengths are in the range of the average length of 3.8 Å for one amino acid. The excluded volume listed in Table 4 generally describes the space that is not accessible for other particles. For Q_{14} the excluded volume was more than 7 times larger compared to the flexible $(GS)_8$ and about 2 times larger compared to the natural EF- and DE-loop. Interestingly, the same behavior was found for the intrachain diffusion coefficient (Table 3), which shows a possible relation between the excluded volume and the chain dynamics.

4.7 Effect of glycerol on the dimensions and dynamics of S_{14} and DE-loop

In chapter 4.1 the influence of the viscogenic co-solute glycerol on different model polypeptide chain was shown. With the empirical equation 4.7 the observed viscosity dependence of the intrachain diffusion coefficient can be well described. The viscosity-sensitive exponent α, determined using equation 4.7, can thus be used to correct viscosity dependence of the intrachain diffusion coefficient. Thereby detailed information regarding the effect of denaturants and temperature and their effects on the viscosities can be considered. The time-resolved FRET measurements for the DE-loop chain are shown and discussed in chapter 4.1. The time-resolved FRET experiment of S_{14} was performed in pure water pH 7.0 ±0.3 at 22.5 °C ranging from 16% to 56% (w/w) glycerol. The pH was adjusted to pH 7.0 ±0.3 with sodium hydroxid.

4.7.1 trFRET measurements on S_{14} and DE-loop at different glycerol concentrations

Time-resolved FRET measurements on S_{14} and DE-loop were performed ranges from 0% to 56% (w/w) glycerol. The DE-loop was measured in 10 mM potassium phosphate buffer ph 7.0 at 22.5 °C (chapter 4.1.3, Figure 4.1.1), whereas the S_{14} was dissolved in pure water without potassium phosphate and the pH was adjusted to 7.0 ± 0.3 with sodium hydroxide (Fig. 4.7.1).

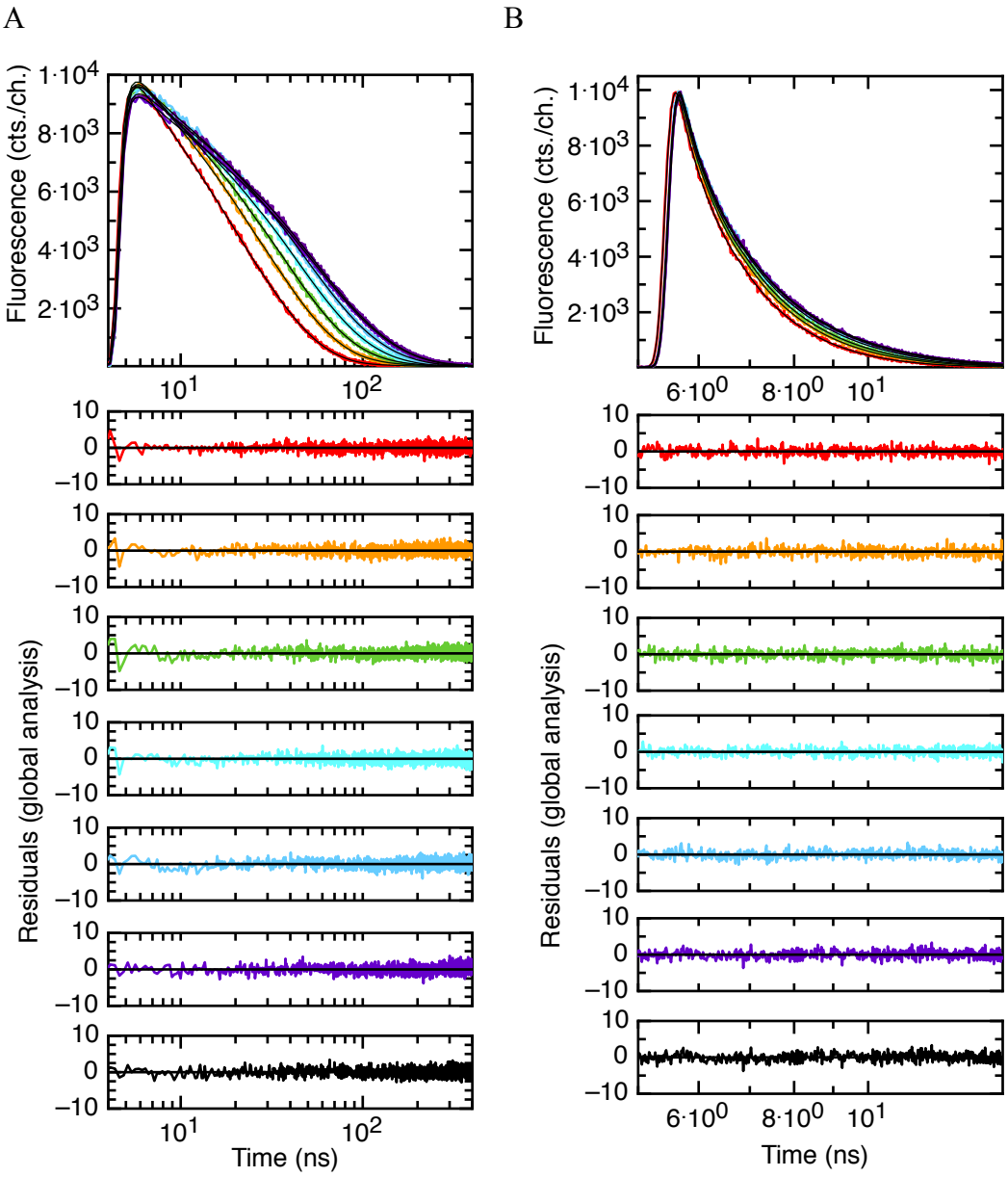


Figure 4.7.1: Time-resolved fluorescence decay curves of S_{14} in dependence of glycerol using donor-acceptor pair **A** dansyl-*pyrene* and **B** DNB-AMCA pH 7.0 at different glycerol concentration ranges from 0% (red), 16% (orange), 28% (green), 40% (turquoise), 48% (purple) to 56% (black) (w/w). The residuals of the global analysis of each concentration point are shown in different colors corresponding to the different concentrations of glycerol for both FRET-pairs. The S_{14} was dissolved in pure water without potassium phosphate and the pH was adjusted to 7.0 ± 0.3 with sodium hydroxide.

As expected the intrachain diffusion coefficient decreased with increasing glycerol concentration (A in Figure 4.7.2). The parameter *b* slightly increased, whereas the parameter σ decreased with increasing glycerol concentration.

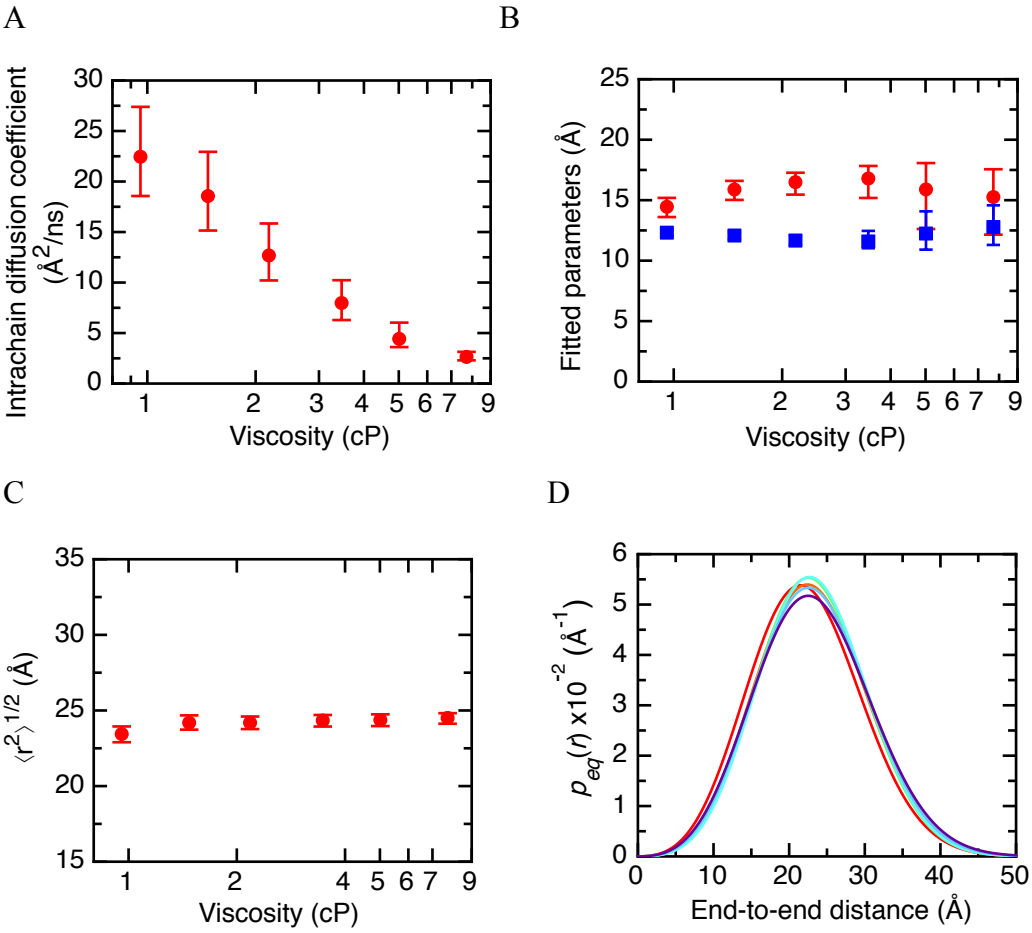


Figure 4.7.2: Glycerol dependence of dynamics and dimensions of S_{14}. **A** Semi-logarithmic illustration of the intrachain diffusion coefficient against the viscosity. **B** Semi-logarithmic illustration of the fitting parameters *b* (red circles) and σ (blue squares). **C** The root-mean square end-to-end distance calculated with equation 1.12 and **D** Linear plot of the end-to-end distance distribution $p_{eq}(r)$ at different glycerol concentrations ranges from 0% (red), 16% (orange), 28% (green), 40% (turquoise), 48% (blue) to 56% (purple) (w/w).

The root mean square end-to-end distance showed a very slight increase with increasing glycerol concentration, which can be justified by the transfer of the polypeptide chains from pure water to the viscogenic co-solute solution (Figure 4.7.2 C)). Nevertheless, this observation is within the calculated error of the root mean square end-to-end distance, which means no significant change of the chain structure

can be observed in the glycerol dependence of S_{14}. The viscosity dependence of the intrachain diffusion coefficient can be analyzed with the empirically equation 4.7, where the viscosity-sensitive exponent α can be determined (Figure 4.7.3). A α-value of 1.04 ± 0.1 for S_{14} and a α-value of 1.14 ± 0.1 for the DE-loop was calculated (chapter 4.1 and Figure 4.7.3).

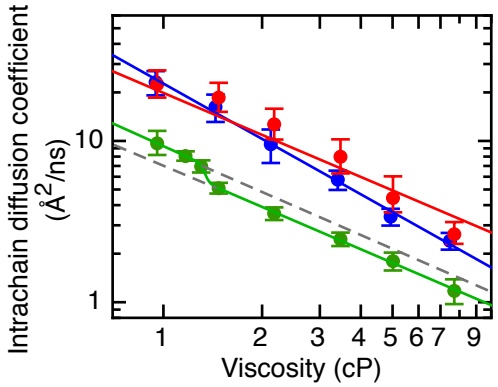


Figure 4.7.3: Double-logarithmic illustration of the intrachain diffusion coefficient against the viscosity of glycerol ranges from 0% to 56% (w/w) glycerol. Q_{14} (green circles), S_{14} (red circles) and DE-loop (blue circles). The data were analyzed using equation 4.7 (corresponding lines).

4.7.2 Summary and Discussion

For the intrachain diffusion coefficient of S_{14} a higher deviation of the fit was observed, but the data does not allow a detailed analysis as performed for Q_{14} (chapter 4.3.1). The α-value for S_{14} showed the behavior of $D = 1/\eta$ predicted by Kramer, as shown for the other non-amyloidic polypeptide chains. Thus, the determined α-values for the S_{14} and the DE-loop can be used for the viscosity correction of the intrachain diffusion coefficient for the temperature and GdmCl dependencies. (detail information for the results of the glycerol dependence of the DE-loop can be found in chapter 4.1).

4.8 Effect of temperature on the dynamics and dimensions of S_{14} and DE-loop

The amyloidic Q_{14} exhibited a significantly high *Arrhenius* activation energy in comparison to the non-amyloidic $(GS)_8$ and the EF-loop, which may explain the comparably slow intrachain diffusion for Q_{14} in water (chapter 4.3).
Time-resolved FRET measurements of S_{14} and DE-loop at different temperatures were performed to get a deeper understanding of the origin of the particular behavior of the amyloidic Q_{14} as discussed in chapter 4.5 and 4.6.

4.8.1 trFRET measurements on S_{14} and DE-loop at various temperature

The time-resolved FRET experiments were carried out in 10 mM potassium phosphate buffer, pH 7.0. The temperature ranges from 2.5 °C (275.65 K) to 42.5 °C (315.65 K) and was globally analyzed with equation 1.12 and 2.1 (Figure 4.4.1 and 4.4.2). The time-resolved FRET experiment of S_{14} was measured in pure water, pH 7.0 ±0.3 and 22.5 °C. The pH was adjusted to pH 7.0 ±0.3 with sodium hydroxid.
As depicted in Figure 4.8.1 and 4.8.2, well-defined time-resolved fluorescence decay curves for both model polypeptide chains and each temperature point were obtained. For the lowest temperature a slight deviation of the fit for the dansyl-*pyrene* fluorescence decay curves was observed.

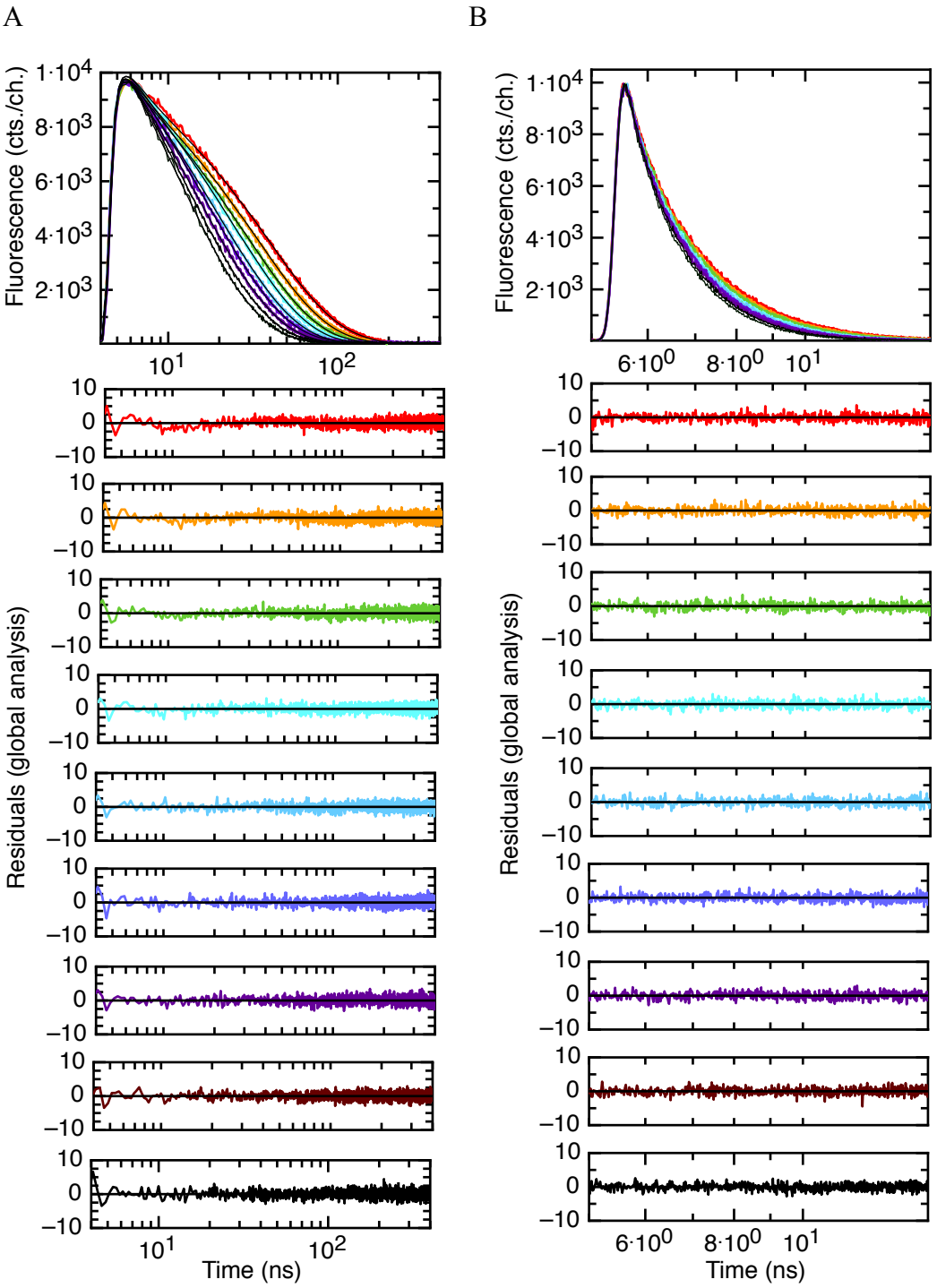


Figure 4.8.1: Time-resolved fluorescence decay curves of S_{14} in dependence of temperature using donor-acceptor pair **A** dansyl-*pyrene* and **B** DNB-AMCA pH 7.0 at different temperature ranges from 2.5 °C (red), 7.5 °C (orange), 12.5 °C (green), 17.5 °C (turquoise), 22.5 °C (blue), 27.5 °C (purple), 32.5 °C (dark purple), 37.5 °C (brown) to 42.5 °C (black). The residuals of the global analysis of each temperature point are shown in different colors corresponding to the different temperature for both FRET-pairs. The S_{14} was dissolved in pure water without potassium phosphate and the pH was adjusted to 7.0 ± 0.3 with sodium hydroxide.

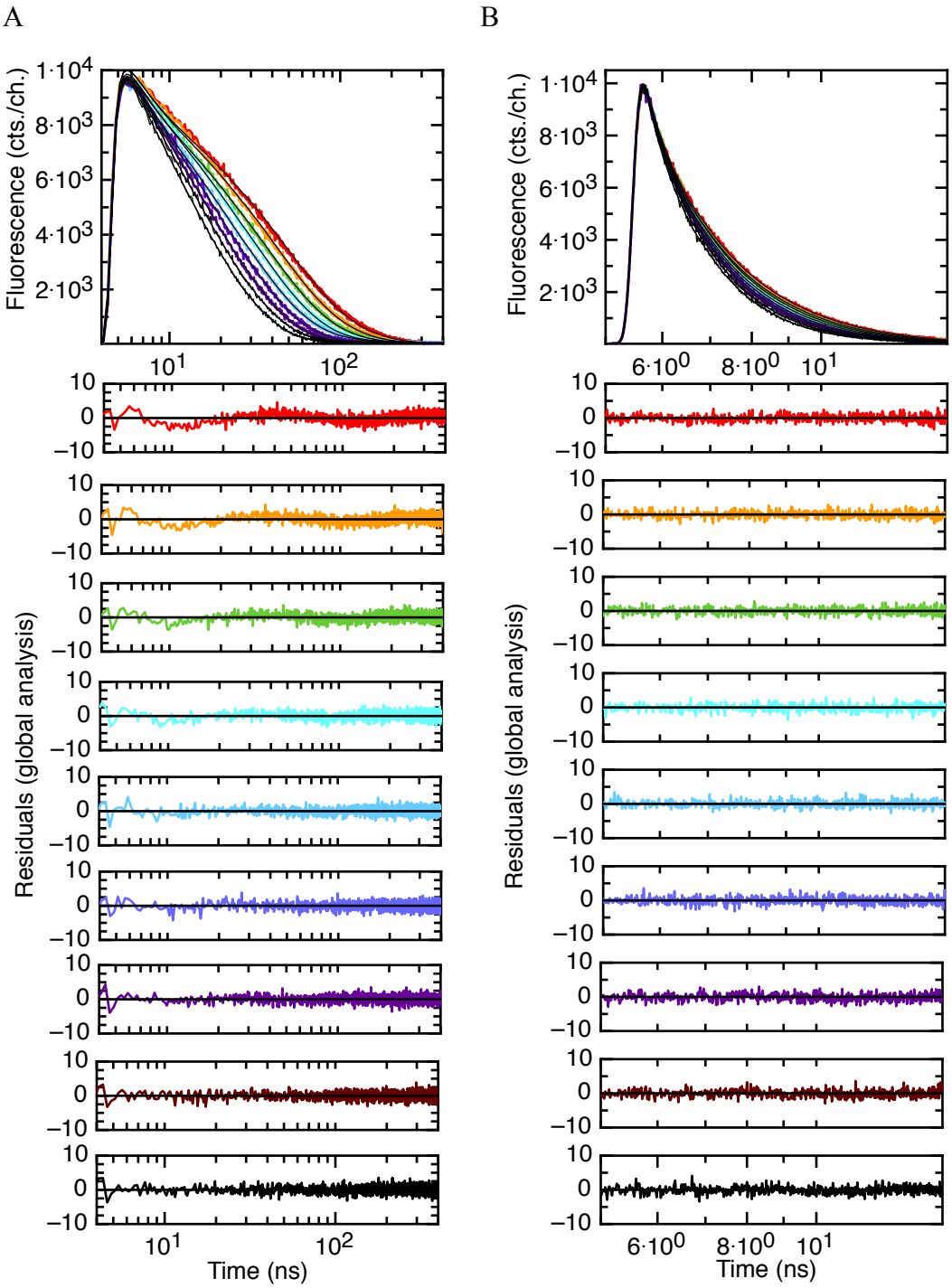


Figure 4.8.2: Time-resolved fluorescence decay curves of DE-loop from the IDP β-carp parvalbumin in dependence of temperature using donor-acceptor pair **A** dansyl-*pyrene* and **B** DNB-AMCA in 10 mM potassium phosphate pH 7.0 at different temperature ranges from 2.5 °C (red), 7.5 °C (orange), 12.5 °C (green), 17.5 °C (turquoise), 22.5 °C (blue), 27.5 °C (purple), 32.5 °C (dark purple), 37.5 °C (brown) to 42.5 °C (black). The residuals of the global analysis of each temperature point are shown in different colors corresponding to the different temperature for both FRET-pairs.

The intrachain diffusion coefficient increased with increasing the temperature as shown in A in Figure 4.8.3 for both model polypeptide chains. The intrachain diffusion coefficient was viscosity-corrected using equation 4.7 and the experimentally determined viscosity-sensitive exponent α (B in Figure 4.8.3).

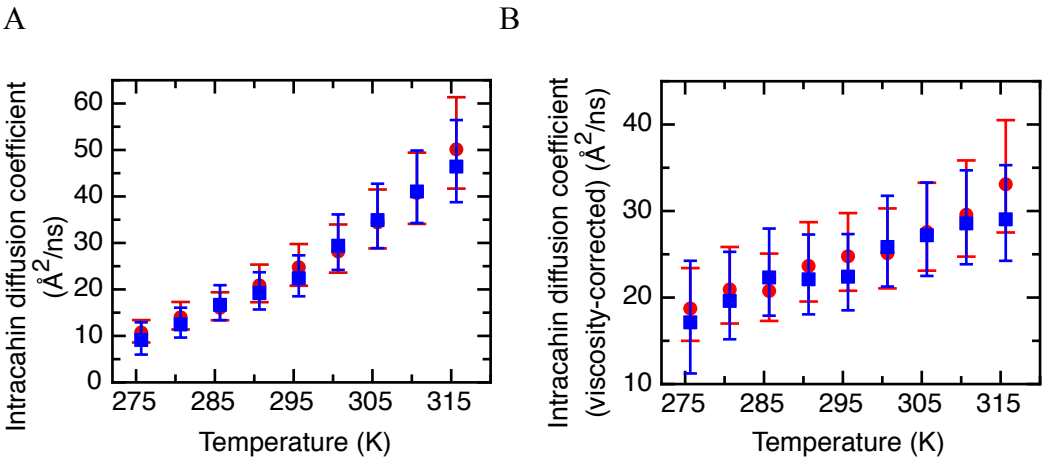


Figure 4.8.3: Temperature dependence of the intrachain diffusion coefficient of S_{14} (red circles) and DE-loop (blue squares). **A** Intrachain diffusion coefficient is plotted against the temperature of both model polypeptide chains. **B** viscosity-corrected intrachain diffusion coefficient was plotted against the temperature and was calculated using the empirical equation 4.7 with the measured α-value of the corresponding model polypeptide chains. The temperature ranges from 2.5 C (275.65 K) to 42.5 °C (315.65 K). The S_{14} was dissolved in pure water without potassium phosphate and the pH is adjusted to 7.0 ± 0.3 with sodium hydroxide.

The fitted parameters *b* and σ obtained from the global analysis are shown in Figure 4.8.4. The plot on the left side shows values for the S_{14} and for the right side shows values for the DE-loop. For both model polypeptide chains the observed fitted parameter *b* decreased with increasing temperature, whereas the fitted parameter *σ* increased with increasing temperature (A in Figure 4.8.4). The population of end-to-end distance of both model polypeptide chains is shown in B in Figure 4.8.4. The corresponding root mean square end-to-end distance showed no significant change with increasing temperature for both model polypeptide chains (C in Figure 4.8.4).

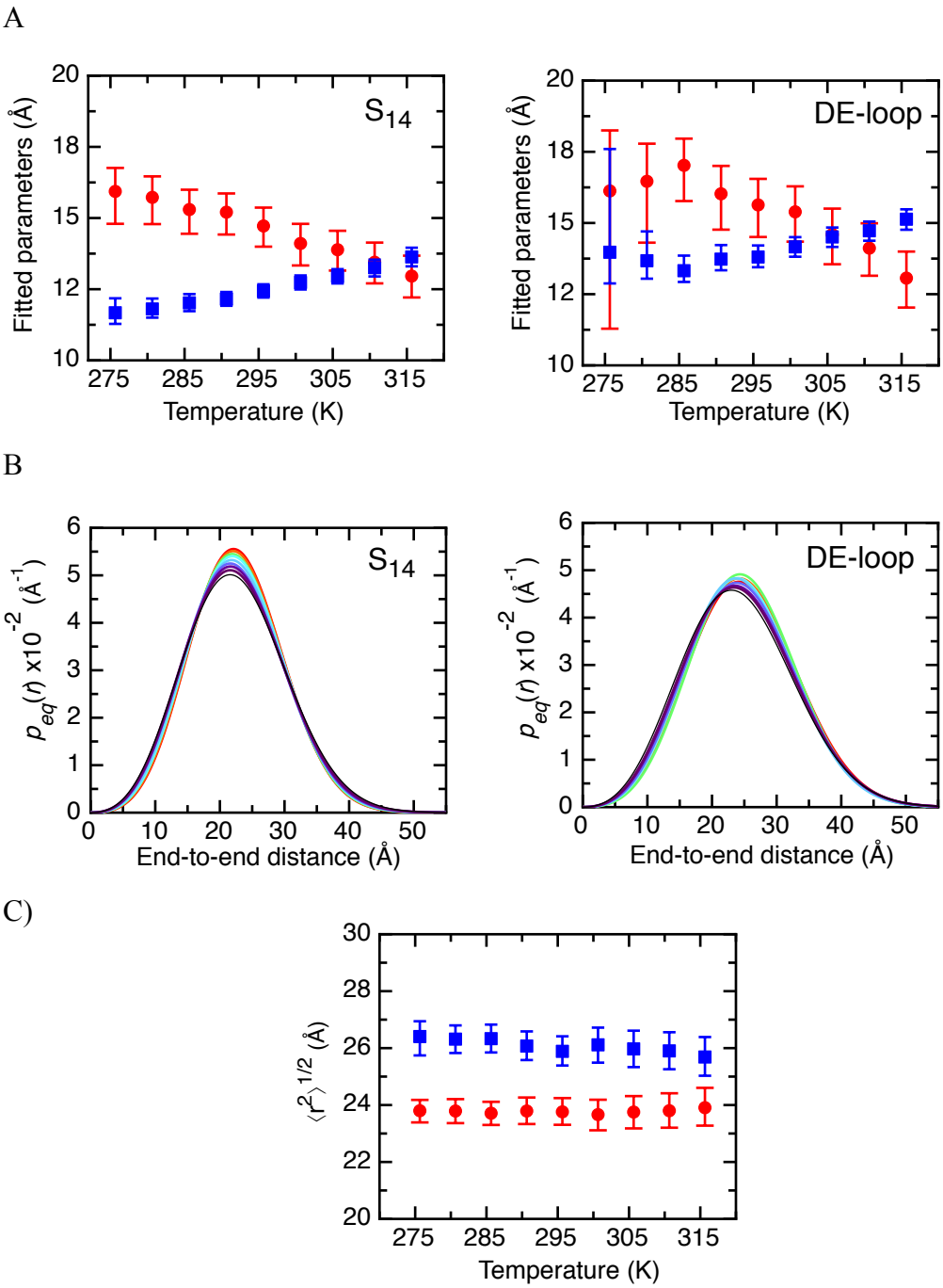


Figure 4.8.4: Global analysis of S_{14} and DE-loop. **A** Fitted parameters *b* (red circles) and σ (blue squares) of both model polypeptide chains. S_{14} left side and DE-loop right side. **B** The end-to-end distance distribution against the end-to-end distance of both model polypeptide chains. Left side S_{14} and right side DE-loop. The temperature ranges from 2.5 °C (red), 7.5 °C (orange), 12.5 °C (green), 17.5 °C (turquoise), 22.5 °C (blue), 27.5 °C (purple), 32.5 °C (dark purple), 37.5 °C (brown) to 42.5 °C (black). **C** Root mean square end-to-end distance calculated from the corresponding fitted parameters *b* and σ and equation 1.12 of S_{14} (red circles) and DE-loop (blue squares). The S_{14} was dissolved in pure water without potassium phosphate and the pH was adjusted to 7.0 ± 0.3 with sodium hydroxide.

For the DE-loop at the low temperature a strong increase of the error of both fitted parameters *b* and *σ* were found. Nevertheless, the root mean square end-to-end distance could be determined with high accuracy. The temperature dependence of the intrachain diffusion coefficient can be described with both the *Arrhenius* formula (equation 4.11) as well as the *Zwanzig* model (equation 4.12) as discussed in chapter 4.4. For the *Arrhenius* formula the viscosity-corrected natural logarithms of the intrachain diffusion coefficient was plotted against *1/T* to calculate the *Arrhenius* activation energy E_A. With the *Zwanzig* model the root mean square barrier height *ε* can be determined, where the viscosity-corrected intrachain diffusion coefficient was plotted against $1/T^2$ (Figure 4.8.5).

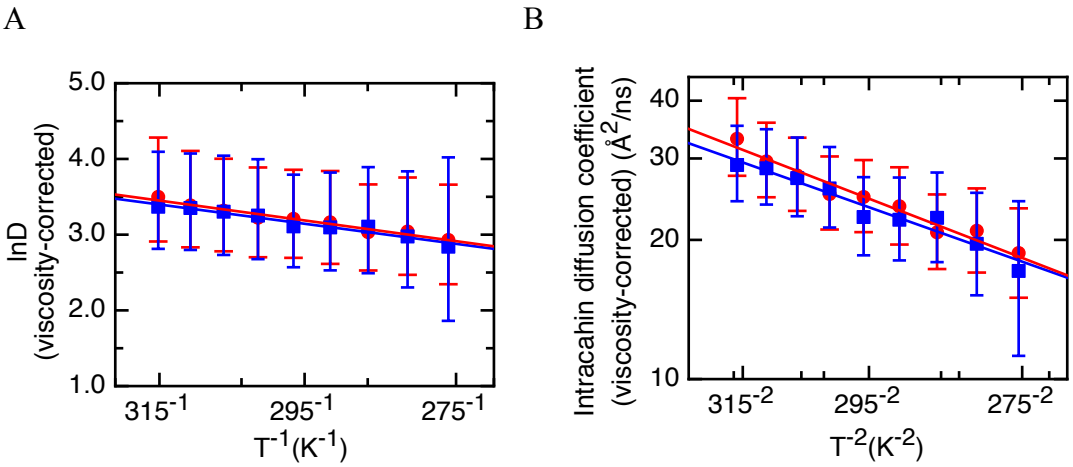


Figure 4.8.5: Temperature dependence of S_{14} (red circles) and the DE-loop (blue squares). **A** For calculation of the *Arrhenius* activation energy the natural logarithmic plot of the viscosity-corrected intrachain diffusion coefficient $\ln D_{corr}$ is plotted against $1/T$. **B** For calculation of the root-mean square height of the energy barrier of the *Zwanzig* model the natural logarithmic of the viscosity-corrected intrachain diffusion coefficient is plotted against the $1/T^2$. The S_{14} was dissolved in pure water without potassium phosphate and the pH was adjusted to 7.0 ± 0.3 with sodium hydroxide.

For the S_{14} and the DE-loop a *Arrhenius* activation energies of 9.5 ± 3.4 kJ/mol and 9.2 ± 3.9 kJ/mol were found, respectively. Fitting the data with the *Zwanzig* model resulted in a root mean square barrier height of ε = 3.5 ± 0.5 kJ/mol and ε = 3.3 ± 0.6 kJ/mol for S_{14} and the DE-loop, respectively. Note that for Q_{14} a *Arrhenius* activation energy of 16.6 ± 5.2 kJ/mol and a root mean square barrier height of ε = 4.6 ± 1.3 kJ/mol was found.

4.8.2 Discussion and Summary

The calculated root mean square end-to-end distance for both non-amyloidic polypeptide chains show no change with increasing temperature similar to the Q_{14} (chapter 4.4, Figure 4.4.4). The *Arrhenius* activation energy as well as the root mean square height of energy barriers ε for both investigated non-amyloidic model polypeptide chains are about 40 % smaller as compared to the value of the amyloidic Q_{14}. Interestingly, the natural DE-loop, which includes more sterically demanding amino acids, showed comparable values to the S_{14}, which includes a high number of hydrogen bond forming units. However, for the Q_{14} significantly high activation energy and corresponding root mean square height of energy barriers was found, which gives evidence for strong hydrogen bond networks and/or evidence for the steric effect resulting from the large glutamine side chains.
With the denaturant dependence the contribution of the hydrogen bonds to the found difference in the activation energies compared to the two non-amyloidic model polypeptide chains can be investigated, as discussed in chapter 4.5.

4.9 The Effect of GdmCl on the dynamics and dimensions of S_{14} and DE-loop

For a better understanding of a possible contribution of a hydrogen bonds network on the low intrachain diffusion coefficient and on the higher activation energy or the root mean square height of energy barriers of Q_{14}, a [GdmCl] dependent time-resolved FRET experiments for both S_{14} and DE-loop were performed.

4.9.1 trFRET measurements on S_{14} and DE-loop at different GdmCl concentrations

Time-resolved fluorescence decay curves of S_{14} and DE-loop in dependence of [GdmCl] were performed (Figure 4.9.1). For S_{14} the 1 M and 2 M [GdmCl] points could not be measured, because sufficient peptide concentration could not be reached. This can be explained by the observed salt effect on the solubility of the peptide

chains shown with potassium phosphate. With higher concentration of GdmCl, the salt effect seems to be dominated by the denaturation effect of GdmCl. A well-defined global analysis for both model polypeptide chains and each GdmCl concentration point were found as depicted in Figure 4.9.1 and 4.9.2.

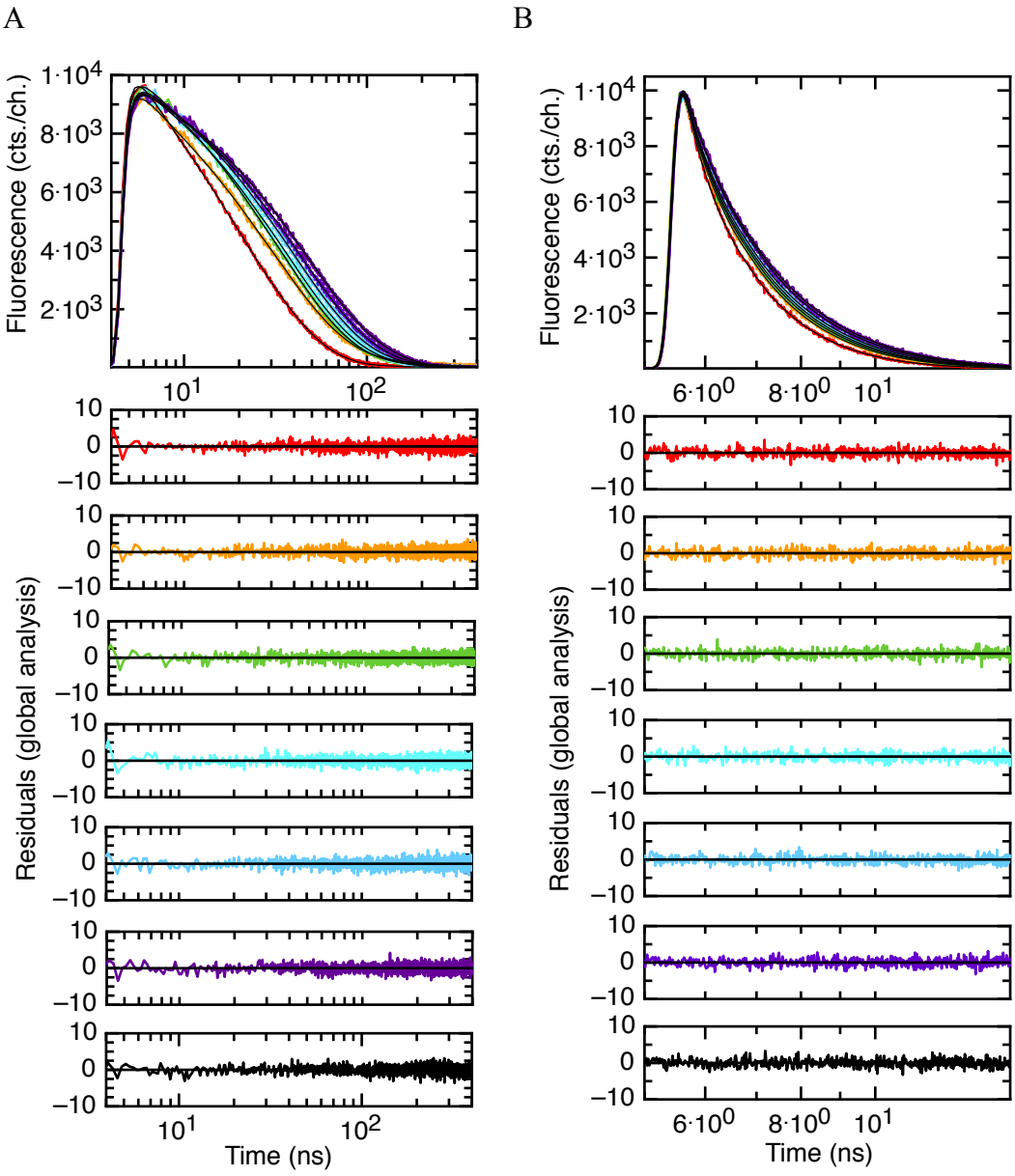


Figure 4.9.1: Time-resolved fluorescence decay curves of S_{14} in dependence of GdmCl using donor-acceptor pair **A** dansyl-*pyrene* and **B** DNB-AMCA pH 7.0 at different GdmCl concentration ranges from 0 M (red), 3 M (orange), 4 M (green), 5 M (turquoise), 6 M (blue), 7 M (dark purple) to 8 M (black). The residuals of the global analysis of each concentration point are shown in different colors corresponding to the different concentrations of GdmCl for both FRET-pairs. The S_{14} was dissolved in pure water without potassium phosphate and the pH was adjusted to 7.0 ± 0.3 with sodium hydroxide.

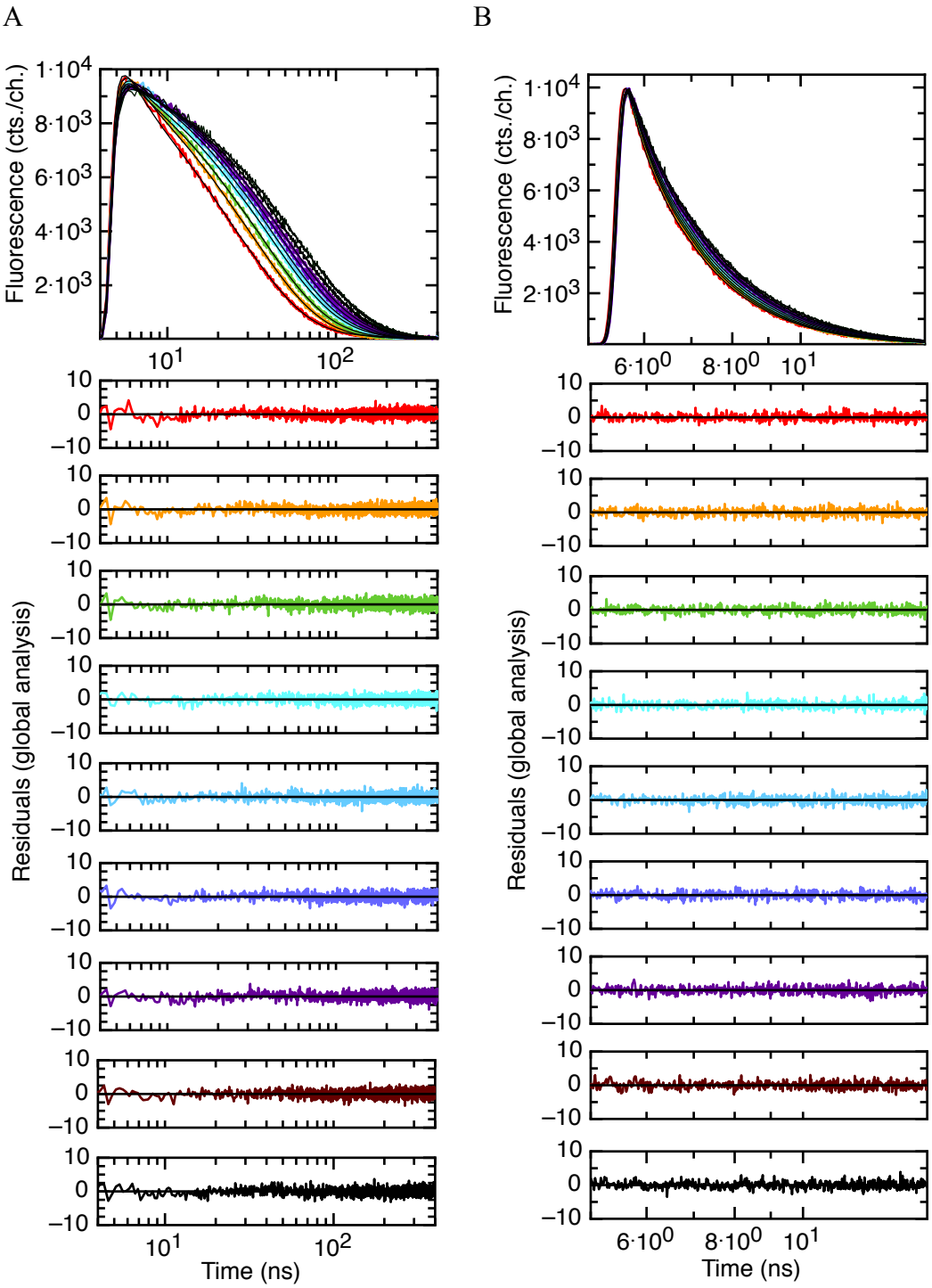


Figure 4.9.2: Time-resolved fluorescence decay curves of DE-loop in dependence of GdmCl using donor-acceptor pair **A** dansyl-pyrene and **B** DNB-AMCA in 10 mM potassium phosphate pH 7.0 at different GdmCl concentrations ranges from 0 M (red), 1 M (orange), 2 M (green), 3 M (turquoise), 4 M (blue), 5 M (purple), 6 M (dark purple), 7 M (brown) to 8 M (black). The residuals of the global analysis of each concentration point are shown in different colors corresponding to the different concentrations of GdmCl for both FRET-pairs.

The resulting intrachain diffusion coefficient and the viscosity-corrected intrachain diffusion coefficient for both model polypeptide chains are shown in Figure 4.9.3 A and B. For the DE-loop a slight increase of the viscosity-corrected intrachain diffusion coefficient with increasing [GdmCl] was observed. For S_{14} a significant increase of the viscosity-corrected intrachain diffusion coefficient with increasing [GdmCl] was found (A and B in Figure 4.9.3).

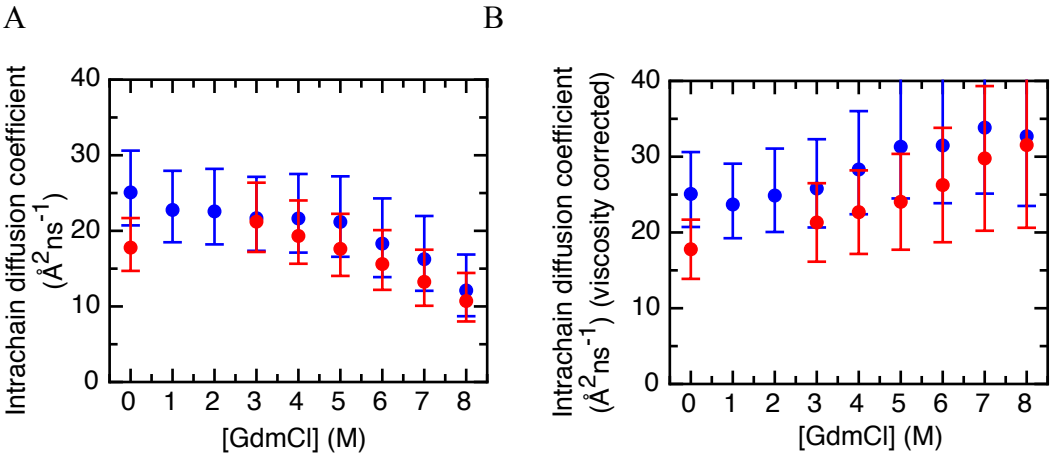
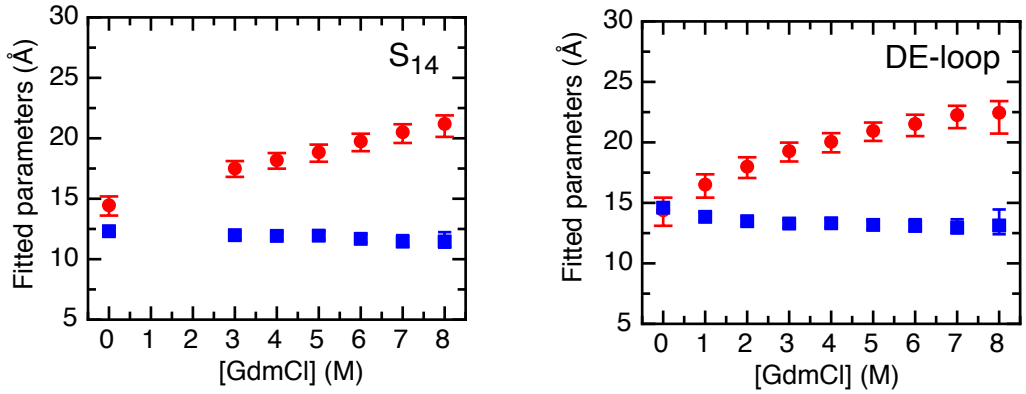


Figure 4.9.3: The intrachain diffusion coefficient of S_{14} (red circles) and DE-loop (blue squares) in dependence of GdmCl. **A** Intrachain diffusion coefficient is plotted against the [GdmCl] of both model polypeptide chains. **B** The viscosity-corrected intrachain diffusion coefficient is plotted against [GdmCl] and was calculated using the empirical equation 4.7 and the measured α-value of the corresponding model polypeptide chain. The GdmCl concentration points ranges from 0 M to 8 M for the DE-loop. For S_{14} 0 M, 3M to 8 M [GdmCl] are measured. The S_{14} was dissolved in pure water without potassium phosphate and the pH was adjusted to 7.0 ± 0.3 with sodium hydroxide.

For both model polypeptide chains a strong increase of the fitted parameter *b* was observed, while the fitted parameter *σ* decreased slightly with increasing [GdmCl] (A in Figure 4.9.4). This was also reflected in the calculated population of end-to-end distance, which sifted towards higher end-to-end distances (B in Figure 4.9.3).

A



B

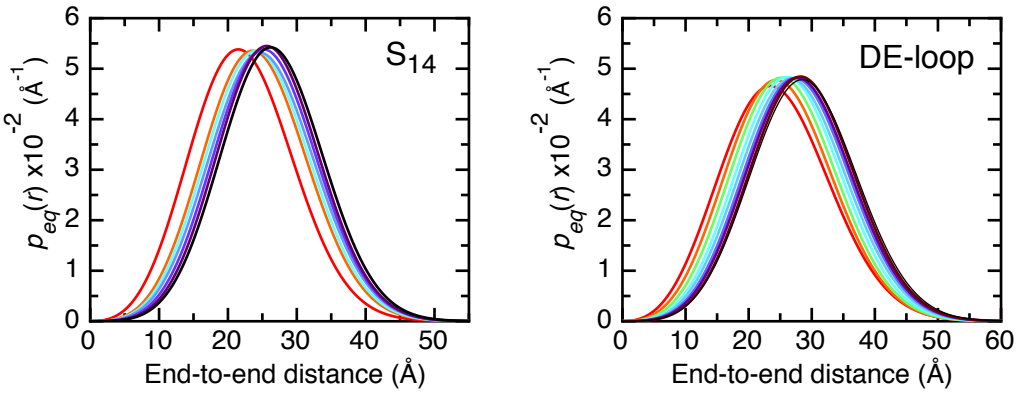


Figure 4.9.4: GdmCl dependence of S_{14} left side and DE-loop right side. **A** Fitted parameters *b* (red circles) and σ (blue squares) of both model polypeptide chains. **B** The end-to-end distance distribution against the end-to-end distance of both model polypeptide chains ranges from 1 M (red lines) to 8 M (black lines) [GdmCl]. The S_{14} was dissolved in pure water without potassium phosphate and the pH was adjusted to 7.0 ± 0.3 with sodium hydroxide.

As discussed in chapter 4.5 the dependence of root mean square end-to-end distance on the GdmCl can be described by the weak-binding model of Schellman (equation 4.16), where the exchange equilibrium constant K_{ex} for replacing water molecules with denaturant molecules can be determined. As shown in Figure 4.9.5, the weak-binding model described the data within the error well. For S_{14} a K_{ex} of 5.7 ± 2.7 and for the DE-loop a K_{ex} of 7.4 ± 4.3 was calculated (Figure 4.9.5).

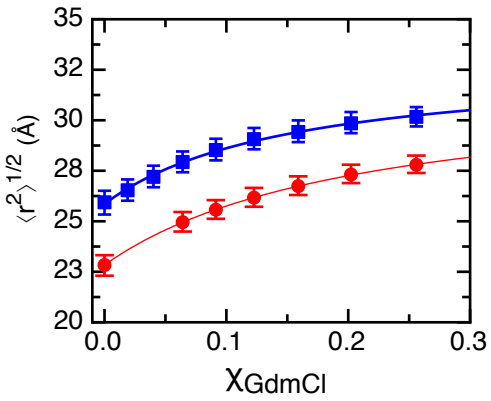


Figure 4.9.5: Root mean square end-to-end distance S_{14} (red circles) and DE-loop (blue squares) in dependence of GdmCl concentration ranges from 0 M to 8M. The asymptotic increase of the root-mean square end-to-end distance was fitted using the weak-binding model of Schellman (equation 4.16).

Nevertheless, the exchange equilibrium constant of both model polypeptide chains could not be determined very accurately, as was also observed for Q_{14} (chapter 4.5). For both model polypeptide chains a significantly smaller exchange constant K_{ex} was found in comparison to the value of 28.2 ± 6.2 for Q_{14}. Additionally, in previous work measured $(GS)_8$ and EF-loop yielded a slightly higher K_{ex} of 11.0 ± 4.3 and 10.9 ± 4.4, respectively (chapter 4.5), but are still in the same range as was found for the S_{14} and the DE-loop, taking into account the error.

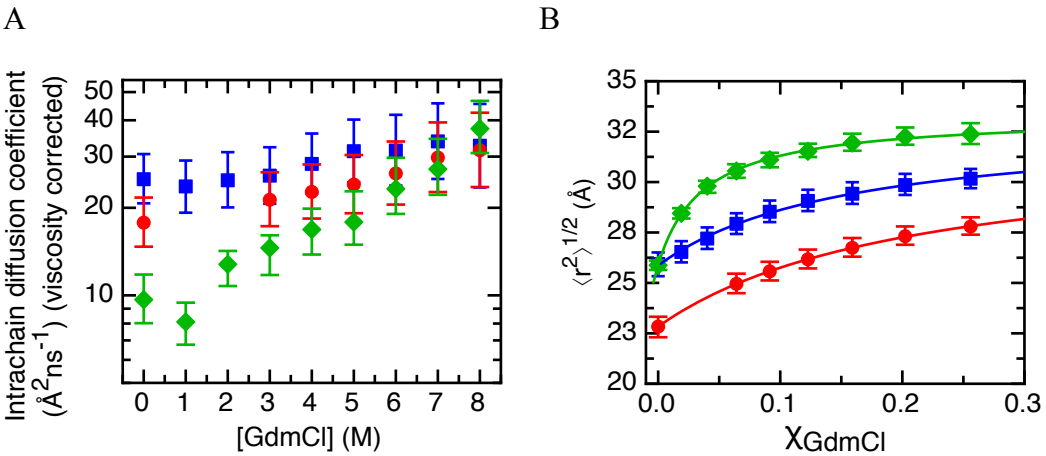


Figure 4.9.6: The effect of GdmCl on the chain dimension and dynamics. **A** The viscosity-corrected intrachain diffusion coefficient is semi-logarithmical illustrated against the GdmCl concentration. **B** The root mean square end-to-end distance is plotted against the mole fraction of GdmCl. Q_{14} is depicted in green squares, S_{14} is shown in red circles and the DE-loop is shown in blue squares. The asymptotic increase of the root-mean square end-to-end distance was fitted using the weak-binding model of Schellman (equation 4.16).

4.9.2 Summary and Discussion

In Figure 4.9.6 the effect of GdmCl on the chain dynamics and chain dimensions of both non-amyloidic chains S_{14} and DE-loop as well as the Q_{14} are shown. For Q_{14} it can be clearly observed that the chain dynamics, represented by the intrachain diffusion coefficient, is significantly influenced by the denaturant GdmCl, whereas for the non-amyloidic DE-loop only a slightly influence can be observed. However, the S_{14} showed a higher effect of denaturant on the chain dynamics, but to a significantly lesser extent as compared to the amyloidic Q_{14}. Thus, a steric effect does not seem to play an essential role, otherwise a corresponding effect would have been observed with the DE-loop (see discussion 4.8.2). The S_{14} consists of the same number of amino acids that can forms hydrogen bond in comparison to the Q_{14}, but includes only one hydrogen donor per side chain. For the Q_{14}, two hydrogen donors and one hydrogen acceptor per side chain are present. Therefore, Q_{14} seems to be able to form a network of hydrogen bonds, which is reflected in the strong sensitivity to GdmCl in comparison with the smaller effect exhibited by the S_{14}. Furthermore, for the root mean square dependence the strongest effect of GdmCl is found for Q_{14} in comparison the all investigated model polypeptide chains.

5. Summary and Outlook

In this work the intrachain dynamics and dimensions of the amyloidic polyglutamine chains in comparison with non-amyloidic model polypeptide chains are investigated with time-resolved FRET method, which provides direct access to these quantities. The Q_{14} was chosen as an amyloidic model chain to identify biophysical properties as a possible disease-related effect of polyglutamine chain at the monomeric level. The polyglutamine chain belongs to the intrinsically disordered polypeptides or proteins (IDP), which are flexible and highly dynamic. This consequently leads to an ensemble of different conformational states that represents a dimensional parameter of the polypeptide chains and must be described by dedicated polymer models. The chain dynamics represent another important parameter to describe flexible and unstructured polypeptide chains and are quantified by the intrachain diffusion coefficient. In this work, the intrachain diffusion coefficient describes the one-dimensional diffusion of one chain-end relative to the other chain-end during the lifetime of the excited FRET-donor. This factor of the polyglutamine chains was compared to that of non-amyloidic polypeptide chains with similar numbers of amino acid, specifically the artificial sequence polyserine chains and the DE-loop as fragment of the IDP β-carp parvalbumin.

The effect of viscogenic co-solute glycerol, temperature as well as the denaturant GdmCl on the chain dynamics and dimension was measured for the amyloidic as well as non-amyloidic model polypeptide chains. An rather slow chain dynamics for amyloidic Q_{14} in comparison to highly flexible poly-(glycine-serine) chains and lesser flexible model polypeptide chains such as the polyserine chains and the natural two fragments of the IDP β-carp parvalbumin was found (Table 5, intrachain diffusion coefficient). Compared to the other model polypeptide chains, no compact structure of Q_{14} can be observed (root mean square end-to-end distance Table 5). Therefore, the slow chain dynamics cannot be associated with a compact nature of Q_{14}. Coarse-grained MC simulations of Q_{14} supports the resulting trFRET dimension and found a distance distribution, which can be described with the here used EDWARDS' polymer model, and yields a comparable root mean square end-to end distance of 24.9 Å [292].

Table 5: Comparison of the amyloidic Q_{14} with the non-amyloidic model polypeptide chains in aqueous solution. Both $(GS)_8$ and the EF-loop are taken from previous work [159] and is depicted in grey. All model polypeptide chains were measured in 10 mM potassium phosphate at 7.0 pH and 22.5 °C. The S_{14} was dissolved in pure water without potassium phosphate and the pH was adjusted to pH 7.0 ± 0.3 with sodium hydroxide.

peptide	*number of amino acid*	*intrachain diffusion coefficient (Å²/ns)*	*root mean square end-to-end distance (Å)*	*Arrhenius activation energy (kJ/mol)*
Q_{14}	14	9.7± 2.1	25.9 ± 0.3	16.6 ± 5.2
S_{14}	14	17.9± 3.4	22.8 ± 0.5	9.5 ± 3.4
DE-loop	14	22.3± 4.5	25.9 ± 0.6	9.2 ± 3.9
EF-loop	16	22.2 ± 3.8	26.5 ± 0.4	12.7 ± 2.8
$(GS)_8$	16	33.9 ± 5.8	23.5 ± 0.4	9.8± 2.8

Also, the conformation distribution of Q_{14} showed no large deviation related to the other investigated model polypeptide chains (chapter 4.2). Further investigation of the temperature dependence of the dynamics of polyglutamine chains yielded significantly higher *Arrhenius* activation energy compared to the non-amyloidic polypeptide chains. This activation energy describes the conformational energy barrier of unstructured polypeptide chains (Table 5. *Arrhenius* activation energy) and may explain the observed slow chain dynamics of the polyglutamine chains in water.
Time-resolved FRET experiments at variable solvent viscosity showed another deviation of Q_{14}. Such investigations were carried out with the viscogenic small-molecular co-solute glycerol, in which the chain dynamics decreased with increasing viscosity, whereby a viscosity-sensitive exponent α of 1 is expected to fulfill Kramers' theory. However, for polyglutamine a α-value smaller than 1 was found. Additionally, two distinguishable regimes for the glycerol dependence for polyglutamine were found, which was up to now not observed for another non-amylodic polypeptide chains (Figure 5.2). These findings could be correlated to transient sub-structures, which shows less sensitivity to the viscosity and follow non-Kramer-like dependence of the chain dynamics.

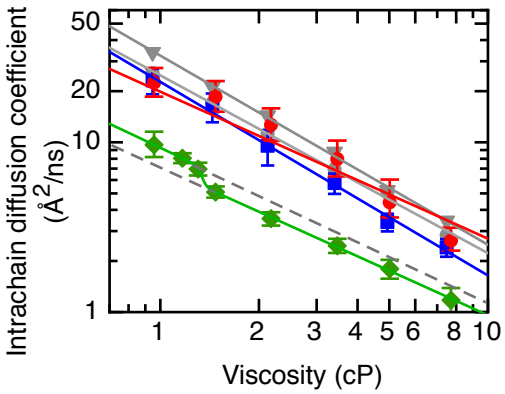


Figure 5.2: Effect of the viscogenic co-solutes on the chain dynamics and dimensions. The double-logarithmical illustration of the intrachain diffusion coefficient against the viscosity of glycerol ranges from 0% to 56% (w/w). The Q_{14} is shown in green diamonds, S_{14} is shown in red circles and the DE-loop is shown in blue squares. The $(GS)_8$ (triangles) and the EF-loop (circles) were measured in previous work and were depicted in grey [159]. The viscosity-sensitive exponent α was fitted using equation 4.7.

Besides transient sub-structures, formed *e.g.* by hydrogen bonds, the steric effects within the polyglutamine chains resulting from the large side chain of the glutamine may also limit the chain dynamics. In addition to other effects (see discussion chapter 4.5), the denaturant GdmCl is particularly interesting because it interrupts hydrogen bonds *e.g.* intramolecular hydrogen bonds formed in polypeptide chains [133], [280]. Therefore, GdmCl dependent time-resolved FRET experiments could elucidate the influence of hydrogen bonds on the chain dynamics. For Q_{14} a strong GdmCl effect on the chain dynamics was observed, whereas for the non-amyloidic model polyserine chain only a small effect of GdmCl on the chain dynamics could be observed (Figure 4.3 A)). Noted that the polyserine chain includes similar numbers of hydrogen bond forming amino acids compared to the polyglutamine chain. Additionally, a strong effect of GdmCl on the chain dimension of polyglutamine chains compared to the non-amylodidc model polypeptide chains was found (B in Figure 5.3). This result gives first evidence that the intramolecular hydrogen bonds are formed between side chains as well as side chain-main chain and explains the high impact on the chain dynamics of the polyglutamine chains and also the low impact of steric effects.

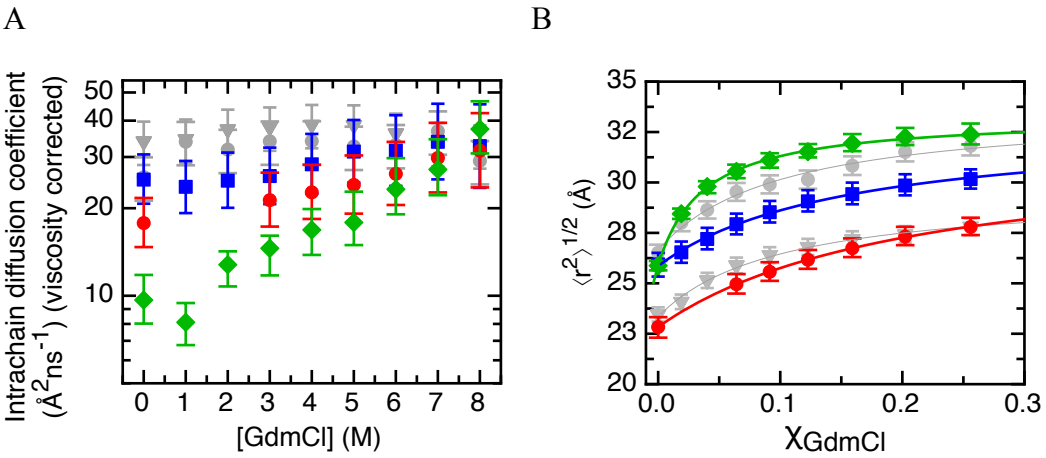


Figure 5.3: The effect of GdmCl on the chain dynamics and dimensions. **A** The semi-logarithmic viscosity-corrected intrachain diffusion coefficient is illustrated against the GdmCl concentration. **B** The root-mean square end-to-end distance is plotted against the mole fraction of GdmCl. The Q_{14} is depicted in green squares, S_{14} is shown in red circles and the DE-loop is shown in blue squares. The $(GS)_8$ (triangles) and the EF-loop (circles) were measured in previous work and are shown in grey [159]. The apparent asymptotic increase of the root-mean square end-to-end distance was fitted using the weak-binding model of Schellman (equation 4.16).

Punihaole *et. al.* found in MD simulations of Q_{10} that hydrogen bonds of both side chain-side chain and side chain-peptide backbone are enthalpically more favorable than those of the glutamine side chain to water [293]. Furthermore, using replica exchange molecular dynamics (REMD) simulations of Q_{10}, it was demonstrated that the multiple configurations of the polyglutamine monomers are due to the formation of side-chain hydrogen bonds [294]. MD simulations and coarse-gained MD sampling of a similar construct, as used in this work, revealed a specific hydrogen bonding pattern characterizing transient structure motifs. Such motifs are found in low amounts and are predicted to serve as first state for nucleation and further aggregation into amyloidic fibrils. Interestingly, an influence of the solubility tail of the peptide on the hydrogen-bond pattern was also discussed [295]. Wang *et. al.* investigated different conformational states of polyglutamine in MD simulations and found 4 to 5 hydrogen bonds between side chains and peptide backbone for Q_{15} [296]. However, glutamine incorporated in helices stabilized those through intramolecular hydrogen bonds of the side chains as shown for different model polypeptide chains and the Ataxin-3 protein (Introduction, chapter 2) [297]–[299].

Moreover, in the amyloidic fibrils the hydrogen bonds between the side chains play an important role. X-ray scattering and computer-generated model as well as solid-state NMR show that the single chains are stabilized by intra- and intermolecular hydrogen bonds formed between side chain-side chains and side chain-peptide backbone [300]–[303]. However, it seems that the hydrogen bonds play an important role not only in the amyloidic fibrils but also for the monomeric state of the polyglutamine chains. In addition, a network of hydrogen bonds between the glutamine side chains may exist, which alters the chain dynamics in the monomeric state as shown in this work.

Nevertheless, in this work it was shown that the polyglutamine chains, represented by the Q_{14}, show significantly different behavior at the monomeric level in comparison to non-amyloidic model polypeptide chains. Of course, more detailed investigations are necessary to validate the first hypothesis of influence of hydrogen bonds on the chain dynamics of polyglutamine chains and to rule out other effects such as the hydrophobic effect or electrostatic effects like dipole-dipole interactions. In future steps, model polypeptide chains can be designed, where one hydrogen donor or acceptor of the glutamine side chain is replaced by a methyl-group, to avoid the formation of hydrogen bonds. Also the distance between the glutamines in the polypeptide chain can be increased by the incorporation of glycine or alanine, which should interrupt the hypothesized hydrogen bonding network and results in a decreased influence of GdmCl on the chain dynamics and in a lower *Arrhenius* activation energy.

Rose *et. al.* discussed the “uniquely pivotal” role of hydrogen bonds in favoring the folded and the unfolded state in poor or good solvent and clarified the importance of hydrogen bonds for folding at both states, the unfolded and the folded state [8], [304].

However, understanding the influence of hydrogen bonds on the chain dynamics at the monomeric level of polyglutamine chains seems to be an important step not only for the understanding of the dynamics and structural behavior at the monomeric level and the formation of fibrils, but also to obtain more information on the molecular mechanism of aggregations inhibitors, such as the QBP1 discussed in the introduction, for *e.g.* future therapeutic drug developments.

This work provides for the first time detailed information on both the intrachain dynamics and dimension of an amyloidic polypeptide chains under various biophysical conditions. In general, the combination of the presented results and

computational approaches could enable a deeper understanding of the conformational space and furthermore structural information about amyloid-prone structures. This is also valid for IDPs, because their functionality are probably also based on their conformational space. Therefore, it is necessary to understand their properties both in the unfolded state and in the state bound to their target protein. In addition, it could be shown that the trFRET method is a suitable technique not only for a better understanding of IDPs but also for the future elucidation of fundamental questions such as protein folding.

6. Material and Methods

All used chemicals were purchased from Sigma Aldrich (St. Louis, MO, USA) or Carl Roth (Karlsruhe, Germany) unless mentioned otherwise in the following text. For the solid-phase peptide synthesis (SPPS) the Fmoc-protected amino acids were purchased from either Iris Biotech GmbH (Marktredwitz, Germany), Merck KGaA (Darmsadt, Germany) or Novarbiochem® (Merck group). All used Resins for the SPPS were purchased from Rapp Polymere GmbH (Tübingen, Germany).

6.1 Solid-phase peptide synthesis and purification

6.1.1 Solid-phase peptide synthesis (SPPS)

All used polypeptide chains were synthesized on ABI 433 with UV monitoring (Applied Biosystems, CA, USA) or Liberty Blue™ CEM Corporation. For all synthesis the Resin Tentagel R RAM from Rapp Polymer were used. For the synthesis on ABI 433 the Fmoc (9-fluoroenylmethyloxycarbonyl) deprotection was performed using 20% piperidine and the couple activation was performed using HBTU (2-(1H-benzotriazol-1-yl)-1,1,3,3-tetramethyluronium hexafluorophosphate) or HATU (O-(7-Azabenzotriazol-1-yl)-N,N,N′,N′-tetramethyluronium hexafluorphosphate). The standard solvent was NMP (N-Methyl-2-pyrrolidone). For the Liberty Blue™ the Fmoc deprotection was performed with a deprotection cocktail consisting of 10% piperazin (w/v) in DMF using 0.1 M HOBT (1-Hydroxybenzotriazole hydrate) and 1:9 ethanol:NMP, and the activation was performed using DIC (N,N'-Diisopropylcarbodiimide) and the additive Oxyma pure (ethyl 2-cyano-2-(hydroxyimino)acetate). The standard solvent was DMF (N,N'-dimethyhlformamide). The side chain of amino acids were protected to avoid side reactions as shown in the following sequences (from N- to C-termini):

$(GS)_8$:

Gly, Ser(tBu), Gly, Ser(tBu), **Gly-Ser(psi(Me,Me)pro)**, Gly, Ser(tBu), Gly, Ser(tBu), Gly, Ser(tBu), **Gly-Ser(psi(Me,Me)pro)**, Gly, Ser(tBu), **Dpr(Mtt)**, Ser(tBu), Arg(Pbf), Gly.

EF-loop chains from β-carp parvalbumin:

Leu, Lys(Boc), Ala, Gly, Asp(OtBu), Ser(tBu), **Asp(OtBu)-(Dmb)Gly**, **Asp(OtBu)-(Dmb)Gly**, Lys(Boc), Ile, Gly, Val, Asp(OtBu), Glu(OtBu), **Dpr(Mtt)**, Ser(tBu), Gly.

DE-loop chains from β-carp parvalbumin:

Lys(Boc), Ala, Asp(OtBu), Ala, Arg(Pbf), **Ala, Leu-Thr(psi(Me,Me)pro)**, **Asp(OtBu)-(Dmb)Gly,** Glu(OtBu), Thr(tBu), Lys(Boc)-Thr(psi Me, Me pro), **Dpr(Mtt)**, Ser(tBu), Gly.

S_{14}:

Ser(tBu)-Ser(psi(Me,Me)pro), Ser(tBu)-Ser(psi(Me,Me)pro), Ser(tBu), Ser(tBu)-Ser(psi(Me,Me)pro), Ser(tBu)-Ser(psi(Me,Me)pro), Ser(tBu)-Ser(psi(Me,Me)pro), Ser(tBu), Ser(tBu)-Ser(psi(Me,Me)pro), Dpr(Mtt), Ser(tBu), Arg(Pbf), Ser(tBu), Arg(Pbf), Gly.

Q_{14}:

Gln(Trt), Gln(Trt), Gln(Trt), Gln(Trt), Gln(Dmcp), Gln(Trt), Gln(Trt), Gln(Trt), Gln(Trt), Gln(Dmcp), Gln(Trt), Gln(Trt), Gln(Trt), Gln(Trt), Dpr(Mtt), Ser(tBu), Arg(Pbf), Ser(tBu), Arg(Pbf), Gly.

Q_3:

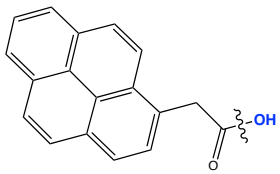
Gln(Trt), Gln(Trt), Gln(Trt), Dpr(Mtt), Ser(tBu), Arg(Pbf), Ser(tBu), Arg(Pbf), Gly.

Double building blocks were used for optimizing synthesis *e.g.* pseudo-prolines Gly-Ser(psi(Me,Me)pro). Double coupling steps for the API 433 synthesis are marked in bold at the sequences above. For peptide synthesis with Liberty Blue™ only Dpr(Mtt) was doubled coupled.

6.1.2 Labeling of chromophores

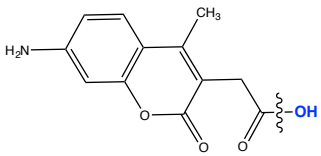
For trFRET measurements the polypeptide chains dansyl-sequence-*pyrene* and DNB-sequence-AMCA were synthesized as described above, where Dpr is 2,3-diaminopropionic acid and *dansyl*, *pyrene*, *1-fluoro-2,4-dinitrobenzene* (DNB*)* and *7-amino-4-methyl-3coumarinylacetic acid* (AMCA*)* are the chromophores as specified in Figure M.1. The protecting group Mtt (4-methyltrityl) of Dpr was selectively removed using 2% TFA (trifluoroacetic acid) and 5% TIPS (triisopropylsilane) in DCM (dichloromethane). The FRET donors AMCA and *pyrene* (1-pyreneacetic acid, both from Sigma Aldrich, St. Louis, MO, USA) were coupled to the free side chain amine of Dpr (Iris Biotech, Marktredwitz, Germany) using HATU in DMF. The FRET acceptors *dansyl* (5-(dimethylamino)naphthalene-1-sulfonyl chloride, Sigma Aldrich, St. Louis, MO, USA**)** and DNB (1-fluoro-2,4-dinitrobenzene, Sigma Aldrich, St. Louis, MO, USA) were coupled to the N-terminus of the polypeptide chain using N,N-diisopropylethylamine in DMF. In donor-only or acceptor-only polypeptide chains the second reactive sites (N-terminus or amine of the side chain of Dpr were acetylated using acetic acid and DIPEA (N-diisopropylethylamine) in DMF.

A



1-Pyreneacetic acid (*pyrene*)

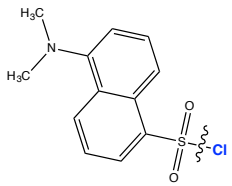
$R_0 \approx 22$ Å; $\tau_D = 150$ ns



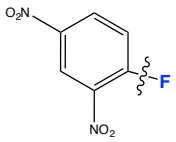
7-Amino-4-methyl-3-coumarinylacetic acid (AMCA)

$R_0 \approx 31$ Å; $\tau_D = 5$ ns

B



5-Dimethylaminonaphthalene -1-sulfonyl chloride (*dansyl*)



1-Fluoro-2,4-dinitrobenzene (DNB)

Figure M.1: Time-resolved FRET chromophores with the corresponding R_0 and donor-only lifetime τ_D. The wavy line depicts the coupling position of the chromophore to the polypeptide chain, while the blue functional group is removed for **A** donor and **B** acceptor.

6.1.3 Purification

All polypeptide chains were cleaved from the Resin by cleavage solution of 90% TFA, 5% TIPS and 5% pure water under strong mixing for 3 hours. The polypeptide chain was precipitated in cold tert-butyl methyl ether (1:5, cleavage solution:tert-butyl methyl ether) and filtered. After drying, the raw polypeptide chains were purified using reversed-phase high-performance liquid chromatography (HPLC) with different RP-columns types and different gradients of ACN (Acetonitrile) and water mixture as described below. For injection in the HPLC-instrument the raw polypeptide chains were dissolved in 5% ACN in pure water. For Q_{14} the raw peptide was dissolved in pure TFA and incubated for 1 hour. The HPLC-solvents included 0.1% TFA (v/v) for both pure water and ACN. For Q_{14} and S_{14} the TFA concentration was changed to 0.01% (v/v).

$(GS)_8$ and EF-loop from β-carp parvalbumin:

1. RP-column, preparative: C8 (Merck, LiChrospher® 100 RP-8 (5 µm) 250-25, gradient: 0.5% ACN/min, started at 10% CAN, Flow rate 20 ml/min
2. RP-column, semi-preparative: C12 (Phenomenex Jupiter 4u Proteo 90A, 250x10 mm), gradient: 0.5% ACN/min, started at 10% CAN, Flow rate 4.7 ml/min

DE-loop from β-carp parvalbumin:

1. RP-column, preparative: C12 (Phenomenex Jupiter 10u Proteo 90A, 250x30 mm), gradient: 0.66% ACN/min, started at 15% ACN, Flow rate 20 ml/min
2. RP-column, semi-preparative: C12 (Phenomenex Jupiter 4u Proteo 90A, 250x10 mm), gradient: 2% ACN/min, started at 15% ACN, Flow rate 4.7 ml/min

S_{14}:

1. RP-column, preparative: C12 (Phenomenex Jupiter 10u Proteo 90A, 250x30 mm)), gradient: 0.3% ACN/min, started at 5% ACN, Flow rate 20 ml/min
2. RP-column, semi-preparative: C18 (Hibar LiChrosorb® RP-18 (5 µm), 250-10), gradient: 0.3% ACN/min, started at 5% ACN, Flow rate 5 ml/min

$Q_{14/3}$:

1. RP-column, preparative: C12 (Phenomenex Jupiter 10u Proteo 90A, 250x30 mm)), gradient: 0.5% ACN/min, started at 10% ACN, Flow rate 20 ml/min
2 RP-column, semi-preparative: C12 (Phenomenex Jupiter 4u Proteo 90A, 250x10 mm), gradient: 0.3% ACN/min, started at 10% ACN, Flow rate 5 ml/min

6.1.4 Determination of the polypeptide chain purity

All HPLC fractions were analyzed with LC-MS (liquid chromatography mass spectrometry, Waters ACQUITY UPLC System (H-CLASS) with PDA eλ Detector and QDa Detector (Quadrupole)). Fraction with a peak purity of ≥ 98% for donor-acceptor polypeptide chains and ≥ 95% for donor- or acceptor-only polypeptide chains were pooled and lyophilized (-50 °C chamber temperature and -80 °C ice condensator temperature, Christ Alpha 2-4 LD plus). Analysis software: Waters MassLynx Software and MagTan-Mag1.

6.2 Sample preparation

For all measurements, the lyophilized peptides were dissolved in the measuring buffer and ultrasonication was done for 10 min. Insoluble particles were removed by two times centrifugation at 13300 rpm for 8 min (Thermo Electron Corporation Heraues Pico 17 Centrifuge). For time-resolved fluorescence measurements the concentration were adjusted for optical density of 0.1 at the excitation wavelength of the donor. Measurements for S_{14} were performed in pure water without any buffer. Therefore, the pH was adjusted with sodium hydroxide (1 mM) to a pH of 7.0 ± 0.3. For Q_{14} a disassembly protocol was used modified from Wetzel and colleagues [183], [246] (see 6.3). As the measurements at low temperature were conducted with a constant stream of nitrogen, because additional influences on the lifetime could be observed due to degassing of the solvents. Therefore, for temperature-dependent experiments, the measuring solution should be oxygen-saturated as shown in previous work [159].

After 6 min incubation at a given temperature, the measuring solution was 10 times bubbled with a standard pipette and further equilibrated for 4 min to reach a thermal equilibrium. For the temperature dependent experiments, round lid fluorescence cuvette was used.

6.3 Disassembly protocol for Q_{14}

In the current work a modified disassembly protocol was used, which was earlier established by Wetzel and colleagues to generate monomeric state of polyglutamine chains [183], [246]. Lyophilized peptides were incubated in 1:1 mixture of TFA:HFIP (trifluoroacetic acid:1,1,1,3,3,3 hexaflouro-2-propanol) over night (>5 hours). Afterwards, the solution was evaporated under mild nitrogen stream and lyophilized for at least 1 hour (-50 °C chamber temperature, -80 °C ice condensator temperature). The lyophilized powder was dissolved in the measuring buffer and centrifuged with ~267000 g (Beckman Coulter® Optima™ TLX Ultracentrifuge using "Festwinkelrotor" TLA-110, middle position of tube) at 20.0 °C for 1.5 hours or more, depending on the viscosity of the measuring buffer. The upper 75% of the supernatant was used.

6.4 Buffer preparation

All measuring buffers were prepared with ultra pure bottled water (Water HPLC Plus Sigma-Aldrich) to prevent contamination and filtered with Anatop 0.2 µm Whatman filters. Before filtration, the pH of 7.0 was adjusted (pH meter Schrott Instrument with SenTix Mic pH electrode, with three point calibration every week) using diluted potassium hydroxide or phosphoric acid solution. If the buffer system allowed freezing, the buffer was aliquoted, frozen with liquid nitrogen and storage at -80 °C. All buffer included 10 mM potassium phosphate (Merck) except for S_{14}, which was dissolved in pure water (0.055 µS/cm, ions exchange Purelab Plus UF). Glycerol was purchased from Roth (Glycerol, ≥99.7%, Ultra-Quality, synthetic Roth) and balanced

with an accuracy of ± 0.01 g for preparation of the different glycerol concentration (w/w). GdmCl was purchased in high quality (AA-grade) from NIGU Chemie GmbH Waldkraiberg, Germany. The resulting GdmCl concentrations were checked and finally adjusted with an automatic refractometer (AR700, Reichert Inc. Depew,NY, USA). The PEG_{6000} was purchased from Merck (zur Synthese). A stock solution of 40.8% (w/v) was incubated with ultra pure activate charcoal (Norit, Sigma Aldrich), centrifuged and the supernatant was filtrated through syringe filters (Whatman, 0.2 μm). The resulting viscosity was measured with Anton Paar DMA 4100M/Lovis 2000 ME.

6.5 Fluorescence measurements

6.5.1 Fluorescence lifetime measurements

For all fluorescence lifetime measurements the fluorescence lifetime spectrometer FluoTime 200 from PicoQuant (Berlin, Germany; with a PDL 800-Dpicosecond pulsed diode laser driver unit; PicoHarp TCSPC; PMA-182 photo multiplier tubes (H5783 photosensor modules, Hamamatsu); polarizers and cuvette holders (TLC 50, Quantum northwest, WA, USA)) were used to determine the donor-only lifetime τ_D and to obtain the fluorescence decay curves of the donor-acceptor labeled polypeptide chains for the global analysis. The peptide concentrations were adjusted to a maximum optical density of 0.1 at the excitation wavelength of the donor (Table 6). The measurements were stopped at a maximal peak counts/channel value of 10000. The corresponding buffer without peptides were measured with the same set up and amount of time, which was determined with the labeled polypeptide chain. The instrument response function (IRF) was determined at the excitation wavelength identical to that of excitation of the donor and stopped at a maximal peak counts/channel value of 40000. The IRF solution was prepared with 1:100 dilutions in pure water of the Ludox solution (Sigma Aldrich) to generated the scattering effect and was measured for each measurement day. For determination of the donor-only

lifetime the company software of PicoQuant FluoFit (Berlin, Germany) based upon a reconvolution fit using a sum of exponentials were employed:

$$F(t) = \int_{-\infty}^{t} IRF(t') \sum_{i=1}^{n} A_i \exp\left(-\frac{t-t'}{\tau_i}\right) dt' \qquad \text{(Eq. 6.1)}$$

Table 6: Set up for fluorescence decay curves for donor-only and donor-acceptor labeled polypeptide chains of the fluorescence lifetime spectrometer FluoTime 200 from PicoQuant.

set up	*pyrene*	*dansyl/pyrene*	*AMCA*	*DNB/AMCA*
light source	LED	LED	Diode laser	doidelaser
puls width	500 ps	500 ps	44 ps	44 ps
λ_{ex}	298 nm	298 nm	372 nm	372 nm
$\Delta\lambda_{ex}$	14 nm	14 nm	1.1 nm	1.1 nm
λ_{em}	398 nm	398 nm	445 nm	445 nm
$\Delta\lambda_{em}$	16 nm	16 nm	4 nm	4 nm
time-resolution	256 ps	256 ps	16 ps	16 ps
filters	LP 345	LP 345	LP 398	LP 398
slits	2.0	2.0	0.5	0.5
Rep.-frequency	I 2/2	I 2/1	I 1/4	I 1/4

6.5.2 Determination of the quantum yield and Förster distance

As shown in equation 2.4, the Förster distance depends strongly on the quantum yield Q and the spectral overlap of the donor fluorescence and the acceptor extinction. For the determination of the quantum yield Q, the absolute intensity of the fluorescence spectra has to be measured. Therefore, fluorescence standards with known quantum yields were used (Table 8).

$$Q = Q_{st} \cdot \frac{F A_{st} n^2}{F_{st} A n_{st}^2} \qquad \text{(Eq. 6.2)}$$

Where Q_{st} represents the quantum yield of standard fluorophore, F represents the integrated fluorescence intensities, where F_{st} stands for the standards. The excitation wavelength is given as A and the refractive index of the solvent is given as n. The accuracy was improved through measurements of three different fluorophore

standards for the determination of the quantum yield of donors, *pyrene* and AMCA (Table 8). The concentrations of the donor-only labeled peptides were adjusted to have an optical density ranging from 0.2 to 0.4 (double beam spectrophotometer, AVIV model 14 UV-VIS, AVIV Biomedicals, NJ, USA) at the maximum wavelength according to the published extinction coefficient (Table 7). For fluorescence measurements with the Fluorlog Tau 2 spectrofluorometer (Jobin Yvon, Horiba, USA with a correction file for intensity dependence on the wavelength) the sample were diluted to obtain optical density of 0.01. For a high accuracy of the resulting concentration, the dilution was balance to calculate the real dilution factor (Satorius AG Göttingen ME235P-OCE, ± 0.01 mg). The extinction spectra $\varepsilon_A(\lambda)$ were determined from the absorbance spectra of the corresponding acceptor-only labeled polypeptide chain with the published extinction coefficient ε_A (ε_A^{max} = 4300 M^{-1} cm^{-1} for *dansyl* [305] and for ε_A^{max} = 15900 M^{-1} cm^{-1} *DNB* [306]). The overlap integral J for the fluorescence spectra F_D and the extinction spectra $\varepsilon_A(\lambda)$ was calculated as follows:

$$J = \frac{\int_0^\infty F_D(\lambda)\varepsilon_A(\lambda)\lambda^4 d\lambda}{\int_0^\infty F_D(\lambda) d\lambda} \qquad \text{(Eq. 6.3)}$$

For the calculation of the Förster distance R_0, the determined quantum yield Q, the overlap integral J and the commonly used orientation factor, κ^2 of 2/3 (see Introduction) were used. The relative errors of the Förster distances were estimated to be about 2.5%, as established in previous work [159]. For all the peptides in the current work the R_0 were measured in buffer of 10 mM potassium phosphate pH 7.0 at 22.5 °C. It should be noted that the dependence of R_0 for Q_{14} on GdmCl concentration and for $(GS)_8$ on PEG_{6000} concentration were determined, confirming previous work in that R_0 is independent of the nature of polypeptide chains under different additive conditions. Therefore, R_0 for all other peptides were theoretically estimated from previously measured R_0 in the 10 mM potassium phosphate buffer at pH 7.0 and 22.5 °C.

Table 7: Set up for the determination of the quantum yield *Q* and the overlap integral *J* from the fluorescence spectra of donor labeled polypeptide chains and the absorbance spectra of the acceptor-only labeled polypeptide chains

set up	*fluorescence spectra (for Q)*		*absorbance spectra (for J)*	
	pyrene	*AMCA*	*dansyl*	*DNB*
λ_{ex}	313 nm	353 nm	420 - 260 nm	440 – 280 nm
λ_{em}	320 – 600 nm	360 – 700 nm	-	-
bandwidth λ_{ex}	2 nm	2 nm	2 nm	2 nm
bandwidth λ_{em}	1 nm	1 nm	-	-
integrations time	0.1 s	0.1 s	0.5 s	0.5 s
wavelength step	1 nm	1 nm	1 nm	1 nm

Table 8: Standard fluorophores for determination of the quantum yield.

fluorescence standard	*corresponding donor*	*quantum yield*	*solvent*
9-anthracenecarbonitrile	AMCA	0.80 [307]	methanol
POPOP (1,4 bis(5-phenyloxazol-yl)benzene)	*pyrene*, AMCA	0.97 [307]	cyclohexane
DPA (9,10-diphenylanthracene)	*pyrene*, AMCA	1.00 [307]	cyclohexane
PPO (2,5-diphenyloxazole)	*pyrene*	0.94 [307]	cyclohexane
perylene	AMCA	0.94 [308]	cyclohexane

6. 6. Global analysis

The applied global analysis routine was programmed using Matlab (MathWorks, MATLAB R2015b), which was developed and described in the previous work by Ursula Zinth [159]. The program was run on a Mac Pro computer. On this basis, the calculation speed was improved to around 1000 times with the help of Jan Ebenhan, where the Matlab function "pdepe" [159], [309] was replaced as described in Chapter 6.6.2, first section. This improvement allows the calculation of the *SPA* for each buffer point and each polypeptide chain. Also the calculation of the root mean square end-to-end distance error by the *SPA* was enables with the new routine as described in Chapter 6.6.3. Additional the second donor lifetime was included as a fit parameter for the global analysis of the Q_{14} fluorescence decay curves

(see next Chapter 6.6.2 and equation 4.6). An additionally output of the amplitudes of all fit parameter were also included as well as a determination of the time range (tear off) of the fluorescence decay curves for testing dependencies. Both the global analysis routine from Ursula Zinth [159] as well as the new routine described in the next Chapters, obtain equal results of D_{DA}, b and σ considering the error.
All improvements and modifications can be used without restriction for further calculations of other polypeptide chains with different fit parameters, which is enabled by the inactivation of the secondary donor lifetime.
The next Chapters 6.6.1, 6.6.2 (second section) and 6.6.3 are based on the descriptions in the dissertation of Ursula Zinth, who developed and described the global analysis routine version, which serves as the initial routine for my work [159].

6.6.1 Data input for the global analysis

The following inputs are necessary for the global analysis: at first, the measured time-resolved fluorescence decay curve of the donor-acceptor labeled polypeptide chains is needed. Second, the time-resolved fluorescence decay curves of the corresponding buffers and IRFs are needed for the calculation.
Both the Förster distance and the donor lifetime at each corresponding buffer point or at the respective corresponding temperature point of the two chromophores *pyrene* and AMCA, as well as the total length of the polypeptide chains r_{max} (number of amino acids plus two times 3.8Å) and r_{min} of 0, are the fixed parameters in this analysis. To achieve a successful calculation, the initial values for b, σ, and the intrachain diffusion coefficient D_{DA} must be included.

6.6.2 Numerical solution of the partial differential equation (PDE)

Section 1 (in discussion with Jan Ebenhan):
The Matlab function "pdepe", which is used by the global analysis routine of Ursula Zinth [159], is no longer used in the new routine. In the new routine the spatial part of the differential equation is discretized at distance r and the derivatives are approximated with 4th order central difference quotients.

Considering reflective boundary conditions (chain "bounces" at r=0 and r=r_{max}), the coefficients are modified in the first and last two columns of the matrix to set the "flux" across the boundary to zero (Neumann boundary conditions). The time part of the differential equation is transform in a 1st order coupled linear differential equation, which can then be solved analytically as the sum of exponential function with the matrix inversion (Eigensystem) as described by [319]. The resulting theoretical fluorescence decay curves are dependent on D_{DA}, b, σ, τ_D, and R_0.

Section 2:

A $\overline{p}^*\left(r,t=0\right)$ of 1 was used as initial conditions and the solution of the PDEs $\overline{p}^*\left(r,t_i\right)$ was calculated for each time t_i in the raw data. This was followed by a reduction of the two-dimensional solution to a one-dimensional time trace $F^{norm}_{D_{DA},b,\sigma}\left(t_i\right)$ (equation 2.12) and with a normalization to a maximum value of 1 [159]. To obtain the reconvolution fit for the final analysis, the measured IRF, which is a non-ideal form, was convolved with the theoretical fluorescence decay function and presented as follows:

$$F^{fit}_{total}\left(t_i\right)=IRF\left(t_i-t_0\right)\otimes\left\{A_1\cdot F^{norm}_{D_{DA},b,\sigma}\left(t_i\right)+A_n\cdot\exp\left(\tau_n t_i\right)+A_{dc}\right\} \qquad \text{(Eq. 6.4)}$$

The parameters describe: t_0 for the position of the fluorescence signal of the excitation pulse in the time trace, A_1 for the main amplitude, A_n is the amplitude of the difference of the dark counts (A_{dc}) between the raw data and the time trace of the buffer signal and additionally phase, τ_n takes into account photobleaching, stray light or stacking of the fluorophores.

It follows that the IRF decay curves were individually modified by time and amplitude compared to the main fluorescence decay curve. The same was done with the buffer signal. In addition, to obtain a good estimate of the buffer signal, for all time-resolved fluorescence measurements with the *pyrene* setup, the observed buffer signal was fitted as a single fluorescence decay with a lifetime of 2.0 ns. To avoid overfitting, for Q_{14} of the GdmCl and glycerol dependence, the secondary donor lifetime fluorescence decay was fitted with the same parameters, as for the main fluorescence decay, for time and amplitude position.

Two ranges for the distribution of values were determined according to the tear-off values. A linear distribution was considered for an early time point and a logarithmic distribution for later time points to reduce the weight of later time points data. A total number of 1000 time points were used. The fit parameters A_n, A_{dc}, A_1 and t_0, included in the reconvolution fit, do not affect the fit parameters D_{DA}, b and σ. A reduced χ^2 was used to optimize the reconvolution fit:

$$\chi^2 = \frac{1}{n-p} \sum_{i=1}^{n=1000} \frac{\left(F_{\exp}(t_i) - F_{total}^{fit}(t_i)\right)^2}{\left(\Delta F_{\exp}(t_i)\right)^2} \qquad \text{(Eq. 6.5)}$$

The following parameters are available: F_{exp} the raw data of the donor-acceptor labeled polypeptide chain with the buffer signal subtracted; $\Delta F_{exp}(t_i)$ represents the error at each time point t_i (including buffer and raw data errors).

A χ^2 close to 1 represents a well-defined analysis, with a χ^2 of 1.3 representing the acceptable upper limit. To find the optimum convolutional fit, a χ^2 was calculated for each FRET pair and the arithmetic mean χ^2_{global} was used. This procedure was implemented using the Matlab program "fminsearch". Based on the Nelder-Mead simplex direct search, the minimized χ^2_{global} was compared with the previous ones for one iteration of the three fitting parameters D_{DA}, b and σ [159]. Three new fitted parameters D_{DA}, b and σ were then estimated. This process was repeated until a minimum χ^2_{global} was found [310].

6.6.3 Support Plane Analysis

As developed and described by Ursula Zinth [159], the Matlab function "fminsearch" could not calculate fitting error directly. So the *Support Plane Analysis (SPA)* method was introduced to estimate the fitting error [159]. *SPA* is a statistical analysis and analyzes only a measured data set, as used for fluorescence lifetime data analysis (PicoQuant software) [158], [159].

Equation 6.6 below shows that the F-value $F(p,v,P)$ was calculated from the found optimized χ^2_{global} of the three fitted parameters divided by χ^2_{global}.

In this process, one of the three fitted parameters was fixed and changed to lower or higher values to obtain the fitting error of the chosen parameter, as shown for the parameter *D*:

$$F(p,\nu,P)=\frac{\chi^2_{global}\left(D_{fix},b_{free}\left(D_{fix}\right),\sigma_{free}\left(D_{fix}\right)\right)}{\chi^2_{global}\left(\left(D_{free},b_{free},\sigma_{free}\right)\right)} \qquad \text{(Eq. 6.6)}$$

D_{fix} denotes the fixed parameter and was change to lower and higher values, to yield a parabolic shape as shown in the Results 4.2.2 and to calculated the error. D_{free} represents the optimized free fitted parameter, which was found in the global analysis. An F-value of 1.02 was used to calculated the error represents for both FRET pair traces, where ν denotes the degrees of freedom, p is the used parameters and P denotes the probability of 0.68 (corresponding to the standard deviation) [159]. For each buffer or temperature point and each polypeptide chain the corresponding *SPA* was done. For the calculation of the root mean square end-to-end distance error the fitted parameter b in the EDWARDS polymer model was replaced by the root mean square end-to-end distance with equation 6.7, which allows direct calculation of the *SPA* for the root mean square end-to-end distance. For this replacement the $0.5< b/\sigma<1.5$ must be true.

$r_0 = 5.41219*(-0.473929*\sigma+\text{sqrt}(0.369536*r_0*\sigma-0.226131*\sigma\text{^}2))$ (Eq. 6.7)

6.7 Other used programs

Besides Matlab, IgorPro8 was used to calculate the population of end-to-end distance as well as the Förster distance programmed by Dr. Ursula Zinth. For further calculation and data analysis of the resulting fitted parameters, like D_{DA} getting from the global analysis, the ProFit (6.2.14) was used. Calculated *SPA* errors were averaged and included in the fit procedure of ProFit under "Error types". The resulting fitting errors were calculated with the help of the "Levenberg-Marquardt" algorithms.
CemDraw® was used to create peptide structure as a 2D-structure for illustration.

6.8 Dynamic light scattering (DLS)

The DLS measurements were performed with an ALV/LSE-5003 Light Scattering Electronics and Multiple Tau Digital Correlator with a compact goniometer system and a 1.8 W, 532 nm LASER (Cobolt Samba™ 500, 532 nm). The ALV Correlator Software V.3.0 was used to calculate the apparent hydrodynamic radius. Twenty independent measurements for 20 s were chosen and averaged, to achieve a sufficient signal to noise ratio. Samples for Q_{14} was prepared with the disassembly protocol to achieve highest concentration of the monomeric peptide in pure water at 22.5 °C (estimated at 70 - 100 µM). To prevent dust particles, all cuvettes were washed with pure methanol and dried under nitrogen stream.

6.9 Synchrotron small-angle X-ray scattering (SAXS)

The measurement was performed in cooperation with PD Dr. Stephan König on the beamline P12 of the storage ring PETRA III of the EMBL on the Deutschen Elektronen Synchrotron (DESY) at Hamburg, Germany [311], [312]. For the

detection a Dectris 2M Pilatus detector was used (Dectris AG, Baden, Switzerland), with a wavelength of 0.15 nm and an exposure time of 30 to 50 ms with a resolution range from 300 to 0.5 nm. The data processing, buffer subtraction and first estimation of the radius of gyration was done with the Software PRIMUS-MAR and AUTOMAR [313]–[315]. For the estimation of radius of gyration, the data were fitted with the Debye formula in combination with a power-law form to consider the assumption of scale invariance (detailed information see Appendix Figure A.2) [316]. It should be noted that the data and analysis from the SAXS experiment are not completely reliable and require further measurements and evaluations. Samples for Q_{14} was prepared with the disassembly protocol to achieved a sufficient concentration of the monomeric Q_{14} in pure water at 20 °C (estimated at 70 – 100 μM). The measurements were performed by Philipp Skorupa and PD Dr. König (Martin-Luther University Halle-Wittenberg, Institute of Biochemistry and Biotechnology). The data was analyzed by Prof. Dr. Saalwächter (Martin-Luther University Halle-Wittenberg, Institute of Physics).

6.10 Circular dichroism (CD)

All CD measurements were carried out on an AVIV 410 or AVIV 410/430 Circular Dichroism spectropolarimeter (AVIV, USA) with bandwidth of 1 nm. Signals were recorded in a 10 mm cuvette for 8 s to measure 4 scans, which were further averaged. The limits of the wavelength ranges were chosen to avoid dynode voltage signal crossing a upper limit of 850 V. The concentration of the investigated peptides were adjusted to 0.87 AU (AnalytikJena Specord S 600) at 200 nm for an optimal signal to noise ratio (the wavelength of around 200 nm represents key structural information of unfolded polypeptide chain for CD measurements).

6.11 Analytical Ultracentrifugation (aUZ)

All measurements were done in cooperation with Prof. Dr. Lilie (Martin-Luther-University Halle Wittenberg, Institute of Biochemistry and Biotechnology) on Beckman Coulter Optima XL-A Analytical Ultracentrifuge orProteomeLab XL-1 Protein Characterization System at 22.5 °C in 10 mM potassium phosphate, pH 7.0 and a optical density of around 0.3 AU at 340 nm. Time scan at 10 min.

6.12 Synthesis of Xanthonic acid, improved protocol

The Xanthonic acid (9-oxoxanthen-2-carboxylic acid) is used for the TTET experiment as donor and synthesized in three steps. Here an improved protocol that increases yield of 300% over previous method of Lewis and Graham [317].
Step I for (4-(2-Methoxybenzoyloxy) benzoate): 7 g anhydrous potassium carbonate, 3.4 g methyl-4-hydroxybenzoate and 150 ml acetone (dried with molecular sieve) were added and carefully heated. After heating, 3.7 g 2-methoxybenzoyl chloride was added and cooked under reflux for 2 hours. Acetone was removed under reduced pressure, which results in yellow crystals. **Step II** for (4-Hydroxy-3-(2-Methoxybenzoyl)benzoic acid): 3 g of 4-(2-Methoxybenzoyloxy) benzoate was dissolved in 300 ml benzol and transferred to a quartz UV-reactor (Groschopp and Co. 4060 Viersen 1). The solution was irradiated with a mercury lamp for 72 hours under reflux. **Step III** (Xanthonic acid): After irradiation and cooling down of the solution, the solution was transfers in a round shank. The solution was further cooled down with an ice bath and 50 - 100 ml 4 M potassium hydroxide was dropwise added under strong stirring. The solution was stored over night at about 4 °C to obtain white crystals. After filtration (frit), white crystals were dissolved in 50 - 100 ml of pure water maintained at 60 °C and then transferred in a separatory funnel. The solution was acidulated to a pH of around 1 (test with indicator paper) with 1 M hydrochloride acid so that a white precipitate was observed. Ethyl acetate was added under careful mixing so that the white precipitate was completely dissolved (50 – 100 ml). The upper phase was isolated and the solvent was removed under reduced pressure and reflux. For the recrystallization, the whitish crystals were nearly completely dissolved

with pure methanol (about 50 – 100 ml, depends on the yield) under reflux. After slow cooling, the solution was transferred to an ice bath and stored over night at about 4 °C. After filtration (frit), the whitish crystals were washed with cool pure methanol. The resulting end product was dried and analyzed with LC-MS (nearly 99% of peak purity was achieved).

6.13 Combined equation for analysis of viscosity dependence of Q_{14}

Equation 4.4 shows a combined equation for analysis of the two separated parts, which were observed in the viscosity dependence of Q_{14}. Therefore, a standard equation for calculation of the Gibbs free enthalpy of denaturant transition curves of proteins was used [318]. Both linear terms of the standard equation were replaced by the equation 4.4 to fit all data points:

$$rise1 = D_0*((\eta/\eta_0)\text{^}(-\alpha1));\ \ rise2 = D_0*((\eta/\eta_0)\text{^}(-\alpha2));\ \ \Delta G=x+y*\eta;$$
$$D = (rise1+rise2*exp(\Delta G/constant))/(1+exp(\Delta G/constant));$$

Where η is the viscosity and η_0 is the viscosity of the buffer. D is the intrachain diffusion coefficient and D_0 is the intrachain diffusion coefficient without viscosity. α_1 and α_2 are viscosity-sensitive exponents. ΔG describes the turning point with the help of x and y [318].

7. Appendix

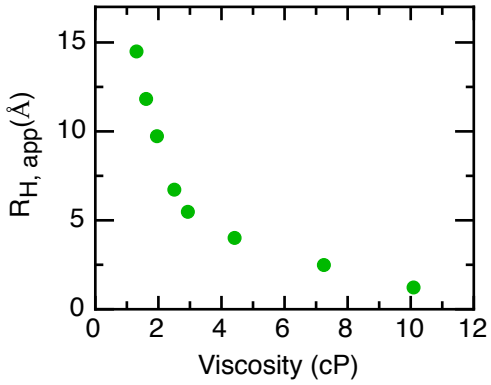


Figure A.1: DLS measurements of PEG_{6000} at different concentrations (0% to 21.1 % (w/w)). The decrease of the apparent hydrodynamic radius with increasing co-solute concentration in a nonlinear manner implies repulsive interaction of the co-solutes and the physical inhomogeneity of the medium with increasing concentration.

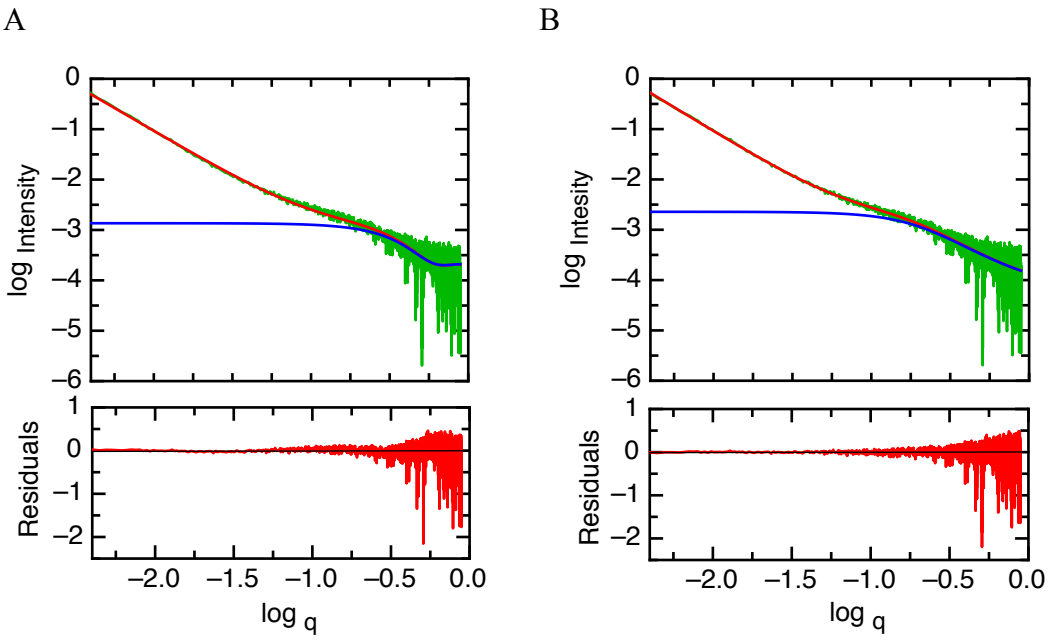


Figure A.2: Determination of the radius of gyration R_G of DNB-Q_{14}-AMCA in pure water. Measurements were done with synchrotron SAXS (EMBL) and performed by Philipp Skorupa and PD Dr. König (Martin-Luther University Halle-Wittenberg, Institute of Biochemistry and Biotechnology). The data was analyzed by Prof. Dr. Saalwächter (Martin-Luther-University Halle-Wittenberg, Institute of Physics), where the scattering intensity is plotted against the scatting vector *q (1/Å)* in a double-logarithmic plot. The data were fitted using **A** the sphere form factor or **B** the Debye formula. In order to consider the assumption of scale invariance, the data were fitted using the power-law form (red lines) and were also fitted without the power-law form (blue lines) [316].

It is shown in Figure A.2 that the implementation of a power-law form is necessary to describe the data completely. The residuals of the fits using the power-law form are plotted at the bottom (red). A R_G of 0.51 ± 0.1 nm for the sphere form factor (power-law form) and a R_G of 0.81 ± 0.1 nm for the Debye formula (power-law form) was found. The resulting errors of the R_G values were estimated. As can be seen from the residuals, the Debye formula in combination with the power-law form yielded the best fit and was used in the chapter Results and Discussion.

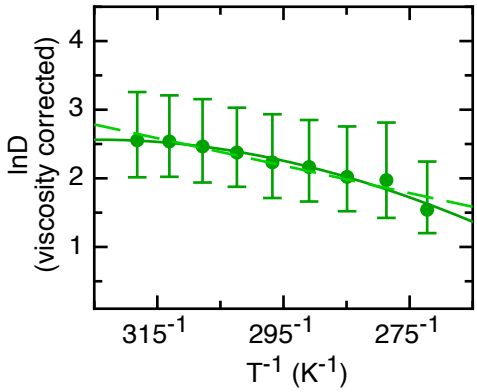
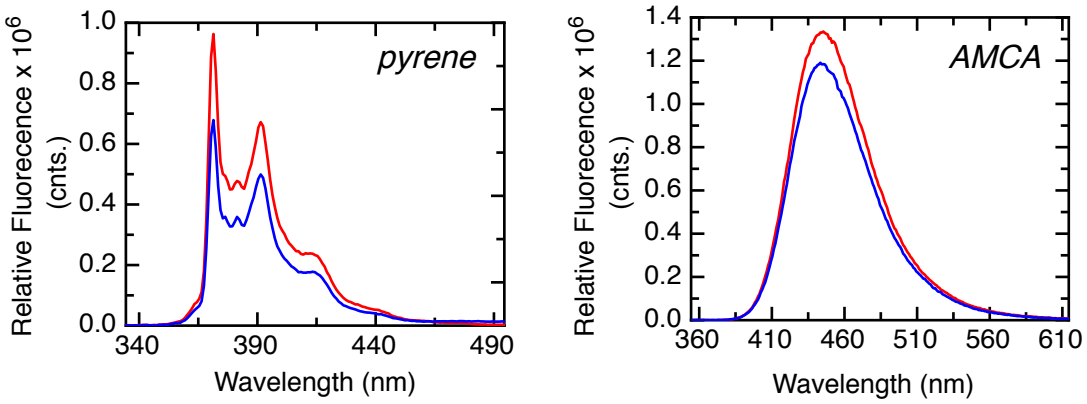


Figure A.3: *Eyring* plot of viscosity-corrected *ln(intrachain diffusion coefficient)* against *1/T* of Q_{14}. A Δc_p of -0.62 ± 0.92 is found (green solid line). In aspect of the high inaccuracy of the found Δc_p resulting from the error of the intrachain diffusion coefficient, an analysis with $\Delta c_p = 0$, as a fixed value, is done and shown as the dashed green line. The resulting hyperbolic slope of the free Δc_p analysis shows a good description of the data, but the analysis with the fixed $\Delta c_p = 0$ also described the data similarly well, relating to the error of the intrachain diffusion coefficient. Therefore, it was concluded that the $\Delta c_p \approx 0$ and no thermodynamic stable sub-structures were found.

A



B

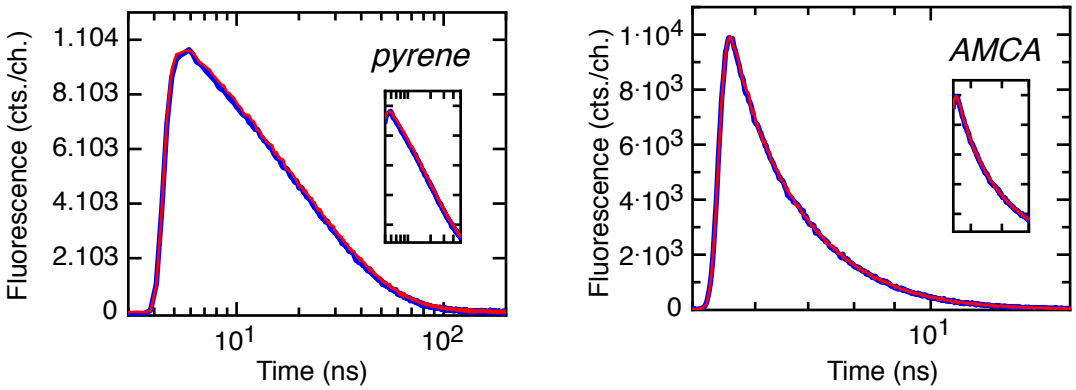


Figure A.4: Analysis of the pH dependence on the Förster distances and the time-resolved fluorescence decay curves of S_{14}. The sample with the pH of 5.0 ± 0.3 are depicted in red and the samples with a pH of 7.0 ± 0.3 (adjusted with sodium hydroxide) are shown in blue for **A** the fluorescence spectra for the determination of the Förster distance of both Ac-S_{14}-*pyrene* (left side) and Ac-S_{14}-AMCA (right side) and **B** the time-resolved fluorescence decay curves of both dansyl-S_{14}-pyrene (left side) and DNB-S_{14}-AMCA (right side). The calculated Förster distance for *pyrene* labeled polypeptide chains is 29.6 Å at pH 7 and 30.2 Å at pH 5 with a relative alteration of 1.9% and for *AMCA* labeled polypeptide chain 20.2 Å at pH 7 and 21.3 Å at pH 5 with a relative alteration of 5.3%. The time-resolved fluorescence decay for both labeled polypeptide chains show no alteration between pH 7 and 5, as depicted in the inlet (zoomed at early times).

Table of content for the attached CD -Data_ThesisPeterEnke-

1. Decay curves_globalFitAllpeptides
2. DLS_Q14
3. GS8inPEG6000
4. Fluorescence decay curves raw data
5. LC-MS data (examples)
6. Picture_Thesis
7. PV(DE)
8. PV(EF)inPEG6000
9. Q3inWater
10. Q12inWater
11 Q14
12. S14
13. SAXS_Q14
14. SPA_AllPeptides
15. Test_MonomerQ14

Explanation of the folders:
All measurement data of the measured polypeptide chains can be found in the respective folders, subdivided according to the different buffers and temperatures. The individual folders contain all evaluated data and the corresponding plots, as well as the calculation of the R_0 value. The subfolder "forFit" contains the raw data of all fluorescence decay curves (where the lowest concentration or temperature starts at xxx_0), as well as the input tables containing all constants and resulting fit parameters (inputsXxxFinal contains all parameters and inputsXxxFinalrms, where *b* is replaced by the root mean square end-to-en distance for the calculation of the error with the SPA). These input tables can also be used directly for global analysis. The input tables are structured as follows. The constants in column 2= R_0 of pyrene; column 3= R_0 of AMCA; column 5= τ_D of pyrene; column 6= τ_D of AMCA and the resulting fit parameters in columns 7= secondary lifetime of Q_{14}; column 28= D_{DA}; column 29= *b*; column 30= σ; column 31= root mean square end-to-end distance; column 32= χ^2_{pyrene}; column 33= χ^2_{AMCA} A; column 35= χ^2_{global}.

In addition, the "forFit" folder contains the resulting *SPA* calculations SuppXXXFinal and SuppXXXFinalrms (rms= root mean square end-to-end distance).
The folder "Picture_Thesis" contains all plots and data used in this thesis. This is again subdivided into the respective polypeptide chains and corresponding buffers or temperature.
The folder "SPA_AllPeptides" contains the calculated error for all polypeptide chains. Furthermore, the folder "Decay curves_globalFitAllpeptides" contains all fluorescence decay curves presented in this work, for the respective polypeptide chains.
The folder "Test_MonomerQ14" contains data for the three different measurement methods (analytical ultracentrifugation; CD and fluorimeter measurements) for the detection of the monomeric state of Q_{14}.
The following programs were used: Matlab, IgorPro8, ProFit, and PicoQuant FluoFit (see Material and Methods).

8. List of abbreviations

Ac-	acetyl-
ACN	Acetonitrile
AMCA	7-amino-4-methyl-3-coumarin
aUZ	analytical ultracentrifugation
CD	circular dichroism
CREB	cAMP response element-binding protein
cts./ch.	counts per channel
DBO	(2,3-dazabicyclo[2.2.2]oct-2-ene)
DCM	dichloromethane
DIC	N,N'-Diisopropylcarbodiimide
DIPEA	N- diisopropylethylamine
DLS	dynamic light scattering
DMF	dimethylformamide
DNB	2,4-dinitrobenzene (bound form 2,4-dinitrophenyl)
Dpr(Mtt)	α, β-diaminopropionic acid (4-methyltrityl)
DRPLA	denatoruubral pallidoluysian atrophy
Fmoc	9-fluoroenylmethyloxycarbonyl
FRET	fluorescence/Förster resonance energy transfer
FTIR	fourier-transform infrared spectroscopy
$(GS)_8$	$(glycine\text{-}serine)_8$ polypeptide chains
GdmCl	guanidine hydrochloride
HATU	O-(7-Azabenzotriazol-1-yl)-N,N,N′,N′-tetramethyluronium hexafluorphosphate
HBTU	2-(1H-benzotriazol-1-yl)-1,1,3,3-tetramethyluronium hexafluorophosphate
HD	Huntington disease
HFIP	hexafluoro-2-propanol
HOBT	1- Hydroxybenzotriazole hydrate
HPLC	high-pressure liquid chromatography
IDR/IDP	intrinsically disordered regions / proteins
IRF	instrument response function
LC-MS	liquid chromatography mass spectrometry
MC	Monte Carlo

NMP	N-Methyl-2-pyrrolidon
NMR	nuclear magnetic resonance
Oxyma	ethyl 2- cyano-2-(hydroxyimino)acetate)
PDE	partial differential equation
PEG_{6000}	Polyethylene glycol 6000 (average molecular weight)
PET-FCS	photoinduced electron transfer and fluorescence correlation spectroscopy
pKID/KIX	phosphorylated kinase-inducible domain/kinase interacting domain
Q_{14}	$(glutamine)_{14}$ polypeptide chains
QPB1	polyQ binding protein 1
RAM	rink-amide
ROA	raman optical activity
RP	reversed phase
S_{14}	$(serine)_{14}$ polypeptide chains
SANS	small-angle neutron scattering
SAXS	small-angle X-ray scattering
SBMA	spinal and bullar muscular atrophy
SCAx	spinocerebellar ataxia x
SPA	*Support Plane Analysis*
SPPS	solid-phase peptide synthesis
SSS-theory	Szabo-Schulten-Schulten theory
TCSPC	time-correlated single photon counting
TFA	trifluoroacetic acid
TIPS	triisopropylsilane
trFRET	time-resolved fluorescence/Förster resonance energy transfer
TTET	Triplet-triplet energy transfer

protection groups:

tBu (tert-butyl); Pbf (2,2,4,6,7-pentamethyl-2,3-dihydrobenzofuran-5-sulfonyl); OtBu (O-t-butyl); Dmb (2,4-dimethoxybenzyl); Boc (tert-butyl); tBu (tert-butyl); Trt (trityl); Me (methyl)

amino acids:

Gly (glycine); Ser (serine); Thr (threonine); Arg (arginine); Leu (leucine); Lys (lysine); Ala (alanine); Asp (asparagine); Val (valine); Glu (glutamate); Gln (glutamine)

9. Bibliography

[1] J. R. Wiśniewski, M. Y. Hein, J. Cox, and M. Mann, "A 'Proteomic Ruler' for Protein Copy Number and Concentration Estimation without Spike-in Standards," *Mol. Cell. Proteomics*, vol. 13, no. 12, pp. 3497–3506, 2014.

[2] J. A. Rivas-Pardo, E. C. Eckels, I. Popa, P. Kosuri, W. A. Linke, and J. M. Fernández, "Work Done by Titin Protein Folding Assists Muscle Contraction," *Cell Rep.*, vol. 14, no. 6, pp. 1339–1347, 2016.

[3] M. Su, Y. Ling, J. Yu, J. Wu, and J. Xiao, "Small proteins: Untapped area of potential biological importance," *Front. Genet.*, vol. 4, pp. 1–9, 2013.

[4] M. I. Galindo, J. I. Pueyo, S. Fouix, S. A. Bishop, and J. P. Couso, "Peptides encoded by short ORFs control development and define a new eukaryotic gene family," *PLoS Biol.*, vol. 5, no. 5, pp. 1052–1062, 2007.

[5] C. B. Anfinsen, "Folding of Protein Chains," *Science*, vol. 181, no. 4096, pp. 223–230, 1973.

[6] J. S. Valastyan and S. Lindquist, "Mechanisms of protein-folding diseases at a glance," *Dis. Model. Mech.*, vol. 7, no. 1, pp. 9–14, 2014.

[7] F. U. Hartl and M. Hayer-Hartl, "Converging concepts of protein folding in vitro and in vivo," *Nat. Struct. Mol. Biol.*, vol. 16, no. 6, pp. 574–581, 2009.

[8] D. W. Bolen and G. D. Rose, "Structure and Energetics of the Hydrogen-Bonded Backbone in Protein Folding," *Annu. Rev. Biochem.*, vol. 77, no. 1, pp. 339–362, 2008.

[9] C. Levinthal, "How to fold graciously. Mossbauer Spectroscopy in Biological Systems: Proceedings of a meeting held at Allerton House," *Monticello, Illinois (Debrunnder JTP, Munck E.., eds.)*, pp. 22–24, 1969.

[10] S. E. Jackson, "How do small single-domain proteins fold?," *Fold. Des.*, vol. 3, no. 4, pp. 81–91, 1998.

[11] A. R. Fersht, "Folding : Delights and Scepticisms," *Nat. Rev. Mol. Cell Biol.*, vol. 9, pp. 650–654, 2008.

[12] J. B. Udgaonkar, "Multiple Routes and Structural Heterogeneity in Protein Folding," *Annu. Rev. Biophys.*, vol. 37, no. 1, pp. 489–510, 2008.

[13] C. B. Anfinsen, "The formation and stabilization of protein structure.," *Biochem. J.*, vol. 128, no. 4, pp. 737–49, 1972.

[14] O. B. Ptitsyn and A. A. Rashin, "Stagewise mechanism of protein folding," *Dokl. Akad. Nauk SSSR*, vol. 213, pp. 473–475, 1973.

[15] P. S. Kim and R. L. Baldwin, "Specific Intermediates in the Folding Reactions of Small Proteins and the Mechanism of Protein Folding," *Annu. Rev. Biochem.*, vol. 51, no. 1, pp. 459–489, 1982.

[16] R. L. Baldwin and G. D. Rose, "Is protein folding hierarchic? II. Folding intermediates and transition states [published erratum appears in Trends Biochem Sci 1999 May;24(5):185]," *Trends Biochem Sci*, vol. 24, no. February, pp. 77–83, 1999.

[17] M. Karplus and D. L. Weaver, "Protein folding dynamics: The diffusion-collision model and

experimental data," *Protein Sci.*, vol. 3, no. 4, pp. 650–668, 1994.

[18] A. R. Fersht, "Nucleation mechanisms in protein folding Alan R Fersht," *Curr. Opin. Struct. Biol.*, 1997.

[19] V. Daggett and A. Fersht, "Protein folding and unfolding at atomic resolution," *Cell*, vol. 108, pp. 573–582, 2002.

[20] O. B. Ptitsyn, "Protein folding: Hypotheses and experiments," *J. Protein Chem.*, vol. 6, no. 4, pp. 273–293, 1987.

[21] O. B. Ptitsyn and V. N. Uversky, "The molten globule is a third thermodynamical state of protein molecules," *FEBS Lett.*, vol. 341, no. 1, pp. 15–18, 1994.

[22] C. Wagner and T. Kiefhaber, "Intermediates can accelerate protein folding," *Proc. Natl. Acad. Sci.*, vol. 96, no. 12, pp. 6716–6721, 1999.

[23] P. Kim, "Intermediates In The Folding Reactions Of Small Proteins," *Annu. Rev. Biochem.*, vol. 59, no. 1, pp. 631–660, 1990.

[24] C. Matthews, "Pathways of Protein Folding," *Annu. Rev. Biochem.*, vol. 62, no. 1, pp. 653–683, 1993.

[25] W. K. Lim, J. Rösgen, and S. W. Englander, "Urea, but not guanidinium, destabilizes proteins by forming hydrogen bonds to the peptide group," *Proc. Natl. Acad. Sci.*, vol. 106, no. 8, pp. 2595–2600, 2009.

[26] K. A. Dill, J. L. Maccallum, and P. Folding, "The Protein-Folding Problem , 50 Years On," *Science*, vol. 338, pp. 1042–1047, 2012.

[27] M. Tsytlonok and L. S. Itzhaki, "The how's and why's of protein folding intermediates," *Arch. Biochem. Biophys.*, vol. 531, no. 1–2, pp. 14–23, 2013.

[28] A. Bachmann and T. Kiefhaber, "Apparent two-state tendamistat folding is a sequential process along a defined route," *J. Mol. Biol.*, vol. 306, no. 2, pp. 375–386, 2001.

[29] I. E. Sánchez and T. Kiefhaber, "Evidence for sequential barriers and obligatory intermediates in apparent two-state protein folding," *J. Mol. Biol.*, vol. 325, no. 2, pp. 367–376, 2003.

[30] E. Fischer, "Einfluss der Configuration auf die Wirkung der Enzyme," *Aus dem I. Berliner Univ.*, vol. 27, pp. 2985–2993, 1894.

[31] D. Shortle, "The denatured state - the other half of the folding N-D," *FASEB J.*, vol. 10, no. 1, pp. 27–34, 1996.

[32] K. A. Dill and D. Shortie, "Denatured states of proteins," *Annu. Rev.*, pp. 795–825, 1991.

[33] A. K. Dunker *et al.*, "Intrinsically disordered protein," vol. 3263, no. 00, pp. 26–59, 2001.

[34] V. N. Uversky, "Natively unfolded proteins A point.pdf," pp. 739–756, 2002.

[35] H. Wu, "Studies on Denaturation of Proteins," *Adv. Protein Chem.*, vol. 46, no. 4, pp. 6–26, 1931.

[36] J. A. Schellman, "Fifty years of solvent denaturation," *Biophys. Chem.*, vol. 96, no. 2–3, pp. 91–101, 2002.

[37] C. Tanford, "Protein denaturation," *This Week's Cit. Class.*, p. 123, 1980.

[38] A. K. Jha, A. Colubri, K. F. Freed, and T. R. Sosnick, "Statistical coil model of the unfolded

state: Resolving the reconciliation problem," *Proc. Natl. Acad. Sci.*, vol. 102, no. 37, pp. 13099–13104, 2005.

[39] J. J. Ward, J. S. Sodhi, L. J. McGuffin, B. F. Buxton, and D. T. Jones, "Prediction and Functional Analysis of Native Disorder in Proteins from the Three Kingdoms of Life," *J. Mol. Biol.*, vol. 337, no. 3, pp. 635–645, 2004.

[40] Y. Cheng, V. N. Uversky, C. J. Oldfield, M. S. Cortese, A. K. Dunker, and C. J. Brown, "Comparing and combining predictors of mostly disordered proteins.," *Biochemistry*, vol. 44, no. 6, pp. 1989–2000, 2005.

[41] Y. Huang and Z. Liu, "Kinetic Advantage of Intrinsically Disordered Proteins in Coupled Folding-Binding Process: A Critical Assessment of the 'Fly-Casting' Mechanism," *J. Mol. Biol.*, vol. 393, no. 5, pp. 1143–1159, 2009.

[42] L. M. Iakoucheva, C. J. Brown, J. D. Lawson, Z. Obradović, and A. K. Dunker, "Intrinsic disorder in cell-signaling and cancer-associated proteins," *J. Mol. Biol.*, vol. 323, no. 3, pp. 573–584, 2002.

[43] M. Wells *et al.*, "Structure of tumor suppressor p53 and its intrinsically disordered N-terminal transactivation domain," *Proc. Natl. Acad. Sci.*, vol. 105, no. 15, pp. 5762–5767, 2008.

[44] J. Dyson, P. E. Wright, and H. J. Dyson, "Intrinsically Unstructured Proteins : Re- Assessing the Protein Structure-Function Paradigm," *J. Mol. Biol.*, vol. 293, pp. 321–331, 1999.

[45] L. M. Iakoucheva *et al.*, "Intrinsic disorder and protein function," *Nuc. Acids Res.*, vol. 41, pp. 6573–6582, 2002.

[46] P. Tompa, "Intrinsically Unstructured Proteins," *Trends Biochem Sci*, vol. 27, no. 10, pp. 527–533, 2002.

[47] P. Tompa, "The interplay between structure and function in intrinsically unstructured proteins," *FEBS Lett.*, vol. 579, no. 15, pp. 3346–3354, 2005.

[48] P. E. Wright *et al.*, "Classification of Intrinsically Disordered Regions and Proteins," *Chem. Rev.*, vol. 114, no. 13, pp. 6589–6631, 2014.

[49] J. Gsponer and M. Madan Babu, "The rules of disorder or why disorder rules," *Prog. Biophys. Mol. Biol.*, vol. 99, no. 2–3, pp. 94–103, 2009.

[50] L. M. Iakoucheva *et al.*, "The importance of intrinsic disorder for protein phosphorylation," *Nucleic Acids Res.*, vol. 32, no. 3, pp. 1037–1049, 2004.

[51] A. L. Fink, "Natively unfolded proteins," *Curr. Opin. Struct. Biol.*, vol. 15, no. 1 SPEC. ISS., pp. 35–41, 2005.

[52] V. N. Uversky, "Multitude of binding modes attainable by intrinsically disordered proteins: A portrait gallery of disorder-based complexes," *Chem. Soc. Rev.*, vol. 40, no. 3, pp. 1623–1634, 2011.

[53] A. K. Dunker and Z. Obradovic, "The protein trinity - Linking function and disorder," *Nat. Biotechnol.*, vol. 19, no. 9, pp. 805–806, 2001.

[54] B. A. Shoemaker, J. J. Portman, and P. G. Wolynes, “Speeding molecular recognition by using the folding funnel: The fly-casting mechanism,” *Proc. Natl. Acad. Sci.*, vol. 97, no. 16, pp. 8868–8873, 2000.

[55] P. Tompa, “Intrinsically disordered proteins : a 10-year recap,” vol. 37, no. 12, pp. 509–516, 2012.

[56] T. Kiefhaber, A. Bachmann, and K. S. Jensen, “Dynamics and mechanisms of coupled protein folding and binding reactions,” *Curr. Opin. Struct. Biol.*, vol. 22, no. 1, pp. 21–29, 2012.

[57] M. Fuxreiter, I. Simon, P. Friedrich, and P. Tompa, “Preformed structural elements feature in partner recognition by intrinsically unstructured proteins,” *J. Mol. Biol.*, vol. 338, no. 5, pp. 1015–1026, 2004.

[58] V. D. Kumar, L. Lee, and B. F. P. Edwards, “Refined crystal structure of calcium-liganded carp parvalbumin 4.25 at 1.5-.ANG. resolution,” *Biochemistry*, vol. 29, no. 6, pp. 1404–1412, 1990.

[59] F. Cordier, B. Hartmann, M. Rogowski, M. Affolter, and S. Grzesiek, “DNA Recognition by the Brinker Repressor - An Extreme Case of Coupling Between Binding and Folding,” *J. Mol. Biol.*, vol. 361, no. 4, pp. 659–672, 2006.

[60] S. Güthe, L. Kapinos, A. Möglich, S. Meier, S. Grzesiek, and T. Kiefhaber, “Very fast folding and association of a trimerization domain from bacteriophage T4 fibritin,” *J. Mol. Biol.*, vol. 337, no. 4, pp. 905–915, 2004.

[61] I. Radhakrishnan *et al.*, “Solution structure of the KIX domain of CBP bound to the transactivation domain of CREB: a model for activator:coactivator interactions,” *Cell*, vol. 91, pp. 741–752, 1997.

[62] R. R. Rustandi, D. M. Baldisseri, and D. J. Weber, “letters Structure of the negative regulatory domain of p53 bound to S100B (ββ),” *America (NY).*, vol. 7, no. 7, pp. 570–574, 2000.

[63] J. L. Avalos, S. Muhammad, M. S. Cosgrove, C. Wolberger, I. Celic, and J. D. Boeke, “Structure of a Sir2 enzyme bound to an acetylated p53 peptide,” *Mol. Cell*, vol. 10, no. 3, pp. 523–535, 2002.

[64] T. R. Weikl and C. Von Deuster, “Selected-fit versus induced-fit protein binding: Kinetic differences and mutational analysis,” *Proteins Struct. Funct. Bioinforma.*, vol. 75, no. 1, pp. 104–110, 2009.

[65] P. E. Wright and H. J. Dyson, “Intrinsically disordered proteins in cellular signalling and regulation,” *Nat. Rev. Mol. Cell Biol.*, vol. 16, no. 1, pp. 18–29, 2015.

[66] H. Zhou, “A Gaussian-chain model for treating residual charge – charge interactions in the unfolded state of proteins,” vol. 99, no. 6, pp. 3569–3574, 2002.

[67] R. A. Pethrick, “Polymer physics. Edited by Michael Rubinstein and Ralph H Colby Oxford University Press, Oxford, 2003. ISBN 019852059X. pp 440,” *Polym. Int.*, vol. 53, no. 9, pp. 1394–1395, 2004.

[68] N. Y. Ithaca, M. Doi, and S. F. Edwards, “The Theory of Polymer Dynamics.” Oxford: Clarendon, 1986.

[69] G. R. Strobl, *The physics of polymers*. Springer, 1997.

[70] G. Fuertes *et al.*, "Decoupling of size and shape fluctuations in heteropolymeric sequences reconciles discrepancies in SAXS vs. FRET measurements," *Proc. Natl. Acad. Sci. U. S. A.*, vol. 114, no. 31, pp. E6342–E6351, 2017.

[71] P. J. Flory, "The Configuration of Real Polymer Chains," *J. Chem. Phys.*, vol. 17, no. 3, pp. 303–310, 1949.

[72] P. J. Flory, *Principles of polymer chemistry*. Cornell University Press, 1953.

[73] I. C. Sanchez, "Phase Transition Behavior of the Isolated Polymer Chain," *Macromolecules*, vol. 12, no. 5, pp. 980–988, 1979.

[74] D. Thirumalai, H. S. Samanta, H. Maity, and G. Reddy, "Universal Nature of Collapsibility in the Context of Protein Folding and Evolution," *Trends Biochem. Sci.*, vol. 44, no. 8, pp. 675–687, 2019.

[75] K. C. Aune, A. Salahuddin, M. H. Zarlengo, and C. Tanford, "Evidence for residual structure in acid- and heat-denatured proteins.," *J. Biol. Chem.*, vol. 242, no. 19, pp. 4486–4489, 1967.

[76] C. Tanford, "Proteins in 6 M Guanidine Hydrochloride," *J. Biol. Chem.*, vol. 241, pp. 1921–1923, 1966.

[77] O. Tcherkasskaya and V. N. Uversky, "Denatured collapsed states in protein folding: Example of apomyoglobin," *Proteins Struct. Funct. Genet.*, vol. 44, no. 3, pp. 244–254, 2001.

[78] J. E. Kohn *et al.*, "Random-coil behavior and dimensions of chemically unfolded proteins," *Proc. Natl. Acad. Sci. USA*, vol. 101, no. 34, pp. 12691–12696, 2004.

[79] D. P. Goldenberg, "Computational simulation of the statistical properties of unfolded proteins," *J. Mol. Biol.*, vol. 326, no. 5, pp. 1615–1633, 2003.

[80] D. K. Wilkins, S. B. Grimshaw, V. Receveur, C. M. Dobson, J. A. Jones, and L. J. Smith, "Hydrodynamic radii of native and denatured proteins measured by pulse field gradient NMR techniques," *Biochemistry*, vol. 38, no. 50, pp. 16424–16431, 1999.

[81] I. S. Millett, S. Doniach, and K. W. Plaxco, "Toward a taxonomy of the denatured state: Small angle scattering studies of unfolded proteins," *Adv. Protein Chem.*, vol. 62, pp. 241–262, 2002.

[82] R. I. Dima and D. Thirumalai, "Asymmetry in the shapes of folded and denatured states of proteins," *J. Phys. Chem. B*, vol. 108, no. 21, pp. 6564–6570, 2004.

[83] J. C. Le Guillou and J. Zinn-Justin, "Critical exponents for the n-vector model in three dimensions from field theory," *Phys. Rev. Lett.*, vol. 39, no. 2, pp. 95–98, 1977.

[84] M. Sadqi, L. J. Lapidus, and V. Muñoz, "How fast is protein hydrophobic collapse?," *Proc. Natl. Acad. Sci. U. S. A.*, vol. 100, no. 21, pp. 12117–12122, 2003.

[85] H. S. Chung, I. V. Gopich, K. McHale, T. Cellmer, J. M. Louis, and W. A. Eaton, "Extracting rate coefficients from single-molecule photon trajectories and FRET efficiency histograms for a fast-folding protein," *J. Phys. Chem. A*, vol. 115, no. 16, pp. 3642–3656, 2011.

[86] J. K. Lum, H. Neuweiler, and A. R. Fersht, "Long-range modulation of chain motions within the intrinsically disordered transactivation domain of tumor suppressor p53," *J. Am. Chem. Soc.*, vol. 134, no. 3, pp. 1617–1622, 2012.

[87] A. Soranno *et al.*, "Quantifying internal friction in unfolded and intrinsically disordered proteins with single-molecule spectroscopy," *Proc. Natl. Acad. Sci. U. S. A.*, vol. 109, no. 44, pp. 17800–17806, 2012.

[88] W. Zheng, G. H. Zerze, A. Borgia, J. Mittal, B. Schuler, and R. B. Best, "Inferring properties of disordered chains from FRET transfer efficiencies," *J. Chem. Phys.*, vol. 148, no. 12, 2018.

[89] B. Schuler and H. Hofmann, "Single-molecule spectroscopy of protein folding dynamics-expanding scope and timescales," *Curr. Opin. Struct. Biol.*, vol. 23, no. 1, pp. 36–47, 2013.

[90] O. Kratky and G. Porod, "Röntgenuntersuchung gelöster fadenmoleküle," *Recl. des Trav. Chim. des Pays-Bas*, vol. 68, no. 12, pp. 1106–1122, 1949.

[91] D. Thirumalai and B.-Y. Ha, "Statistical Mechanics of Semiflexible Chains: A Meanfield Variational Approach," *Mathmatical Model. Polym. Res.*, pp. 1–35, 1997.

[92] C. Bouchiat, M. D. Wang, J. F. Allemand, T. Strick, S. M. Block, and V. Croquette, "Estimating the persistence length of a worm-like chain molecule from force-extension measurements," *Biophys. J.*, vol. 76, no. 1 I, pp. 409–413, 1999.

[93] E. M. Puchner, G. Franzen, M. Gautel, and H. E. Gaub, "Comparing proteins by their unfolding pattern," *Biophys. J.*, vol. 95, no. 1, pp. 426–434, 2008.

[94] F. Berkemeier *et al.*, "Fast-folding α-helices as reversible strain absorbers in the muscle protein myomesin," *Proc. Natl. Acad. Sci. U. S. A.*, vol. 108, no. 34, pp. 14139–14144, 2011.

[95] L. J. Lapidus, P. J. Steinbach, W. A. Eaton, A. Szabo, and J. Hofrichter, "Effects of chain stiffness on the dynamics of loop formation in polypeptides. Appendix: Testing a 1-dimensional diffusion model for peptide dynamics," *J. Phys. Chem. B*, vol. 106, no. 44, pp. 11628–11640, 2002.

[96] M. Buscaglia, L. J. Lapidus, W. A. Eaton, and J. Hofrichter, "Effects of denaturants on the dynamics of loop formation in polypeptides," *Biophys. J.*, vol. 91, no. 1, pp. 276–288, 2006.

[97] J. Wilhelm and E. Frey, "Radial Distribution Function of Semiflexible Polymers," *Phys. Rev. Lett.*, vol. 77, no. 12, pp. 2581–2584, 1996.

[98] S. D. Stellman and P. J. Gans, "Computer Simulation of Polymer Conformation. II. Distribution Function for Polymers with Excluded Volume," *Macromolecules*, vol. 5, no. 6, pp. 720–729, 1972.

[99] E. Haas, M. Wilchek, E. Katchalski-Katzir, and I. Z. Steinberg, "Distribution of end-to-end distances of oligopeptides in solution as estimated by energy transfer.," *Proc. Natl. Acad. Sci. U. S. A.*, vol. 72, no. 5, pp. 1807–1811, 1975.

[100] E. Haas, E. Katchalski-Katzir, and I. Z. Steinberg, "Brownian motion of the ends of oligopeptide chains in solution as estimated by energy transfer between the chain ends," *Biopolymers*, vol. 17, no. 1, pp. 11–31, 1978.

[101] J. M. Beechem and E. Haas, "Simultaneous determination of intramolecular distance distributions and conformational dynamics by global analysis of energy transfer measurements," *Biophys. J.*, vol. 55, no. 6, pp. 1225–1236, 1989.

[102] A. Navon, V. Ittah, P. Landsman, H. A. Scheraga, and E. Haas, “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*, vol. 40, no. 1, pp. 105–118, 2001.

[103] A. Moglich, K. Joder, and T. Kiefhaber, “End-to-end distance distributions and intrachain diffusion constants in unfolded polypeptide chains indicate intramolecular hydrogen bond formation,” *Proc. Natl. Acad. Sci.*, vol. 103, no. 33, pp. 12394–12399, 2006.

[104] A. Grupi and E. Haas, “Segmental conformational disorder and dynamics in the intrinsically disordered protein α-synuclein and its chain length dependence,” *J. Mol. Biol.*, vol. 405, no. 5, pp. 1267–1283, 2011.

[105] A. Grupi and E. Haas, “Time-resolved FRET detection of subtle temperature-induced conformational biases in ensembles of α-synuclein molecules,” *J. Mol. Biol.*, vol. 411, no. 1, pp. 234–247, 2011.

[106] F. Huang *et al.*, “Multiple conformations of full-length p53 detected with single-molecule fluorescence resonance energy transfer,” *Proc. Natl. Acad. Sci. U. S. A.*, vol. 106, no. 49, pp. 20758–20763, 2009.

[107] S. F. Edwards, “The statistical mechanics of polymers with excluded volume,” *Proc. Phys. Soc.*, vol. 85, no. 4, pp. 613–624, 1965.

[108] G. Damaschun, H. Damaschun, K. Gast, C. Gernat, and D. Zirwer, “Acid denatured apo-cytochrome c is a random coil: evidence from small-angle X-ray scattering and dynamic light scattering.” pp. 289–295, 1991.

[109] A. Nöppert, K. Gast, M. Mfiller-frohne, D. Zirwer, and G. Damaschun, “Reduced-denatured ribonuclease A is not in a compact state,” vol. 380, pp. 179–182, 1996.

[110] M. Nygaard, B. B. Kragelund, E. Papaleo, and K. Lindorff-Larsen, “An Efficient Method for Estimating the Hydrodynamic Radius of Disordered Protein Conformations,” *Biophys. J.*, vol. 113, no. 3, pp. 550–557, 2017.

[111] E. E. Lattman, “Small angle scattering studies of protein folding,” *Curr. Opin. Struct. Biol.*, vol. 4, no. 1, pp. 87–92, 1994.

[112] S. Doniach, J. Bascle, T. Garel, and H. Orland, “Partially folded states of proteins: Characterization by X-ray scattering,” *J. Mol. Biol.*, vol. 254, no. 5, pp. 960–967, 1995.

[113] L. J. Smith, K. M. Fiebig, H. Schwalbe, and C. M. Dobson, “The concept of a random coil : Residual structure in peptides and denatured proteins,” *Fold. Des.*, vol. 1, no. 5, pp. 95–106, 1996.

[114] L. J. Lapidus, W. A. Eaton, and J. Hofrichter, “Measuring the rate of intramolecular contact formation in polypeptides,” *Proc. Natl. Acad. Sci.*, vol. 97, no. 13, pp. 7220–7225, 2000.

[115] L. J. Lapidus and V. R. Singh, “The Intrinsic Stiffness of Polyglutamine Peptides,” *J. Phys. Chem. B*, vol. 112, no. 42, pp. 13172–13176, 2008.

[116] I. C. Yeh and G. Hummer, “Peptide loop-closure kinetics from microsecond molecular dynamics simulations in explicit solvent,” *J. Am. Chem. Soc.*, vol. 124, no. 23, pp. 6563–6568, 2002.

[117] O. Bieri, J. Wirz, B. Hellrung, M. Schutkowski, M. Drewello, and T. Kiefhaber, "The speed limit for protein folding measured by triplet-triplet energy transfer," *Proc. Natl. Acad. Sci. U. S. A.*, vol. 96, no. 17, pp. 9597–9601, 1999.

[118] F. Krieger, B. Fierz, O. Bieri, M. Drewello, and T. Kiefhaber, "Dynamics of unfolded polypeptide chains as model for the earliest steps in protein folding," *J. Mol. Biol.*, vol. 332, no. 1, pp. 265–274, 2003.

[119] H. Satzger *et al.*, "Ultrafast quenching of the xanthone triplet by energy transfer: New insight into the intersystem crossing kinetics," *J. Phys. Chem. A*, vol. 108, no. 46, pp. 10072–10079, 2004.

[120] B. Heinz *et al.*, "On the unusual fluorescence properties of xanthone in water," *Phys. Chem. Chem. Phys.*, vol. 8, no. 29, pp. 3432–3439, 2006.

[121] I. J. Chang, J. C. Lee, J. R. Winkler, and H. B. Gray, "The protein-folding speed limit: Intrachain diffusion times set by electron-transfer rates in denatured Ru(NH3)5(His-33)-Zn-cytochrome c," *Proc. Natl. Acad. Sci. U. S. A.*, vol. 100, no. 7, pp. 3838–3840, 2003.

[122] R. R. Hudgins, F. Huang, G. Gramlich, and W. M. Nau, "A fluorescence-based method for direct measurement of submicrosecond intramolecular contact formation in biopolymers: An exploratory study with polypeptides," *J. Am. Chem. Soc.*, vol. 124, no. 4, pp. 556–564, 2002.

[123] F. Huang, R. R. Hudgins, and W. M. Nau, "Primary and secondary structure dependence of peptide flexibility assessed by fluorescence-based measurement of end-to-end collision rates," *J. Am. Chem. Soc.*, vol. 126, no. 50, pp. 16665–16675, 2004.

[124] H. Neuweiler, A. Schulz, M. Böhmer, J. Enderlein, and M. Sauer, "Measurement of submicrosecond intramolecular contact formation in peptides at the single-molecule level," *J. Am. Chem. Soc.*, vol. 125, no. 18, pp. 5324–5330, 2003.

[125] H. Neuweiler and M. Sauer, "Using Photoinduced Charge Transfer Reactions to Study Conformational Dynamics of Biopolymers at the Single-Molecule Level," *Curr. Pharm. Biotechnol.*, vol. 5, no. 3, pp. 285–298, 2004.

[126] H. Neuweiler, M. Löllmann, S. Doose, and M. Sauer, "Dynamics of Unfolded Polypeptide Chains in Crowded Environment Studied by Fluorescence Correlation Spectroscopy," *J. Mol. Biol.*, vol. 365, no. 3, pp. 856–869, 2007.

[127] H. Neuweiler, C. M. Johnson, and A. R. Fersht, "Direct observation of ultrafast folding and denatured state dynamics in single protein molecules," *Proc. Natl. Acad. Sci. U. S. A.*, vol. 106, no. 44, pp. 18569–18574, 2009.

[128] M. H. Jacob *et al.*, "Method-Unifying View of Loop-Formation Kinetics in Peptide and Protein Folding," *J. Phys. Chem. B*, vol. 122, no. 16, pp. 4445–4456, 2018.

[129] E. Nir *et al.*, "Shot-noise limited single-molecule FRET histograms: Comparison between theory and experiments," *J. Phys. Chem. B*, vol. 110, no. 44, pp. 22103–22124, 2006.

[130] U. B. Choi, J. J. McCann, K. R. Weninger, and M. E. Bowen, "Beyond the random coil: Stochastic conformational switching in intrinsically disordered proteins," *Structure*, vol. 19, no. 4, pp. 566–576, 2011.

[131] E. Sisamakis, A. Valeri, S. Kalinin, P. J. Rothwell, and C. A. M. Seidel, "Accurate Single-Molecule FRET Studies Using Multiparameter Fluorescence Detection," *Methods Enzymol.*, vol. 475, no. C, pp. 455–514, 2010.

[132] J. R. Lakowicz *et al.*, "Resolution of end-to-end diffusion coefficients and distance distributions of flexible molecules using fluorescent donor-acceptor and donor-quencher pairs," *Biopolymers*, vol. 31, no. 12, pp. 1363–1378, 1991.

[133] A. Möglich, F. Krieger, and T. Kiefhaber, "Molecular basis for the effect of urea and guanidinium chloride on the dynamics of unfolded polypeptide chains," *J. Mol. Biol.*, vol. 345, no. 1, pp. 153–162, 2005.

[134] M. H. Jacob, R. N. Dsouza, I. Ghosh, A. Norouzy, T. Schwarzlose, and W. M. Nau, "Diffusion-enhanced förster resonance energy transfer and the effects of external quenchers and the donor quantum yield," *J. Phys. Chem. B*, vol. 117, no. 1, pp. 185–198, 2013.

[135] T. Förster, "Intermolecular energy transfer and fluorescence.," *Ann. Phys. Leipzig.*, vol. 2, pp. 55–75, 1948.

[136] R. P. . Haugland, J. Yguerabide, and L. Stryer, "Dependence of the kinetics of singlet-singlet energy transfer on spectral overlap," *Proc. Natl. Acad. Sci. U. S. A.*, no. 2, pp. 23–30, 1969.

[137] I. Z. Steinberg, E. Haas, and E. Katchalski-Katzir, "Long-range nonradiative transfer of electronic excitation energy," in *Time-Resolved Fluorescence Spectroscopy in Biochemistry and Biology*, Springer, 1983, pp. 411–450.

[138] B. W. van der Meer, "Kappa-squared: From nuisance to new sense," *Rev. Mol. Biotechnol.*, vol. 82, no. 3, pp. 181–196, 2002.

[139] I. Z. Steinberg and E. Katchalski, "Theoretical analysis of the role of diffusion in chemical reactions, fluorescence quenching, and nonradiative energy transfer," *J. Chem. Phys.*, vol. 48, no. 6, pp. 2404–2410, 1968.

[140] I. Z. Steinberg, "Long-Range Nonradiative Transfer of Electronic Excitation Energy in Proteins and Polypeptides," *Annu. Rev. Biochem.*, vol. 40, no. 1, pp. 83–114, 1971.

[141] R. E. Dale, J. Eisinger, and W. E. Blumberg, "The orientational freedom of molecular probes. The orientation factor in intramolecular energy transfer," *Biophys. J.*, vol. 26, no. 2, pp. 161–193, 1979.

[142] L. M. S. Loura, "Simple estimation of förster resonance energy transfer (FRET) orientation factor distribution in membranes," *Int. J. Mol. Sci.*, vol. 13, no. 11, pp. 15252–15270, 2012.

[143] L. Stryer and R. P. Haugland, "Energy transfer: aspectroscopic ruler," *Proc. Natl. Acad. Sci. U. S. A.*, vol. 58, pp. 719–726, 1967.

[144] J. Baumann and M. D. Fayer, "Excitation transfer in disordered two-dimensional and anisotropic three-dimensional systems: Effects of spatial geometry on time-resolved observables," *J. Chem. Phys.*, vol. 85, no. 7, pp. 4087–4107, 1986.

[145] B. Schuler, E. A. Lipman, P. J. Steinbach, M. Kumkell, and W. A. Eaton, "Polyproline and the 'spectroscopic ruler' revisited with single-molecule flourescence," *Proc. Natl. Acad. Sci. U. S. A.*, vol. 102, no. 8, pp. 2754–2759, 2005.

[146] H. Langhals, A. J. Esterbauer, A. Walter, E. Riedle, and I. Pugliesi, "Förster resonant energy transfer in orthogonally arranged chromophores," *J. Am. Chem. Soc.*, vol. 132, no. 47, pp. 16777–16782, 2010.

[147] L. Stryer, "Fluorescence Energy Transfer as a Spectroscopic Ruler," *Annu. Rev. Biochem.*, vol. 47, no. 1, pp. 819–846, 1978.

[148] R. M. Clegg, "Fluorescence Resonance Energy Transfer Probes," *Compr. Biomed. Phys.*, vol. 6, pp. 103–110, 1995.

[149] M. Bates, T. R. Blosser, and X. Zhuang, "Short-range spectroscopic ruler based on a single-molecule optical switch," *Phys. Rev. Lett.*, vol. 94, no. 10, pp. 1–4, 2005.

[150] C. G. . Dos Remedios and P. D. J. Moens, "Fluorescence Resonance Energy Transfer Spectroscopy Is a Reliable 'Ruler' for Measuring Strucutal Changes in Proteins," *J. Struct. Biol.*, vol. 115, pp. 175–185, 1995.

[151] D. Amir and E. Haas, "Determination of intramolecular distance distributions in a globular protein by nonradiative excitation energy transfer measurements," *Biopolymers*, vol. 25, no. 2, pp. 235–240, 1986.

[152] S. Kalinin *et al.*, "A toolkit and benchmark study for FRET-restrained high-precision structural modeling," *Nat. Methods*, vol. 9, no. 12, pp. 1218–1225, 2012.

[153] Z. Gryczynski, T. Tenenholz, and E. Bucci, "Rates of energy transfer between tryptophans and hemes in hemoglobin, assuming that the heme is a planar oscillator," *Biophys. J.*, vol. 63, no. 3, pp. 648–653, 1992.

[154] A. Elofsson, R. Rigler, L. Nilsson, J. Roslund, G. Krause, and A. Holmgren, "Motion of Aromatic Side Chains , Picosecond Fluorescence , and Internal Energy in Escherichia coli Thioredoxin Studied by Site-Directed Mutagenesis, Time-resolved Fluorescence Spectroscopy, and Molecular Dynamics Simulations," *Am. Chem. Soc.*, vol. 30, no. 40, pp. 9648–9656, 1991.

[155] B. Hellenkamp *et al.*, "Precision and accuracy of single-molecule FRET measurements—a multi-laboratory benchmark study," *Nat. Methods*, vol. 15, no. 9, p. 669, 2018.

[156] T. O. Peulen, O. Opanasyuk, and C. A. M. Seidel, "Combining Graphical and Analytical Methods with Molecular Simulations to Analyze Time-Resolved FRET Measurements of Labeled Macromolecules Accurately," *J. Phys. Chem. B*, vol. 121, no. 35, pp. 8211–8241, 2017.

[157] J. J. McCann, U. B. Choi, L. Zheng, K. Weninger, and M. E. Bowen, "Optimizing methods to recover absolute FRET efficiency from immobilized single molecules," *Biophys. J.*, vol. 99, no. 3, pp. 961–970, 2010.

[158] J. R. Lakowicz, *Principles of fluorescence spectroscopy*. Springer, 2006.

[159] U. G. Zinth, "End-to-End Distance Distribution and Intra-Chain Diffusion in Unfolded Polypeptide Chains Determined by Time-Resolved FRET Measurements," 2013.

[160] P. Bernadó, L. Blanchard, P. Timmins, D. Marion, R. W. H. Ruigrok, and M. Blackledge, "A structural model for unfolded proteins from residual dipolar couplings and small-angle x-ray scattering," *Proc. Natl. Acad. Sci. U. S. A.*, vol. 102, no. 47, pp. 17002–17007, 2005.

[161] V. Ozenne *et al.*, "Flexible-meccano: A tool for the generation of explicit ensemble descriptions of intrinsically disordered proteins and their associated experimental observables," *Bioinformatics*, vol. 28, no. 11, pp. 1463–1470, 2012.

[162] M. Schwalbe *et al.*, "Predictive atomic resolution descriptions of intrinsically disordered hTau40 and α-synuclein in solution from NMR and small angle scattering," *Structure*, vol. 22, no. 2, pp. 238–249, 2014.

[163] A. Piai *et al.*, "Just a Flexible Linker? the Structural and Dynamic Properties of CBP-ID4 Revealed by NMR Spectroscopy," *Biophys. J.*, vol. 110, no. 2, pp. 372–381, 2016.

[164] R. Wetzel, "Physical chemistry of polyglutamine: Intriguing tales of a monotonous sequence," *J. Mol. Biol.*, vol. 421, no. 4–5, pp. 466–490, 2012.

[165] F. Saudou and S. Humbert, "The Biology of Huntingtin," *Neuron*, vol. 89, no. 5, pp. 910–926, 2016.

[166] G. P. Bates and C. Benn, "The polyglutamine diseases," *OXFORD Monogr. Med. Genet.*, vol. 45, no. 1, pp. 429–474, 2002.

[167] B. Wilburn *et al.*, "An Antisense CAG Repeat Transcript at JPH3 Locus Mediates Expanded Polyglutamine Protein Toxicity in Huntington's Disease-like 2 Mice," *Neuron*, 2011.

[168] C. Zuccato, M. Valenza, and E. Cattaneo, "Molecular Mechanisms and Potential Therapeutical Targets in Huntington's Disease," *Physiol. Rev.*, 2010.

[169] H. Okazawa, "Polyglutamine diseases: A transcription disorder?," *Cell. Mol. Life Sci.*, vol. 60, no. 7, pp. 1427–1439, 2003.

[170] J. Shao and M. I. Diamond, "Polyglutamine diseases: Emerging concepts in pathogenesis and therapy," *Hum. Mol. Genet.*, vol. 16, no. R2, pp. 115–123, 2007.

[171] C. A. Ross and S. J. Tabrizi, "Huntington's disease: From molecular pathogenesis to clinical treatment," *The Lancet Neurology*. 2011.

[172] B. Burnett, F. Li, and R. N. Pittman, "The polyglutamine neurodegenerative protein ataxin-3 binds polyubiquitylated proteins and has ubiquitin protease activity," *Hum. Mol. Genet.*, 2003.

[173] K. M. Donaldson, W. Li, K. A. Ching, S. Batalov, C.-C. Tsai, and C. A. P. Joazeiro, "Ubiquitin-mediated sequestration of normal cellular proteins into polyglutamine aggregates," *Proc. Natl. Acad. Sci.*, 2003.

[174] Y. Chai, S. S. Berke, R. E. Cohen, and H. L. Paulson, "Poly-ubiquitin Binding by the Polyglutamine Disease Protein Ataxin-3 Links Its Normal Function to Protein Surveillance Pathways," *J. Biol. Chem.*, 2004.

[175] B. O. Evert *et al.*, "Ataxin-3 Represses Transcription via Chromatin Binding, Interaction with Histone Deacetylase 3, and Histone Deacetylation," *J. Neurosci.*, 2006.

[176] B. G. Burnett and R. N. Pittman, "The polyglutamine neurodegenerative protein ataxin 3 regulates aggresome formation," *Proc. Natl. Acad. Sci.*, 2005.

[177] J. M. Warrick *et al.*, "Ataxin-3 suppresses polyglutamine neurodegeneration in Drosophila by a ubiquitin-associated mechanism," *Mol. Cell*, 2005.

[178] F. Li, T. Macfarlan, R. N. Pittman, and D. Chakravarti, "Ataxin-3 is a histone-binding protein with two independent transcriptional corepressor activities," *J. Biol. Chem.*, 2002.

[179] P. D. Knight, T. K. Karamanos, S. E. Radford, and A. E. Ashcroft, "Identification of a novel site of interaction between ataxin-3 and the amyloid aggregation inhibitor polyglutamine binding peptide 1," *Eur. J. Mass Spectrom.*, vol. 24, no. 1, pp. 129–140, 2017.

[180] S. L. Crick, M. Jayaraman, C. Frieden, R. Wetzel, and R. V. Pappu, "Fluorescence correlation spectroscopy shows that monomeric polyglutamine molecules form collapsed structures in aqueous solutions," *Proc. Natl. Acad. Sci.*, vol. 103, no. 45, pp. 16764–16769, 2006.

[181] Y. Nagai *et al.*, "A toxic monomeric conformer of the polyglutamine protein," *Nat. Struct. Mol. Biol.*, vol. 14, no. 4, pp. 332–340, 2007.

[182] F. Ferrone, "Analysis of protein aggregation kinetics," *Methods Enzymol.*, 1999.

[183] K. Kar, M. Jayaraman, B. Sahoo, R. Kodali, and R. Wetzel, "Critical nucleus size for disease-related polyglutamine aggregation is repeat-length dependent," *Nat. Struct. Mol. Biol.*, vol. 18, no. 3, pp. 328–336, 2011.

[184] E. L. Altschuler, N. V. Hud, J. A. Mazrimas, and B. Rupp, "Random coil conformation for extended polyglutamine stretches in aqueous soluble monomeric peptides," *J. Pept. Res.*, 2009.

[185] S. Chen, V. Berthelier, W. Yang, and R. Wetzel, "Polyglutamine aggregation behavior in vitro supports a recruitment mechanism of cytotoxicity," *J. Mol. Biol.*, vol. 311, no. 1, pp. 173–182, 2001.

[186] L. Masino, G. Kelly, K. Leonard, Y. Trottier, and A. Pastore, "Solution structure of polyglutamine tracts in GST-polyglutamine fusion proteins," *FEBS Lett.*, vol. 513, no. 2–3, pp. 267–272, 2002.

[187] L. Dougan, J. Li, C. L. Badilla, B. J. Berne, and J. M. Fernandez, "Single homopolypeptide chains collapse into mechanically rigid conformations," *Proc. Natl. Acad. Sci.*, 2009.

[188] C. Peters-Libeu *et al.*, "Disease-associated polyglutamine stretches in monomeric huntingtin adopt a compact structure," *J. Mol. Biol.*, vol. 421, no. 4–5, pp. 587–600, 2012.

[189] M. Karplus, "The Levinthal paradox: Yesterday and today," *Fold. Des.*, vol. 2, no. 4, pp. 69–75, 1997.

[190] K. W. Plaxco and D. Baker, "Limited internal friction in the rate-limiting step of a two-state protein folding reaction," *Proc. Natl. Acad. Sci. U. S. A.*, vol. 95, no. 23, pp. 13591–13596, 1998.

[191] D. K. Klimov and D. Thirumalai, "Viscosity dependence of the folding rates of proteins," *Phys. Rev. Lett.*, vol. 79, no. 2, pp. 317–320, 1997.

[192] S. Takada and P. G. Wolynes, "Microscopic theory of critical folding nuclei and reconfiguration activation barriers in folding proteins," *J. Chem. Phys.*, vol. 107, no. 22, pp. 9585–9598, 1997.

[193] S. A. Pabit, H. Roder, and S. J. Hagen, "Internal friction controls the speed of protein folding from a compact configuration," *Biochemistry*, vol. 43, no. 39, pp. 12532–12538, 2004.

[194] S. J. Hagen, "Solvent Viscosity and Friction in Protein Folding Dynamics," *Curr. Protein Pept. Sci.*, vol. 999, no. 999, pp. 1–11, 2010.

[195] S. S. Plotkin and J. N. Onuchic, "Understanding protein folding with energy landscape theory part I: basic concepts," *Q. Rev. Biophys.*, vol. 35, no. 2, pp. 111–167, 2002.

[196] S. S. Plotkin and J. N. Onuchic, “Understanding protein folding with energy landscape theory Part II: Quantitative aspects,” *Q. Rev. Biophys.*, vol. 35, no. 3, pp. 205–286, 2002.

[197] B. A. Chrunyk and C. R. Matthews, “Role of Diffusion in the Folding of the α Subunit of Tryptophan Synthase from Escherichia coli,” *Biochemistry*, vol. 29, no. 8, pp. 2149–2154, 1990.

[198] M. Jacob, T. Schindler, J. Balbach, and F. X. Schmid, “Diffusion control in an elementary protein folding reaction,” *Proc. Natl. Acad. Sci. U. S. A.*, vol. 94, no. 11, pp. 5622–5627, 1997.

[199] M. Jacob, M. Geeves, G. Holtermann, and F. X. Schmid, “Diffusional barrier crossing in a two-state protein folding reaction,” *Nat. Struct. Biol.*, vol. 6, no. 10, pp. 923–926, 1999.

[200] R. P. Bhattacharyya and T. R. Sosnick, “Viscosity dependence of the folding kinetics of a dimeric and monomeric coiled coil,” *Biochemistry*, vol. 38, no. 8, pp. 2601–2609, 1999.

[201] O. Bilsel and R. Matthews, “folding of the polypeptide . Third , the dramatic increase in the size of reaction and the well-known hydrophobic effect arising from the in- native state and exposed when unfolded imply that entropy is a very nificantly over the last five years . These c,” *Rev. Lit. Arts Am.*, vol. 53, 2000.

[202] H. A. Kramers, “Brownian motion in a field of force and the diffusion model of chemical reactions,” *Physica*, vol. 7, no. 4, pp. 284–304, 1940.

[203] H. S. Chung and W. A. Eaton, “Single-molecule fluorescence probes dynamics of barrier crossing,” *Nature*, vol. 502, no. 7473, pp. 685–688, 2013.

[204] L. Qiu and S. J. Hagen, “A Limiting Speed for Protein Folding at Low Solvent Viscosity,” *J. Am. Chem. Soc.*, vol. 126, no. 11, pp. 3398–3399, 2004.

[205] G. S. Jas, W. A. Eaton, and J. Hofrichter, “Effect of viscosity on the kinetics of α-helix and β-hairpin formation,” *J. Phys. Chem. B*, vol. 105, no. 1, pp. 261–272, 2001.

[206] R. B. Best and G. Hummer, “Diffusive model of protein folding dynamics with Kramers turnover in rate,” *Phys. Rev. Lett.*, vol. 96, no. 22, pp. 2–5, 2006.

[207] B. Zagrovic and V. Pande, “Solvent viscosity dependence of the folding rate of a small protein: Distributed computing study,” *J. Comput. Chem.*, vol. 24, no. 12, pp. 1432–1436, 2003.

[208] R. Narayanan, L. Pelakh, and S. J. Hagen, “Solvent Friction Changes the Folding Pathway of the Tryptophan Zipper TZ2,” *J. Mol. Biol.*, vol. 390, no. 3, pp. 538–546, 2009.

[209] H. S. Chung, S. Piana-Agostinetti, D. E. Shaw, and W. A. Eaton, “Structural origin of slow diffusion in protein folding,” *Science (80-.).*, vol. 349, no. 6255, pp. 1504–1510, 2015.

[210] R. Cerf, “On the stochastic approach to irreversible statistical mechanics of chain macromolecules,” *Chem. Phys. Lett.*, vol. 16, no. 1, pp. 42–44, 1972.

[211] R. Cerf, “Possible existence of two classes of conformational changes in linear polymeric chain molecules,” *Chem. Phys. Lett.*, vol. 22, no. 3, pp. 613–615, 1973.

[212] J. J. Portman, S. Takada, and P. G. Wolynes, “Microscopic Theory of Protein Folding Rates.II: Local Reaction Coordinates and Chain Dynamics,” *J. Chem. Phys.*, vol. 114, no. 5082, pp. 1–18, 2001.

[213] A. Erbaş, D. Horinek, and R. R. Netz, “Viscous friction of hydrogen-bonded matter,” *J. Am. Chem. Soc.*, vol. 134, no. 1, pp. 623–630, 2012.

[214] P.-G. De Gennes, *Scaling concepts in polymer physics*. Cornell University Press, 1979.

[215] D. De Sancho, A. Sirur, and R. B. Best, "Molecular origins of internal friction effects on protein-folding rates," *Nat. Commun.*, vol. 5, pp. 1–10, 2014.

[216] J. C. F. Schulz, M. S. Miettinen, and R. R. Netz, "Unfolding and Folding Internal Friction of β-Hairpins Is Smaller than That of α-Helices," *J. Phys. Chem. B*, vol. 119, no. 13, pp. 4565–4574, 2015.

[217] F. Krieger, "Dynamics in Unfolded Polypeptide Chains as Model for Elementary Steps in Protein Folding," Universität Basel, 2004.

[218] K. N. Joder, "Intramolecular and intermolecular diffusion processes in protein folding and assembly," 2011.

[219] a. Einstein, "Über die von der molekularkinetischen Theorie der Wärme geforderte Bewegumg von in ruhenden Flüssigkeiten suspendierten Teilchen," *Ann. d. Phys.*, vol. 322, no. 8, pp. 549–560, 1905.

[220] A. Einstein, *Investigations on the Theory of the Brownian Movement*. Courier Corporation, 1956.

[221] M. von Smoluchowski, "Zur kinetischen Theorie der Brownschen Molekularbewegung und Suspensionen," *Ann. d. Phys.*, vol. 60, no. 19, pp. 549–560, 1906.

[222] J. A. Dix and A. S. Verkman, "Crowding Effects on Diffusion in Solutions and Cells," *Annu. Rev. Biophys.*, vol. 37, no. 1, pp. 247–263, 2008.

[223] Y. Wang, C. Li, and G. J. Pielak, "Effects of proteins on protein diffusion," *J. Am. Chem. Soc.*, vol. 132, no. 27, pp. 9392–9397, 2010.

[224] P. hung Wang, I. Yu, M. Feig, and Y. Sugita, "Influence of protein crowder size on hydration structure and dynamics in macromolecular crowding," *Chem. Phys. Lett.*, vol. 671, pp. 63–70, 2017.

[225] A. B. Goins, H. Sanabria, and M. N. Waxham, "Macromolecular crowding and size effects on probe microviscosity," *Biophys. J.*, vol. 95, no. 11, pp. 5362–5373, 2008.

[226] A. Soranno *et al.*, "Single-molecule spectroscopy reveals polymer effects of disordered proteins in crowded environments," *Proc. Natl. Acad. Sci. U. S. A.*, vol. 111, no. 13, pp. 4874–4879, 2014.

[227] N. Kozer and G. Schreiber, "Effect of Crowding on Protein-Protein Association Rates: Fundamental Differences between Low and High Mass Crowding Agents," *J. Mol. Biol.*, vol. 336, no. 3, pp. 763–774, 2004.

[228] Y. Y. Kuttner, N. Kozer, E. Segal, G. Schreiber, and G. Haran, "Separating the contribution of translational and rotational diffusion to protein association," *J. Am. Chem. Soc.*, vol. 127, no. 43, pp. 15138–15144, 2005.

[229] D. C. Latshaw II and C. K. Hall, "Aggregation," *Biophys. J.*, vol. 109, no. July, pp. 124–134, 2015.

[230] K. Sozański, F. Ruhnow, A. Wišniewska, M. Tabaka, S. Diez, and R. Hołyst, "Small Crowders Slow Down Kinesin-1 Stepping by Hindering Motor Domain Diffusion," *Phys. Rev. Lett.*, vol. 115, no. 21, pp. 1–5, 2015.

[231] W. Kuhn and H. Kuhn, “Bedeutung beschränkt freier Drehbarkeit für die Viskosität und Strömungsdoppelbrechung von Fadenmolekellösungen II,” *Helv. Chim. Acta*, vol. 29, no. 71, p. 94, 1946.

[232] B. S. Khatri and T. C. B. McLeish, “Rouse model with internal friction: A coarse grained framework for single biopolymer dynamics,” *Macromolecules*, vol. 40, no. 18, pp. 6770–6777, 2007.

[233] R. R. Cheng, A. T. Hawk, and D. E. Makarov, “Exploring the role of internal friction in the dynamics of unfolded proteins using simple polymer models,” *J. Chem. Phys.*, vol. 138, no. 7, 2013.

[234] W. Zheng, D. De Sancho, and R. B. Best, “Modulation of Folding Internal Friction by Local and Global Barrier Heights,” *J. Phys. Chem. Lett.*, vol. 7, no. 6, pp. 1028–1034, 2016.

[235] S. P. Velsko and G. R. Fleming, “Photochemical isomerization in solution. Photophysics of diphenyl butadiene,” *J. Chem. Phys.*, vol. 76, no. 7, pp. 3553–3562, 1982.

[236] S. P. Velsko, D. H. Waldeck, and G. R. Fleming, “Breakdown of Kramers theory description of photochemical isomerization and the possible involvement of frequency dependent friction,” *J. Chem. Phys.*, vol. 78, no. 1, pp. 249–258, 1983.

[237] J. C. F. Schulz, L. Schmidt, R. B. Best, J. Dzubiella, and R. R. Netz, “Peptide chain dynamics in light and heavy water: Zooming in on internal friction,” *J. Am. Chem. Soc.*, vol. 134, no. 14, pp. 6273–6279, 2012.

[238] P. W. Fenimore, H. Frauenfelder, B. H. McMahon, and F. G. Parak, “Slaving: Solvent fluctuations dominate protein dynamics and functions,” *Proc. Natl. Acad. Sci. U. S. A.*, vol. 99, no. 25, pp. 16047–16051, 2002.

[239] P. W. Fenimore, H. Frauenfelder, B. H. McMahon, and R. D. Young, “Bulk-solvent and hydration-shell fluctuations, similar to α- and β-fluctuations in glasses, control protein motions and functions,” *Proc. Natl. Acad. Sci. U. S. A.*, vol. 101, no. 40, pp. 14408–14413, 2004.

[240] R. F. Grote and J. T. Hynes, “The stable states picture of chemical reactions. II. Rate constants for condensed and gas phase reaction models,” *J. Chem. Phys.*, vol. 73, no. 6, pp. 2715–2732, 1980.

[241] B. Bagchi and D. W. Oxtoby, “The effect of frequency dependent friction on isomerization dynamics in solution,” *J. Chem. Phys.*, vol. 78, no. 5, pp. 2735–2741, 1983.

[242] G. Barshtein, A. Almagor, S. Yedgar, and B. Gavish, “Inhomogeneity of viscous aqueous solutions,” *Phys. Rev. E*, vol. 52, no. 1, pp. 555–557, 1995.

[243] S. Yedgar, C. Tetreau, B. Gavish, and D. Lavalette, “Viscosity dependence of O2 escape from respiratory proteins as a function of cosolvent molecular weight,” *Biophys. J.*, vol. 68, no. 2, pp. 665–670, 1995.

[244] R. H. Walters and R. M. Murphy, “Examining Polyglutamine Peptide Length: A Connection between Collapsed Conformations and Increased Aggregation,” *J. Mol. Biol.*, vol. 393, no. 4, pp. 978–992, 2009.

[245] R. Laghaei and N. Mousseau, “Spontaneous formation of polyglutamine nanotubes with molecular dynamics simulations,” *J. Chem. Phys.*, vol. 132, no. 16, 2010.

[246] B. O'Nuallain *et al.*, "Kinetics and Thermodynamics of Amyloid Assembly Using a High-Performance Liquid Chromatography-Based Sedimentation Assay," *Methods Enzymol.*, vol. 413, no. 06, pp. 34–74, 2006.

[247] M. L. Tiffany and S. Krimm, "Extended conformations of polypeptides and proteins in urea and guanidine hydrochloride," *Biopolymers*, vol. 12, no. 3, pp. 575–587, 1973.

[248] R. W. Woody, "Circular Dichroism," *Methods Enzymol.*, vol. 246, pp. 34–71, 1995.

[249] W. C. Johnson, "Protein secondary structure and circular dichroism: A practical guide," *Proteins Struct. Funct. Bioinforma.*, vol. 7, no. 3, pp. 205–214, 1990.

[250] N. Greenfield and G. D. Fasman, "Computed Circular Dichroism Spectra for the Evaluation of Protein Conformation," *Biochemistry*, vol. 8, no. 10, pp. 4108–4116, 1969.

[251] A. J. Adler, R. Hoving, J. Potter, M. Wells, and G. D. Fasman, "Circular Dichroism of Polypeptides. Poly(hydroxyethyl-L-glutamine) compared to poly(L-glutamic acid)," *J. Am. Chem. Soc.*, vol. 90, pp. 4736–4738, 1968.

[252] V. N. Uversky *et al.*, "Natively unfolded human prothymosin α adopts partially folded collapsed conformation at acidic pH," *Biochemistry*, vol. 38, no. 45, pp. 15009–15016, 1999.

[253] S. M. Kelly and N. C. Price, "The application of circular dichroism to studies of protein folding and unfolding," vol. 1338, pp. 161–185, 1997.

[254] E. W. Ronish and S. Krimm, "Theoretical calculation of the circular dichroism of unordered polypeptide chains," *Biopolymers*, vol. 11, no. 9, pp. 1919–1928, 1972.

[255] S. Brahms and J. Brahms, "Determination of protein secondary structure in solution by vacuum ultraviolet circular dichroism," *J. Mol. Biol.*, vol. 138, no. 2, pp. 149–178, 1980.

[256] R. W. Woody, *Circular dichroism of intrinsically disordered proteins*. John Wiley & Sons, Hoboken, NJ, 2010.

[257] S. Chen, F. A. Ferrone, and R. Wetzel, "Huntington's disease age-of-onset linked to polyglutamine aggregation nucleation," *Proc. Natl. Acad. Sci. U. S. A.*, vol. 99, no. 18, pp. 11884–11889, 2002.

[258] J. Wägele, S. De Sio, B. Voigt, J. Balbach, and M. Ott, "How Fluorescent Tags Modify Oligomer Size Distributions of the Alzheimer Peptide," *Biophys. J.*, vol. 116, no. 2, pp. 227–238, 2019.

[259] S. Albaugh and R. F. Steiner, "Determination of distance distribution from time domain fluorometry," *J. Phys. Chem.*, vol. 93, no. 24, pp. 8013–8016, 1989.

[260] J. L. Digambaranath *et al.*, "An accurate model of polyglutamine," *Proteins Struct. Funct. Bioinforma.*, vol. 79, no. 5, pp. 1427–1440, 2011.

[261] A. Szabo, K. Schulten, and Z. Schulten, "First passage time approach to diffusion controlled reactions," *J. Chem. Phys.*, vol. 72, no. 8, pp. 4350–4357, 2003.

[262] B. Yawen and N. Ruth, *Protein Folding Protocols*. 2006.

[263] C. Hyeon and D. Thirumalai, "Can energy landscape roughness of proteins and RNA be measured by using mechanical unfolding experiments?," *Proc. Natl. Acad. Sci. U. S. A.*, vol. 100, no. 18, pp. 10249–10253, 2003.

[264] L. Milanesi *et al.*, “Measurement of energy landscape roughness of folded and unfolded proteins,” *Proc. Natl. Acad. Sci. U. S. A.*, vol. 109, no. 48, pp. 19563–19568, 2012.

[265] J. A. Ihalainen *et al.*, “Folding and unfolding of a photoswitchable peptide from picoseconds to microseconds,” *Proc. Natl. Acad. Sci. U. S. A.*, vol. 104, no. 13, pp. 5383–5388, 2007.

[266] R. Nevo, V. Brumfeld, R. Kapon, P. Hinterdorfer, and Z. Reich, “Direct measurement of protein energy landscape roughness,” *EMBO Rep.*, vol. 6, no. 5, pp. 482–486, 2005.

[267] D. Nettels, I. V. Gopich, A. Hoffmann, and B. Schuler, “Ultrafast dynamics of protein collapse from single-molecule photon statistics,” *Proc. Natl. Acad. Sci. U. S. A.*, vol. 104, no. 8, pp. 2655–2660, 2007.

[268] S. A. Waldauer *et al.*, “Ruggedness in the folding landscape of protein L,” *HFSP J.*, vol. 2, no. 6, pp. 388–395, 2008.

[269] F. Krieger, A. Möglich, and T. Kiefhaber, “Effect of proline and glycine residues on dynamics and barriers of loop formation in polypeptide chains,” *J. Am. Chem. Soc.*, vol. 127, no. 10, pp. 3346–3352, 2005.

[270] D. Roccatano, H. Sahoo, M. Zacharias, and W. M. Nau, “Temperature dependence of looping rates in a short peptide,” *J. Phys. Chem. B*, vol. 111, no. 10, pp. 2639–2646, 2007.

[271] S. J. Hagen and W. A. Eaton, “Two-state expansion and collapse of a polypeptide 1 1Edited by A. R. Fersht,” *J. Mol. Biol.*, vol. 301, no. 4, pp. 1019–1027, 2000.

[272] A. Abyzov *et al.*, “Identification of Dynamic Modes in an Intrinsically Disordered Protein Using Temperature-Dependent NMR Relaxation,” *J. Am. Chem. Soc.*, vol. 138, no. 19, pp. 6240–6251, 2016.

[273] E. Pollak, A. Auerbach, and P. Talknerz, “Observations on rate theory for rugged energy landscapes,” *Biophys. J.*, vol. 95, no. 9, pp. 4258–4265, 2008.

[274] M. Volk, L. Milanesi, J. P. Waltho, C. A. Hunter, and G. S. Beddard, “The roughness of the protein energy landscape results in anomalous diffusion of the polypeptide backbone,” *Phys. Chem. Chem. Phys.*, vol. 17, no. 2, pp. 762–782, 2015.

[275] G. Scherer, M. L. Kramer, M. Schutkowski, U. Reimer, and G. Fischer, “Barriers to rotation of secondary amide peptide bonds,” *J. Am. Chem. Soc.*, vol. 120, no. 22, pp. 5568–5574, 1998.

[276] R. Zwanzig, “~ ~ Y,” vol. 85, no. April, pp. 2029–2030, 1988.

[277] J. D. Bryngelson and P. G. Wolynes, “Intermediates and barrier crossing in a random energy model (with applications to protein folding),” *J. Phys. Chem.*, vol. 93, no. 19, pp. 6902–6915, 1989.

[278] H. A. Scheraga, “Calculations of Conformations of Polypeptides,” *Adv. Phys. Org. Chem.*, vol. 6, no. C, pp. 103–184, 1968.

[279] G. N. T. Ramachandran and V. Sasisekharan, “Conformation of polypeptides and proteins,” in *Advances in protein chemistry*, vol. 23, Elsevier, 1968, pp. 283–437.

[280] A. Möglich, K. Joder, and T. Kiefhaber, “End-to-end distance distributions and intrachain diffusion constants in unfolded polypeptide chains indicate intramolecular hydrogen bond formation,” *Proc. Natl. Acad. Sci. U. S. A.*, vol. 103, no. 33, pp. 12394–12399, 2006.

[281] A. Tischer, H. Lilie, M. Auton, and C. Lange, “Oxidative refolding of rPA in l -ArgHCl and in ionic liquids: A correlation between hydrophobicity, salt effects, and refolding yield,” *Biopolymers*, vol. 101, no. 11, pp. 1129–1140, 2014.

[282] N. Carolina, “of Amino Guanidine - I,” *J. Biol. Chem.*, no. 7, pp. 1648–1652, 1970.

[283] E. P. O’Brien, R. I. Dima, B. Brooks, and D. Thirumalai, “Interactions between hydrophobic and ionic solutes in aqueous guanidinium chloride and urea solutions: Lessons for protein denaturation mechanism,” *J. Am. Chem. Soc.*, vol. 129, no. 23, pp. 7346–7353, 2007.

[284] D. R. Robinson and W. P. Jencks, “The Effect of Compounds of the Urea-Guanidinium Class on the Activity Coefficient of Acetyltetraglycine Ethyl Ester and Related Compounds,” *J. Am. Chem. Soc.*, vol. 87, no. 11, pp. 2462–2470, 1965.

[285] P. E. Mason, J. W. Brady, G. W. Neilson, and C. E. Dempsey, “The interaction of guanidinium ions with a model peptide,” *Biophys. J.*, vol. 93, no. 1, pp. 4–6, 2007.

[286] G. I. Makhatadze and P. L. Privalov, “Protein interactions with urea and guanidinium chloride,” *J. Mol. Biol.*, vol. 226, no. 2, pp. 491–505, 2004.

[287] P. E. Mason *et al.*, “The structure of aqueous guanidinium chloride solutions,” *J. Am. Chem. Soc.*, vol. 126, no. 37, pp. 11462–11470, 2004.

[288] P. H. von Hippel and T. Schleich, “Ion Effects on the Solution Structure of Biological Macromolecules,” *Acc. Chem. Res.*, vol. 2, no. 9, pp. 257–265, 1969.

[289] J. A. Schellman, “Solvent Denaturation*,” vol. 17, no. 2, pp. 1305–1322, 1978.

[290] J. A. Schellman, “Selective Binding and Solvent Denaturation,” *Biopolymers*, vol. 26, pp. 549–559, 1987.

[291] J. A. Schellman, “Protein stability in mixed solvents: A balance of contact interaction and excluded volume,” *Biophys. J.*, vol. 85, no. 1, pp. 108–125, 2003.

[292] A. Böker, “Simulations of protein thermodynamics and structures with the PRIME20 model Dissertation,” Martin-Luther-Universität Halle-Wittenberg, 2019.

[293] D. Punihaole, R. S. Jakubek, R. J. Workman, and S. A. Asher, “Interaction Enthalpy of Side Chain and Backbone Amides in Polyglutamine Solution Monomers and Fibrils,” *J. Phys. Chem. Lett.*, vol. 9, no. 8, pp. 1944–1950, 2018.

[294] M. Nakano, H. Watanabe, S. M. Rothstein, and S. Tanaka, “Erratum: Comparative characterization of short monomeric polyglutamine peptides by replica exchange molecular dynamics simulation (Journal of Physical Chemistry B (2010) 114 (7056)),” *J. Phys. Chem. B*, vol. 114, no. 31, p. 10234, 2010.

[295] S. Pylaeva, A. Böker, H. Elgabarty, W. Paul, and D. Sebastiani, “The Conformational Ensemble of Polyglutamine-14 Chains: Specific Influences of Solubility Tail and Chromophores,” *ChemPhysChem*, vol. 19, no. 21, pp. 2931–2937, 2018.

[296] X. Wang, A. Vitalis, M. A. Wyczalkowski, and R. V. Pappu, “Characterizing the conformational ensemble of monomeric polyglutamine,” *Proteins Struct. Funct. Genet.*, vol. 63, no. 2, pp. 297–311, 2006.

[297] D. Roy and J. J. Dannenberg, “The Effects of Regularly Spaced Glutamine Substitutions on Alpha-Helical Peptide Structures. A DFT/ONIOM Study,” *Chem. Phys. Lett.*, vol. 512, pp. 255–257, 2011.

[298] V. A. Zhemkov, A. A. Kulminskaya, I. B. Bezprozvanny, and M. Kim, “The 2.2-Angstrom resolution crystal structure of the carboxy-terminal region of ataxin-3,” *FEBS Open Bio*, vol. 6, no. 3, pp. 168–178, 2016.

[299] A. Escobedo *et al.*, “Side chain to main chain hydrogen bonds stabilize a polyglutamine helix in a transcription factor,” *Nat. Commun.*, vol. 10, no. 1, p. 2034, 2019.

[300] M. F. Perutz, J. T. Finch, J. Berriman, and A. Lesk, “Amyloid fibers are water-filled nanotubes,” *Proc. Natl. Acad. Sci. U. S. A.*, vol. 99, no. 8, pp. 5591–5595, 2002.

[301] J. A. Plumley and J. J. Dannenberg, “The importance of hydrogen bonding between the glutamine side chains to the formation of amyloid VQIVYK parallel β-sheets: An ONIOM DFT/AM1 study,” *J. Am. Chem. Soc.*, vol. 132, no. 6, pp. 1758–1759, 2010.

[302] P. Sikorski and E. Atkins, “New model for crystalline polyglutamine assemblies and their connection with amyloid fibrils,” *Biomacromolecules*, vol. 6, no. 1, pp. 425–432, 2005.

[303] C. L. Hoop *et al.*, “Huntingtin exon 1 fibrils feature an interdigitated β-hairpin-based polyglutamine core,” *Proc. Natl. Acad. Sci. U. S. A.*, vol. 113, no. 6, pp. 1546–1551, 2016.

[304] L. L. Porter and G. D. Rose, “Redrawing the Ramachandran plot after inclusion of hydrogen-bonding constraints,” *Proc. Natl. Acad. Sci. U. S. A.*, vol. 108, no. 1, pp. 109–113, 2011.

[305] G. Weber, “Polarization of the fluorescence of macromolecules. 2. Fluorescent conjugates of ovalbumin and bovine serum albumin,” *Biochem. J.*, vol. 51, no. 2, pp. 155–167, 1952.

[306] C. G. Knight, F. Willenbrock, and G. Murphy, “A novel coumarin-labelled peptide for sensitive continuous assays of the matrix metalloproteinases,” *FEBS Lett.*, vol. 296, no. 3, pp. 263–266, 1992.

[307] N. Boens *et al.*, “Fluorescence lifetime standards for time and frequency domain fluorescence spectroscopy,” *Anal. Chem.*, vol. 79, no. 5, pp. 2137–2149, 2007.

[308] I. Berlman, *Handbook of florescence spectra of aromatic molecules*. Elsevier, 1971.

[309] I. Z. Steinberg, “Brownain Motion of End-to-end Distance in Oligopeptide Molecules: Numerical Solution of the Diffusion Equations as Coupled First Order Linear Differential Equations,” *J. theor. Biol.*, pp. 173–187, 1994.

[310] M. L. Johnson and L. M. Faont, “[1] Parameter Estimation by Least-Square Methods,” *Methods Enzymol.*, vol. 210, pp. 1–37, 1992.

[311] C. E. Blanchet *et al.*, “Versatile sample environments and automation for biological solution X-ray scattering experiments at the P12 beamline (PETRA III, DESY),” *J. Appl. Crystallogr.*, vol. 48, pp. 431–443, 2015.

[312] A. G. Kikhney and D. I. Svergun, “A practical guide to small angle X-ray scattering (SAXS) of flexible and intrinsically disordered proteins,” *FEBS Lett.*, vol. 589, no. 19, pp. 2570–2577, 2015.

[313] P. V. Konarev, V. V. Volkov, A. V. Sokolova, M. H. J. Koch, and D. I. Svergun, “PRIMUS: A Windows PC-based system for small-angle scattering data analysis,” *J. Appl. Crystallogr.*, vol. 36, no. 5, pp. 1277–1282, 2003.

[314] M. V Petoukhov, P. V Konarev, G. Kikhney, and I. Dmitri, “ATSAS 2.1 - Supported Small-Angle Scattering Data Analysis,” *Appl. Crystallogr.*, vol. 40, pp. s223–s228, 2007.

[315] D. Franke, A. G. Kikhney, and D. I. Svergun, “Automated acquisition and analysis of small angle X-ray scattering data,” *Nucl. Instruments Methods Phys. Res. Sect. A Accel. Spectrometers, Detect. Assoc. Equip.*, vol. 689, pp. 52–59, 2012.

[316] T. Freltoft, J. K. Kjems, and S. K. Sinha, “Power-law correlations and finite-size effects in silica particle aggregates studied by small-angle neutron scattering,” *Phys. Rev. B*, vol. 33, no. 1, pp. 269–275, 1986.

[317] R. Graham and J. R. Lewis, “J.C.S. Pekin I,” pp. 876–881, 1978.

[318] P. Enke, “Beiträge zur Reinigung und biophysikalischen Charakterisierung der zellulären Proteine NF45 und NF90,” Martin-Luther Universität Halle-Wittenberg, 2011.

[319] I. Z. Steinberg, "Brownian Motion of the End-to-end Distnace in Oligopeptide Molecules: Numerical Solution of the Diffusion Equations Coupled First Order Linear Differential Equations," *J. theor. Biol.*, vol. 166, pp. 173-187, 1994.

Publikationsliste/Poster/Vorträge

Poster

Peter Enke, Svetlana Pylaeva, Daniel Sebastiani, Arne Böker, Wolfgang Paul, Thomas Kiefhaber, "*Dynamic and Structural Properties of Polyglutamine*" 29th Faltertage, Halle 2018

Peter Enke, Michael Schleeger, Thomas Kiefhaber, "*Comparing Structural and Kinetic Properties of Q_{14} and S_{14}*", 5th Minisymposium SFB/TRR102, Leipzig 2018

Peter Enke, Michael Schleeger, Thomas Kiefhaber, "*Dynamic and Structural Properties of Polyglutamine*", IDMPC SFB/TRR102, Wittenberg 2017

Peter Enke, Michael Schleeger, Thomas Kiefhaber, "*Dynamic and Structural Properties of Polyglutamine*", 4th Minisymposium SFB/TRR102, Halle 2017

Peter Enke, Michael Schleeger, Stefan Wicht, Thomas Kiefhaber, "*Dynamic and Structural Properties of Amyloid-Forming and Intrinsically Disordered Polypeptide Chains*", 1st Minisymposium SFB/TRR102, Leipzig 2015

Peter Enke, Ursula Zinth, Thomas Kiefhaber, "*Effect of Co-Solutes on Dimensions and Intrachain Dynamics in Unfolded Polypeptide Chains*", 26th Faltertage, Halle 2015

Peter Enke, Thomas Kiefhaber, "*Dynamic and Structural Properties of Amyloid Forming and Intrinsically Disordered Polypeptide Chains*", SFB/TRR102, Leipzig 2014

Vorträge

"*Dynamic and Structural Properties of $(Glutamine)_{14}$*" iRTG meeting SFB/TRR102, Halle 2018

"*time-resolved ensemble FRET*", 3rd iRTG Workshop SFB/TRR102, Wittenberg 2017

"*Effect of Different Crowding Agents on Structure and Dynamics of Unfolded Proteins*", iRTG meeting SFB/TRR102, Halle 2017

"*Effect of Different Crowding Agents on Structure and Dynamics of Unfolded Proteins*", 27th Faltertage, Halle 2016

"*Dimension and Dynamics of Unfolded Polypeptide Chains measured by time-resolved ensemble FRET*", iRTG Workshop SFB/TRR102, Wittenberg 2015

Angaben zur Person

Persönliche Daten

Familienname	Enke
Vorname	Peter
Geburtstag	26.04.1985
Geburtsort	Plauen (Vogtl.)
Nationalität	Deutsch
Anschrift	Wipperweg 9, 06122 Halle (Saale), Deutschland

Bildungsgang

08/1997 – 08/2003	Mittelschule Jocketa, Realschule, technisches Profil
09/2003 – 07.2005	Ausbildung zum "Staatlich geprüften Technischen Assitent für chemische und biologische Laboratorien" an der Berufsfachschule für Technik an dem Bildungsinstitut PSCHER gGmbH
08/2005 – 07.2008	Berufliches Gymnasium, IWB (Institut für Wissen und Bildung) Plauen, Abitur
10/2008 – 09/2011	Bachelorstudium Biochemie, Martin-Luther-Universität Halle-Wittenberg Bachelorarbeit: "*Beiträge zur Reinigung und biophysikalischen Charakterisierung der zellulären Proteine NF45 und NF90*" in der Arbeitsgruppe von Prof. Dr. Sven-Erik Behrens
10/2011 – 10.2013	Masterstudium Biochemie, Martin-Luther-Universität Halle-Wittenberg Masterarbeit: "*Baculovirale Expression von Me31B und assoziierten Proteinen*"

Promotion

12/2013 – 11/2019	Thema "*Dynamic and Structural Properties of Amyloid Forming and Intrinsically Disordered Polypeptide Chains Determined by Time-Resolved FRET Measurements.*" in der Arbeitsgruppe von Prof. Dr. Thomas Kiefhaber, Institut für Physikalische Chemie an der Technische Universität München und Institut der Biochemie und Biotechnologie an der Martin-Luther-Universität Halle-Wittenberg

Eidesstattliche Erklärung

Ich erkläre an Eides statt, dass ich die Arbeit selbstständig und ohne fremde Hilfe verfasst, keine anderen als die von mir angegebenen Quellen und Hilfsmittel benutzt und die den benutzten Werken wörtlich oder inhaltlich entnommenen Stellen als solche kenntlich gemacht habe.

Halle (Saale), den _____2020 Peter Enke

Acknowledgment

The presented work was carried out from December 2013 until Februar 2020 in the lab from Prof. Dr. Thomas Kiefhaber at the Department of biophysical chemistry TUM in Garching and later at the Institute of Biochemistry and Biotechnology of the Martin-Luther University Halle-Wittenberg.

At first, I would like to thank Prof. Kiefhaber for giving me the opportunity to work in his group, the remarkably supervision, for giving me important scientific experience and the possibility for independent work.

I want to thank Prof. Wolfgang Paul for mentoring, scientific discussion and his support as my Mentor in the SFB TRR102. Also I want to especially thank Prof. Andrea Sinz and Prof Kay Saalwächter for support and valuable discussions.

I want to especially thank PD Dr. Ralph Golbik for scientific discussion and unusually great support. Also I want to specially thank PD Dr. Hauke Lilie for an important support and scientific discussion.

I thank Dr. Maria Ott and PD Dr. Stephan König for the opportunity for measurements and scientific discussion.

Many thanks to Jan Ebenhan, I could never finished my thesis without his support in programming.

Many, many thanks to Dr. Martin Voigtmann and Malte Neudorf for scientific discussion, support and the force to finished my thesis.

I want to acknowledge all former and current group members in the lab at Garching: Dr. Ursula Zinth, Peter Kämmerer, PD Dr. A. Ogrodnik and at Halle: especially Uwe Fandrich, Kerstin Stein, Uta Best and Dr. Vipul Kumar.

At last but not least, many, many thanks to my girlfriend Josepha Xenia Glück for a great support and help to finish my work and to survival. Also many, many thanks to my mother and of course to all my friends from my home town.