

TECHNISCHE UNIVERSITÄT MÜNCHEN

Lehrstuhl für Biophysikalische Chemie

The role of a hydrophobic cluster for stability and assembly kinetics
of the trimeric foldon domain from T4 fibrinin

Tobias Schümmer

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

Doktors der Naturwissenschaften

genehmigten Dissertation.

Vorsitzender: Univ.-Prof. Dr. Chr. F. W. Becker

Prüfer der Dissertation:

1. Univ.-Prof. Dr. T. Kiefhaber
2. Univ.-Prof. Dr. M. Groll

Die Dissertation wurde am 06. 09. 2011 bei der Technischen Universität München
eingereicht und durch die Fakultät für Chemie
am 11. 10. 2011 angenommen.

Perfektion ist nicht dann erreicht,
wenn es nichts mehr hinzu zu fügen gibt,
sondern wenn man nichts mehr weglassen kann.

Antoine de Saint-Exupéry

Für meine Eltern

Abstract

Foldon is the trimerization domain of T4 phage fibrin. It is of great interest for folding studies as it is highly optimized for rapid folding and association. The isolated foldon domain, consisting of 27 amino acid, can form native-like trimers, which show both high stability and fast association kinetics. Association is initiated after rapid formation of a monomeric intermediate, which is already structured. This structure resembles the structure of a monomer in the trimeric native state, as known from an NMR study. The monomeric state of this intermediate is stabilized by hydrogen bonds of the β -hairpin, and hydrophobic interactions of several amino acids which form a hydrophobic cluster. In order to study the effect of these hydrophobic interactions, the contribution of each amino acid in the hydrophobic cluster is probed by insertion of at least one mutation. The effect of each mutation is investigated by determination of the fluorescence spectra of the different states, determination of the free folding enthalpy $\Delta G^0(H_2O)$, unfolding and refolding kinetics, and the crystal structure of the native state. All inserted mutations in the hydrophobic cluster lead to a decrease of the association kinetics, despite having little to no effect on both the crystal structure and fluorescence spectra. The strongest effect is observed for the dimerization reaction, which is decelerated by a factor of up to 10,000. This deceleration is accompanied by a loss of stability of the native state of up to $50 \text{ kJ} \cdot \text{mol}^{-1}$. Two prolines, which are part of a poly-proline II structure in the N-terminal part, exert an effect on the association reaction. Their function could be the induction of this structure in the intermediate state. Only the proline at position 7 is responsible for the observed *cis-trans* isomerization reaction found in foldon wild-type. Mutations at position 6 are even able to change the reaction mechanism. Dimerization of foldon monomers with a *cis* conformation of the Xaa6-Pro7 peptide bond is not possible for Xaa being any amino acid except the wild-type alanine. From the experimental results it becomes clear that the fast association of foldon is guided by the fast equilibrium between the unfolded and the intermediate state. Secondary structure motifs formed early during the folding reaction allow efficient formation of the dimeric encounter complex. The interactions between different regions of the monomeric foldon are ensured by hydrophobic interactions and hydrogen bonds.

Contents

Abstract	I
1. Introduction	1
1.1. Proteins	1
1.2. Protein stability	3
1.3. Protein folding kinetics	5
1.4. Folding and assembly of dimeric and trimeric proteins	8
1.5. The fibritin foldon domain	10
2. Aim of thesis	19
3. Material and methods	21
3.1. Peptide synthesis and purification	21
3.2. Expression and purification from <i>E.coli</i>	23
3.3. Crystallization of foldon variants and X-ray structure determination	28
3.4. Fluorescence spectra determination	30
3.5. Denaturant transitions	31
3.6. Equilibrium <i>N</i> -test of foldon in guanidinium chloride	35
3.7. Refolding kinetics	36
3.8. Reaction order determination	38
3.9. Unfolding kinetics	40
3.10. Rate constant determination	42
3.11. Interrupted refolding kinetics	45
4. Results	47
4.1. Design of foldon variants	47
4.2. Foldon P4T	50
4.3. Foldon A6Abu	58

Contents

4.4. Foldon A6S	67
4.5. Foldon A6V	74
4.6. Foldon P7F	78
4.7. Foldon P7V	78
4.8. Foldon P7Nva	81
4.9. Foldon P4T P7Nva	87
4.10. Foldon Y13F	92
4.11. Foldon W20Nal	99
4.12. Foldon W20H	102
4.13. Foldon L22F	109
4.14. Denaturant transitions	118
4.15. Unfolding kinetics	122
4.16. Determination of α_D -values	124
4.17. Folding kinetics	128
5. Discussion	137
6. Summary	151
A. Appendix	155
A.1. Crystal structure data	155
A.2. GdmCl transitions	161
A.3. Urea transitions	166
A.4. GdmCl unfolding kinetics	169
A.5. Urea unfolding kinetics	176
A.6. Initial slope determination	181
A.7. Reaction order from the half-life times	185
List of Figures	189
List of Tables	193
Bibliography	195

1. Introduction

1.1. Proteins

Proteins are the most versatile group amongst biological molecules. They are responsible for a wealth of different tasks in cells and organisms, from metabolism and signal transduction to the formation of the cytoskeleton. Proteins facilitate transport processes and act as hormones or anti-bodies during humoral immune response¹. All proteins are composed of a set of 20 canonical α -L-amino acids. These amino acids feature different chemical properties of their side-chains, which can be hydrophobic or hydrophilic, polar or non-polar, negatively/positively charged or uncharged. The backbone is uniform for all amino acids, with an amine- and a carboxyl group. The sequence of the amino acids constituting a protein is called the primary structure. The information for this primary structure is stored within the genetic code by the 4 different nucleotides which form the DNA. This genetic information is first transcribed into the mRNA. Ribosomes translate the mRNA into the polypeptide chain by the help of tRNA molecules, each loaded with a specific amino acid². The ribosomes are large RNA-protein complexes which are able to catalyze the formation of the peptide bonds between the individual amino acids and the growing polypeptide chain in a highly efficient and almost error-free process, reaching a speed of up to 20 amino acid couplings per second³. The newly formed polypeptide chain is released from the active site of the ribosome through a narrow exit tunnel in the large ribosomal subunit into the cytosol of the cell⁴. Folding of isolated structural elements (e.g. α -helices) can already occur within this tunnel⁵. However, it was shown that the ribosome is not necessary for the protein folding reaction⁶. Proteins consist of secondary structure elements like α -helices, β -pleated sheets and turns⁷⁻⁹. These structures are mostly stabilized by hydrogen bonds between backbone atoms¹⁰, but side chain interactions can significantly contribute to the stability¹¹. These interactions can be intra-segmental within a secondary structure motif, as in α -helical backbone hydrogen bonds, or inter-segmental between secondary structure motifs, like the hydrogen bonds between individual strands of a β -pleated sheet. Secondary structure elements can be combined in different ways to form

1. Introduction

domains¹⁰. This tertiary structure of proteins is stabilized by non-covalent interactions like hydrogen bonds, hydrophobic and van der Waals interactions, as well as ionic interactions or disulfide bonds¹². Small proteins consist of a single domain, while large proteins can contain two or more domains. The most common proteins consist of 100 to 700 amino acids, and the size of individual domains can range from 29 to 300 amino acids¹³⁻¹⁵. Individual proteins can associate into higher-order complexes, forming the quaternary structure. This quaternary structure is stabilized by non-covalent interactions like hydrogen or ionic bonds or hydrophobic interactions between the subunits of the complex¹⁶. Additionally, the protein can be prone to posttranslational modifications like glycosylation or phosphorylation, as well as binding of cofactors or prosthetic groups¹⁷. However, the process of finding the correct tertiary or quaternary structure is the first task most proteins have to undergo in order to become biologically active. An exception are intrinsically unfolded proteins, which are partially or completely unfolded and only fold upon binding to other proteins, DNA or prosthetic groups¹⁸. Proteins can fail to find their correctly folded structure. Misfolded proteins are normally degraded into their amino acids by peptidases and the proteasome^{19,20}, and these amino acids can be used as building blocks for new proteins. Non-degradable misfolded proteins can lead to the formation of aggregates, which can cause severe diseases like Alzheimer's, Parkinson's or several forms of encephalopathies as BSE, scrapie or Creutzfeld-Jacob's disease^{21,22}.

Finding of the correct three-dimensional structure is of vital importance for a protein to become functional. The folding process is independent of any other factors in most cases. Anfinsen and co-workers had found that the information for correct folding is stored within the primary structure of the polypeptide chain⁶. Later, Levinthal postulated that there must be an underlying mechanism for the folding of the polypeptide chain, as by pure chance the search for the correct conformation even for a small protein of 100 amino acids would take up to about 10^{27} years²³. He calculated this number by assuming that every ψ - and ϕ -angle in the protein backbone could adopt two different conformations, and that two peptide conformations per second could be searched. In nature, however, protein folding occurs in the time span of milliseconds to minutes, making the existence of folding pathways highly probable.

1.2. Protein stability

As proteins fold without the input of energy, protein folding is an exergonic process. The folding free enthalpy $\Delta G^0(H_2O)$ of monomeric proteins is only around -20 to $-40 \text{ kJ} \cdot \text{mol}^{-1}$ ²⁴, which equals the stability gained from a few newly formed hydrogen bonds or salt bridges^{25,26}. The native conformation of a protein on the other hand is stabilized by a vast number of several different interactions¹², as discussed in 1.1. The stability of proteins is so low due to compensating effects of the enthalpy and entropy. Folding of a polypeptide chain has a janiform effect on entropy: Conformational freedom of the polypeptide chain is decreased, while the solvent entropy increases. In the unfolded state, exposition of the hydrophobic side-chains to the solvent leads to the formation of a rigid hydrogen-bond network of the solvent molecules around the side-chains. The formation of these "iceberg-like" structures lowers the water entropy in the system. Upon folding, the hydrophobic amino acid side-chains get buried in the core of the protein, increasing the solvent entropy due to the melting of the "iceberg-like" water structure. This increase tips the scale towards favoring the folded state²⁷. The Gibbs-Helmholtz-equation allows the calculation of the free enthalpy ΔG^0 of any reaction at a certain temperature T ²⁸:

$$\Delta G^0 = \Delta H^0 - T \cdot \Delta S^0 \quad (1.1)$$

where ΔH^0 is the enthalpy and ΔS^0 is the entropy of the reaction. The free enthalpy can also be expressed by the fundamental equation, which can be derived from the fundamental equations of the intrinsic energy²⁸:

$$dG = V dp - S dT + \sum_i \mu_i dn \quad (1.2)$$

with V being the volume, p the pressure, S the entropy, T the temperature, μ the chemical potential and n the amount of substance of the system. The position of equilibrium K between two reactants is directly dependent on the free enthalpy.

$$\Delta G^0 = -RT \ln K \quad (1.3)$$

with R being the constant of the ideal gas equation. The position of the equilibrium of a protein between the unfolded state U and the native state N can be shifted by changes in pressure p , temperature T and the composition of the system n , as shown in equation 1.2. Changing the systems composition is in most cases done by addition of co-solvents. Substances like guan-

1. Introduction

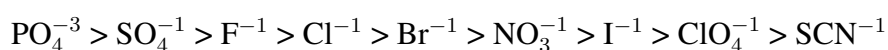
dinium chloride (GdmCl) and urea in aqueous solutions can denature proteins by stronger destabilizing the native state compared to the unfolded state²⁹. These co-solvents cause an abrupt change from the native to the unfolded state of a protein over a relatively small concentration range of the denaturant by forming a complex of the protein with a large number of denaturant molecules in solution, thereby disrupting the hydrophobic core^{30,31}. Strong binding of the denaturant to the proteins was proposed³². Schellman and co-workers refined this model in their thermodynamic studies by showing that binding of the denaturants to the protein is weak³³. The co-solvent concentrations need to be high to successfully unfold a protein, thus it is also possible that by chance the solvent molecule bound to the protein is a denaturant molecule, and not due to preferential binding; thus the theory was expanded to the solvent-exchange model³⁴.

Probing of a peptide solution with varying denaturant concentrations by spectroscopic methods like absorption, circular dichroism absorption or fluorescence reveals the shift of the equilibrium of the folded state N to the unfolded state U . As the equilibrium constant can be determined for each denaturant concentration from the baselines of the native and unfolded state, plotting of the free folding enthalpy against the denaturant concentrations reveals a linear dependence of the free folding enthalpy ΔG^0 on the denaturant concentration³⁵. The m_{eq} -value is a measure of this linear dependence^{36,37}. This allows the linear extrapolation to zero molar denaturant and the determination of $\Delta G^0(\text{H}_2\text{O})$ by the linear extrapolation method first used by Alexander, Pace and Tanford^{38,39}. Santoro and Bolen developed the non-linear least squares fit to determine the stability and m_{eq} -value of a protein by direct fitting of the obtained denaturant transition curves⁴⁰. Myers et al. claimed that the m_{eq} -value is also directly proportional to the change in accessible surface area ΔASA and the change in heat capacity upon unfolding ΔC_p ⁴¹.

The linear extrapolation method should give identical free folding enthalpies regardless of the used denaturant. It was however found that the results obtained by both urea and GdmCl can deviate by 25 to 50 %³⁶. Makhatadze showed that the urea transition curves are in good agreement with thermal transition curves while GdmCl transitions can deviate from the proposed linear behavior⁴². The large ionic strength of GdmCl solutions even at low GdmCl concentration has to be compensated for, otherwise deviation from linear behavior can occur for GdmCl solution below 1.5 M GdmCl concentration⁴³. These findings show that additional measurements have to be carried out to confirm deviations between urea and GdmCl transitions. These

measurement can ideally probe the native and unfolded conformation for their thermodynamic parameters⁴⁴.

Next to co-solvents, dissolved salts can have a strong effect on the stability of proteins⁴⁵. The effect of the cations and anions can be ranked according to the Hofmeister series^{46,47}. A incomplete ranking of the most common anions is:



The ranking is from most stabilizing to most destabilizing. The Hofmeister series is also dependent on the isoelectric point (pI) of the protein. At a pH higher than the pI, the salts order in the Hofmeister series (as shown above), while the opposite order is commonly observed for a pH lower than the pI^{48,49}. Stabilizing salts cause a salting-out of the protein, strengthening the hydrophobic interactions^{50,51}. The opposite effect is seen for the destabilizing salts, which cause a salting-in, accompanied by a better solubility of the solvent for the proteins and weakening of the hydrophobic interactions^{52,53}. The ions act on the proteins not by effecting the bulk water properties^{54,55}, but rather by direct interaction of the ions with the effected macromolecules. This is derived from binding studies of anions with model molecules^{52,53}.

1.3. Protein folding kinetics

Protein folding kinetics give additional information to thermodynamic data, as they can give rise to detailed knowledge of the folding mechanism^{56,57}. Stopped-flow measurements are the most common method to study protein folding, where a reaction can be followed spectroscopically after rapid mixture of two solutions^{58,59}. Aromatic amino acid side-chain often strongly change their fluorescence intensity when transferred from the hydrophobic environment within to the protein to a the solvent^{38,60,61}. The change in circular dichroism absorption upon formation or disruption of secondary structure motifs is another method to follow a folding or unfolding reaction⁶²⁻⁶⁴. A rapid change of the buffer system can be induced by the mixing process either by dilution, addition of cosolvents, or change of the pH-value. Upon mixing the equilibrium of the protein is disturbed, which will consequently relax into a new equilibrium⁵⁷. The stopped-flow method allows detection of fast reactions, but has the draw-back of a dead-time of about 1 ms⁵⁸. Fast reactions which occur within the dead-time thus cannot be followed and evaluated by the stopped-flow measurements. Methods of choice in these

1. Introduction

cases are continuous-flow, temperature- or pressure-jump experiments^{57,65}. All these experiments can be followed by absorption, fluorescence or circular dichroism absorption, as well as time-resolved nuclear magnetic resonance (NMR) detection⁶⁶.

The most simple model for protein folding often found for small proteins or isolated protein domains is a two-state system, where the unfolded and folded state are in rapid equilibrium⁶⁷. Measuring the apparent rate constant k_{app} by rapid dilution of denatured protein into varying final concentrations of denaturant and plotting the common logarithm of the individual rate constants against the corresponding denaturant concentrations yields a V-shaped curve, often termed chevron plot^{68,69}. The two arms of this plot correspond to the refolding and unfolding reaction of the protein⁷⁰. The rate constants k_f and k_u depend linearly on the denaturant concentration⁶⁸, as seen for the free folding enthalpy ΔG^0 . By a linear fit of the arms and extrapolation to zero denaturant, the rate constants of the folding and unfolding reaction can be determined in water. The slopes of these linear fits are the kinetic m -values. At the minimum of the graph, both the unfolding and refolding reaction are in equilibrium, thus the denaturant concentration at the minimum equals the transition midpoint in a denaturant transition if using the same buffer and experiment conditions. Additionally, the sum of the kinetic m -values yields the equilibrium m_{eq} -value^{41,71}. Non-linear behavior in chevron plots is often a consequence from deviations from ideal two-state behavior. This can be found for either in the unfolding, the refolding or both limbs, and there are several reasons for this behavior⁷²⁻⁷⁴. In an ideal two-state folding mechanism, both the unfolded and the native conformation should be uniform. This means for the unfolded state, which shows a large heterogeneity of its conformations, that all proteins can readily interconvert and that there are no sub-populations which are separated by a barrier^{75,76}. Even a small percentage of unfolded proteins in a different ensemble can cause an additional phase in the folding kinetics reaction due to a slower refolding reaction^{77,78}.

Proline residues are one of the major sources for complex folding kinetics and formation of folding intermediates⁷⁹. In most cases, all proline residues need to be in the *trans* conformation to allow for a rapid collapse of the peptide chain. In the unfolded state, the Xaa-Pro peptide bond can exist in both the *cis* and the *trans* conformation, thus, in many cases, isomerization of the peptide bond has to precede the folding reaction⁸⁰. Proline *cis-trans* isomerization occurs with time constants varying from 10 to 60 s at 25 °C, which is in most cases significantly slower than the protein folding reaction⁸¹. The slow isomerization reaction is caused

by the high activation enthalpy of isomerization of approximately $80 \text{ kJ} \cdot \text{mol}^{-1}$ ⁸². The variation in the time constants, accompanied by a variation of the *cis*-content from 7 to 36 % can be caused by the varying amino acids at the N-terminal position of the proline⁸¹, and local structure elements in the unfolded state can have additional impacts on the isomerization reaction. Proline *cis-trans* isomerization can be sped up by peptidyl-prolyl *cis-trans* isomerases which catalyze the isomerization step⁸³. A replacement of the proline by mutation can cause disappearance of the *cis-trans* isomerization reaction⁸⁴. However, not all proline residues in their non-native conformation lead necessarily to a deceleration of the folding reaction. In bovine pancreatic RNase A, a protein with four proline residue, the proline at position 114 was found to be non-essential for the folding kinetics under favorable conditions, regardless of its conformation^{85,86}. Protein refolding can even lack all detectable *cis-trans* isomerization reactions despite the presence of several prolines. The fibronectin III domain type 10 (¹⁰FNIII) contains 7 proline residues and nevertheless refolds within 1 s without any detectable proline *cis-trans* isomerization when refolded from a completely unfolded state without any residual structure⁸⁷.

Deviations from ideal two-state behavior can also be due to aggregation of the protein, and shifts of the transition state^{88,89}. An often observed reason for deviations from two-state behavior is the existence of folding intermediates. These intermediates can be either on- or off-pathway of the folding reaction⁹⁰. Direct evidence for folding intermediates like additional phases in the unfolding or refolding reactions are only observed if the intermediate is stabilized with regard to the native or unfolded state and thus at least transiently populated during the refolding reaction^{91,92}. The concentration of these transiently populated folding intermediates can be detected by interrupted refolding experiments⁹³, often allowing the determination of all microscopic rate constants^{94,95}. Formation of secondary structure is readily determined by detection of the change in circular dichroism absorption during the folding reaction⁹⁶, while the change of the size of the protein can be measured by small angle X-ray scattering (SAXS)⁹⁷. Both methods can be used efficiently if coupled to a stopped-flow or continuous-flow mixing setup. For many small proteins, several or even all of these methods have been applied to determine the folding pathway and all microscopic rate constants^{57,65,98}.

1.4. Folding and assembly of dimeric and trimeric proteins

Folding of oligomeric proteins is in general studied by the same methods as folding of monomeric proteins. The stability can be readily determined by denaturant or thermal transitions and evaluated by a two-state model³⁶. This evaluation using a two-state model can only be applied to proteins constituted of several subunits as long as the equilibrium transition between the folded oligomer and the unfolded monomers is fast and reversible. Folding kinetics of many dimeric proteins also follow a two-state model⁶⁷.

When two-state behavior for dimer folding is observed, the apparent rate constant can either be a first- or second-order rate constant. Folding of p22 arc repressor or general control protein (GCN)-4 are diffusion-limited and thus second order^{99,100}. Repressor of dimer (ROP) on the other hand shows a unimolecular rate constant, making the rearrangement after association rate-limiting¹⁰¹. In dimeric proteins, one has to clearly distinct the association of homo- and heterodimeric proteins. The association of hetero-dimeric proteins was extensively studied by mutating amino acids within the dimerization interface for several protein pairs. Mutations of charged residues within the binding interface have the largest effect of all mutations on the association reaction¹⁰²⁻¹⁰⁴. The association of hetero-dimeric proteins can be governed by charge interactions, with the two binding partners having opposing net charges¹⁰⁵⁻¹⁰⁷. An increase of the ionic strength in the solvent leads to a decrease of the association rate constant for these hetero-dimeric proteins as the charges get shielded by the ions in solution^{108,109}. The opposing effect is observed for homo-dimeric proteins. As the net-charge of the proteins is identical, a high ionic strength dampens the repulsion between the subunits, making association faster¹¹⁰⁻¹¹². Although the association reaction is strongly influenced by diffusion, no dimerization reaction between two proteins determined so far reached the diffusion-limit of $5 \cdot 10^9 \text{ M}^{-1} \cdot \text{s}^{-1}$, as not 100 % of all encounter complexes lead to association¹¹³.

While it is known that binding precedes folding for intrinsically unfolded proteins¹¹⁴, for most homo- and hetero-dimeric proteins the case is less clear. There are two possible scenarios, and both are found in nature. The association of S-peptide to S-protein is shown to be only dependent on native-like hydrophobic interactions and formation of native-state secondary structure is a fast step after association¹¹⁵. General control protein (GCN)-4 forms helices

prior to binding, and this secondary structure formation is necessary for association¹¹⁶. An increase of the helix content of the monomeric state leads to an increase of the association speed¹¹⁷. Similar effects had been observed for several other proteins^{116–120}. The rapid secondary structure formation ensures fast formation of interaction interfaces on the monomer surface and thus fast assembly to other monomers if an encounter occurs¹²¹. The necessity to efficiently find and bind to form higher order complexes is very likely the reason for the rapid collapse in the monomers, making the detection of this reaction almost impossible with rapid kinetic measurements¹²¹. Insertion of hydrophobic amino acids within the hydrophobic core or the dimerization interface can have an accelerating effect on the folding of dimeric proteins, as shown for ROP and p22 arc repressor^{26,101}. The acceleration is caused by a decreased transition state barrier, as burying of hydrophilic or charged side chains within a hydrophobic environment prior to hydrogen bond or salt-bridge formation is accompanied by an energetic penalty²⁶. This finding is in good agreement with a capacious study of 136 dimeric protein interfaces, which shows that the interaction interface of the monomers is constituted of hydrophobic patches combined with inter-subunit polar interactions¹²². Exchange of polar side chains within the dimer interface can both increase or decrease stability and catalytic activity, as shown for several proteins^{123,124}.

In contrast to folding of dimeric proteins, only few trimeric proteins have been studied with respect to their folding stability and kinetics. In many cases a simple $3U \rightleftharpoons N_3$ mechanism is used for the evaluation of denaturant or thermal transitions^{125–127}, although monomeric and dimeric intermediates are observed as well in some cases, like in apo γ -carbonic anhydrase or HSC70^{128,129}. The trimeric globular protein tumor necrosis factor $\text{TNF}\alpha$ folds through a molten-globule state after rapid accumulation of a transient dimeric state¹³⁰, as determined by intrinsic and anilinonaphthalenesulfonic acid (ANS) fluorescence studies and chemical cross-linking. $\text{TNF}\alpha$ monomers undergo several distinct folding processes in the monomeric state, after which rapid association occurs. Individual rate constants for the association reaction could not be determined¹³⁰. In contrast, folding of the trimeric proteins HIV gp41 and SIV gp41 is described by a simple monomer-trimer equilibrium^{131,132}. Both gp41 variants form trimeric coiled-coil structures, which are often used as models for trimer formation and kinetics. Folding of isolated trimeric coiled-coils can be described by two sequential second order reactions: The first step involves dimer formation from two monomers, while a third monomer binds to the dimer in a reversible manner in the second reaction step^{125,133}.

Trimerization of proteins can be initiated by specific trimerization domains. The C-terminal domain of P22 tailspike protein (TSP) forms a prism-like structure by association of β -sheets from the three monomers and was shown to be essential for the trimerization of the whole protein^{134,135}. The folding mechanism involves the formation of a β -helix within a stable monomeric state of the C-terminal domain, followed by a first dimerization and consequent monomer binding to a trimeric intermediate state^{136,137}. Specific ionic interactions are finally formed in a last rearrangement step from the trimeric intermediate to the native state¹³⁸. The folding mechanism shows a high resemblance to the folding of the T4 fibrin foldon domain⁹⁵ (discussed in detail in Chapter 1.5). However, kinetic measurements cannot be carried out for TSP as unfolding is irreversible for high temperatures or denaturant concentrations¹³⁵.

1.5. The fibrin foldon domain

Fibrin is part of the T4 phage structure, forming the collar and the whiskers of the T4 phage. It is composed of three chains of the protein gpwac, a 52 kDa product from the *wac* (whisker antigen control) gene and forms fibers of 53 nm length¹³⁹. Sequencing the 486 amino acids from the nucleotide sequence indicated three different domains, with the largest one, comprised of 408 or approximately 80 % of the total amino acid sequence, being the central domain¹⁴⁰. This central domain is mainly composed of 12 heptad repeat amino acid sequence motifs, an α -helical motif of the general sequence (abcdefg)_n. Positions "a" and "d" are preferably occupied by non-polar amino acids. This structure motif is able to form both dimeric and trimeric coiled-coil structures¹⁴¹; in fibrin, the coiled-coil forms trimers, as known from sequence analysis and structural data¹⁴. The heptad-repeat motifs are separated by linker regions richer in both glycine and proline residues than the heptad-repeat regions¹⁴. In total, 47 complete and two non-complete heptad repeats are present in the central domain of fibrin¹⁴⁰. The N-terminal domain of fibrin, composed of 47 amino acids, forms the collar at the neck region of the T4 phage; lacking of the N-terminal domain does not have an impact on the folding reaction of fibrin, as shown by reversible temperature-induced denaturation¹³⁹. The C-terminal domain however is known to be essential for correct folding of fibrin, as expression of C-terminal deletion mutants of fibrin in *E.coli* leads to formation of inclusion bodies¹³⁹. The C-terminal domain is composed of 29 amino acids and forms a β -hairpin, connected by a β -turn to the last α -helical turn of the central domain. Three subunits form a β -propeller

structure, which was first determined by X-ray crystallography of different length mutants of fibritin, named fibritin E and fibritin M, composed of the 119 and 78 C-terminal amino acids of fibritin, respectively^{14,142}. This C-terminal structure is named foldon¹⁴³, a name first used by the group of Peter G. Wolynes and describes quasi-independent folding units which can act as an initiator for the folding reaction¹⁴⁴. The trimerisation reaction of fibritin is initiated by the foldon domain, followed by zipper-like formation of the coiled-coil structure¹⁴⁵. Foldon catalyzes the correct folding of fibritin by two different means: First, by acting as a folding initiator, as it allows correct and in-register association of the three different peptide strands. Foldon also gives the different strands the right orientation, and enlarges the local concentration to an approximate concentration of 1 M, as the three C-terminal ends are all in a volume of 1 nm³. This apparent concentration-increase enhances the concentration-dependent association kinetics drastically¹⁴³.

The 29 amino acids at the C-terminal end of fibritin, the foldon domain, are known to be essential for the trimerization reaction, shown by partial or complete deletion of this domain from fibritin^{139,145}. Mutating the core tryptophan residue to a leucine has the same effect as complete deletion of the foldon domain, while the insertion of a hydrophobic amino acid at this position promotes correct fibritin formation¹⁴. Foldon is able to initiate trimerization of both homo- and hetero-trimers if fused to domains or even complete proteins. The foldon can be located either at the N- or C-terminal end or replace existing trimerization domains. This was proven with a collagen-mimicking sequence (GlyProPro)₁₀, HIV gp140, and gp26 of bacteriophage 22, where foldon was inserted as a new trimerization domain^{146,147}. In adenovirus fibre shaft protein, foldon can act as a replacement for the original trimerization domain¹⁴⁸. Fusion of the foldon domain to the (GlyProPro)₁₀ sequence leads to a stabilization of 42 °C for thermal transition, compared to the trimer formed by this sequence without foldon fused to it. Similar effects are also shown for the other chimeric proteins.

The structure of the isolated foldon domain in solution was determined by NMR spectroscopy⁹⁵. This isolated foldon domain consists of the first 27 amino acids of the foldon domain (residues 457 to 483, named residues 1 to 27 hereinafter)⁹⁵. Its sequence is given in Table 1.1.

This structure resembles the structure determined of truncated fibritin variants, fibritin E and M. Only slight deviations at the N-terminal end are observed, which has a higher flexibility in the isolated foldon domain⁹⁵. This foldon NMR structure is shown in Figure 1.1A. The β -sheets of the three monomers form a β -propeller structure with a three-fold symmetry axis.

1. Introduction

1	11	21
GYIPEAPRDG	QAYVRKDGEW	VLLSTFL

Table 1.1.: The amino-acid sequence of the isolated foldon wild-type domain, as used for the determination of the NMR structure.

Each foldon monomer of the native trimer structure starts with left-handed polyproline II helix structure, formed between the residues Pro4 and Pro7, which is part of the extended N-terminal region (residues Gly1 to Gln11). Both proline residues are in the *trans* conformation in the native state of the foldon trimer. This first N-terminal segment is connected to the central β -hairpin (residues Ala12 to Leu23) by a classical type II β -bend structure with a X-Asp-Gly-X sequence⁹. This motif is also found within the β -hairpin: The β -hairpin itself is stabilized by 5 hydrogen bonds and shows a larger twist than normally found in β -sheets. This twist could prevent the formation of larger β -sheets, which would finally result in aggregation¹⁴⁹. The foldon structure ends in a 3_{10} -helix, formed by the residues Leu23 to Leu27. The residues Tyr2, Ile3, Val14, Trp20, Leu22, Phe26 and Leu27 form a hydrophobic core along the symmetry axis, and the β -propeller is stabilized by backbone hydrogen bonds between Tyr13 and Arg15 of a neighbouring peptide chain¹⁴. Furthermore, the hydrocarbon sidechain of Arg15 is packed against the Trp20 sidechain of a neighbouring peptide. The loss of this interaction upon mutation of Trp20 to a non-aromatic amino acid is believed to be responsible for the formation of inclusion bodies upon expression of these mutants in *E.coli*¹⁴.

The high stability of foldon against thermal denaturation and high SDS-concentrations was known before¹⁴⁵. The free folding enthalpy of the foldon trimer is known from guanidinium chloride and urea transitions^{95,150}. The folding stability was found to be $\Delta G^0(H_2O) = (-89.2 \pm 0.6) \text{ kJ} \cdot \text{mol}^{-1}$, with an m_{eq} -value of $m_{eq}(\text{GdmCl}) = (10.2 \pm 0.2) \text{ kJ} \cdot \text{mol}^{-1} \cdot \text{M}^{-1}$ ⁹⁵ in GdmCl. The same stability is found by the urea transition ($\Delta G^0(H_2O) = (-89.2 \pm 0.3) \text{ kJ} \cdot \text{mol}^{-1}$), with an m_{eq} -value of $m_{eq}(\text{urea}) = (5.4 \pm 0.1) \text{ kJ} \cdot \text{mol}^{-1} \cdot \text{M}^{-1}$ ¹⁵⁰. However, this stability corresponds to a monomer concentration of 1 M, whereas the concentration in the living cell is approximately 5 μM . The actual stability of foldon at this concentration is $\Delta G^0(H_2O)_{5\mu\text{M}} = -29.7 \text{ kJ} / \text{mol}$ ⁹⁵.

A major contribution to the stability of the native foldon structure arises from an intermolecular salt-bridge between the residues Glu5 and Arg15 of the neighboring foldon monomer; each of the three salt-bridges contributes up to 16 kJ / mol to the folding stability¹⁴⁹. This value was

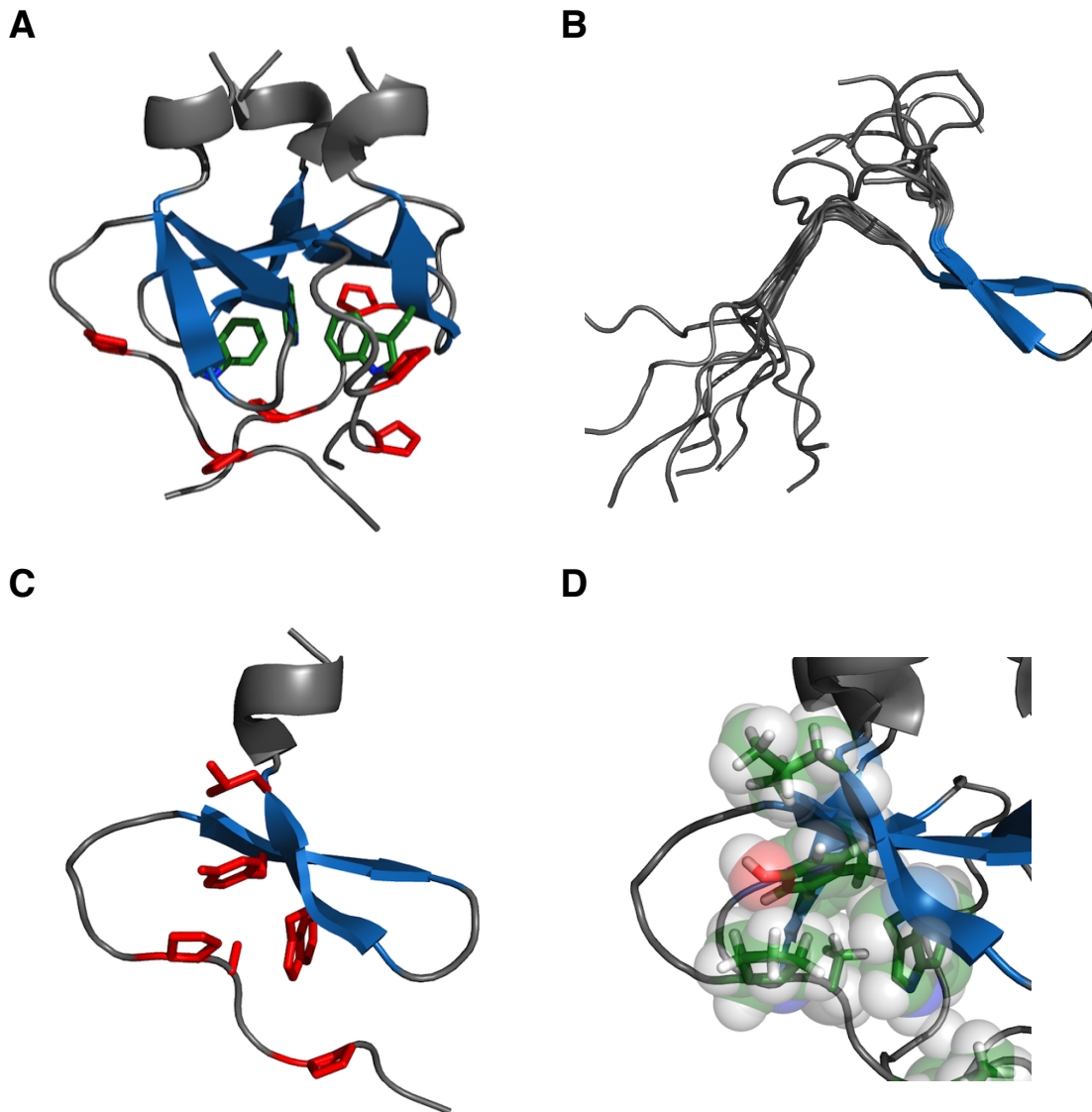


Figure 1.1.: NMR structures of foldon wild-type. A) NMR structure of foldon wild-type in ribbon representation, with the β -strands highlighted in blue, while the 3_{10} -helix and the other regions of the structure are colored in grey. The side chains of the two proline residues at position 4 and 7 and the tryptophan at position 20 are shown in stick representation and the side-chains colored in red and green, respectively. B) The ten most stable conformation of the acid-denatured A-state of foldon. The β -strands are highlighted in blue, while the rest of the foldon molecule is colored in gray. C) The NMR structure of the foldon E5R variant. The side-chains of the amino acids which form the hydrophobic cluster, Pro4, Ala6, Pro7, Tyr13, Trp20 and Leu22, are shown in stick representation and are colored in red. D) The van-der-Waals interactions of the hydrophobic cluster are shown in a sphere representation, with the side chains shown in stick representation, and the carbon atoms colored green, the oxygen atoms in red, nitrogens in blue and hydrogen atoms in white. The other regions apart from the hydrophobic cluster of the foldon molecule are colored as in A).

1. Introduction

determined by NMR studies of the pK_a of the glutamic acid side chain, which was found to be 1.7¹⁴⁹. Disruption of this salt-bridge has major consequences on the stability of foldon, as was shown by two different experiments: In buffer conditions with pH 2.0, the trimer dissociates into monomers, which still have a near-native fold within the β -hairpin, but show a higher flexibility in the N- and C-terminal regions. The ten lowest energy CNS conformers are determined and deposited in the RCSB Protein Data Bank (entry number 1U0P)¹⁴⁹. But despite the fact that the N-terminal segment is no longer packed against the β -hairpin, this segment still shows a near-native backbone conformation. Only the β -hairpin twist is more relaxed than in the native trimer¹⁴⁹. Dissociation of the trimer by disruption of the salt-bridge can occur in this case although the pH-value is higher than the pK_a as all spontaneous dissociation will result in a hydration of the glutamic acid side chain. This results in a shift of the equilibrium towards the unfolded state.

An even more efficient way to disrupt the salt-bridge is by an amino-acid exchange of the glutamate at position 5 to an arginine. The replacement of the negatively charged glutamic acid to a positively charged arginine leads to electrostatic repulsion and thus hinders successful trimer formation at low salt concentrations¹⁵¹. At high salt concentrations charges can be shielded by solvent ions and dimerization/trimerization can occur in this mutant (A. Reiner, unpublished results). The NMR-structure of this E5R variant was determined (Protein Data Bank accession number 2KBL)¹⁵¹. This structure closely resembles the structure of the monomers in the native trimer. Additionally, the formation of a hydrophobic cluster is observable, consisting of the residues Pro4, Ala6, Pro7, Tyr13, Trp20 and Leu22. The folding pathway of wild-type foldon and the rates for each individual association and folding step were determined by stopped-flow refolding experiments, in combination with an interrupted refolding experiment and consequent numerical fitting to the used folding model (Figure 1.3)⁹⁵. The fitted refolding and interrupted refolding traces are given in Figure 1.2. The model with all determined rate constants is given in Figure 1.3.

The formation of a burst-phase intermediate within the dead-time of the stopped-flow measurements is observable. Only the following steps from the intermediate I over the dimer D to trimer T and finally to the native trimer N can thus be resolved by stopped-flow experiments⁹⁵. A concentration-independent phase can be assigned to proline *cis-trans* isomerization by initiating the refolding reaction in the presence of the proline isomerase cyclophilin⁹⁵. This increases the rate of the slow reaction by one order of magnitude. It is not known if this effect

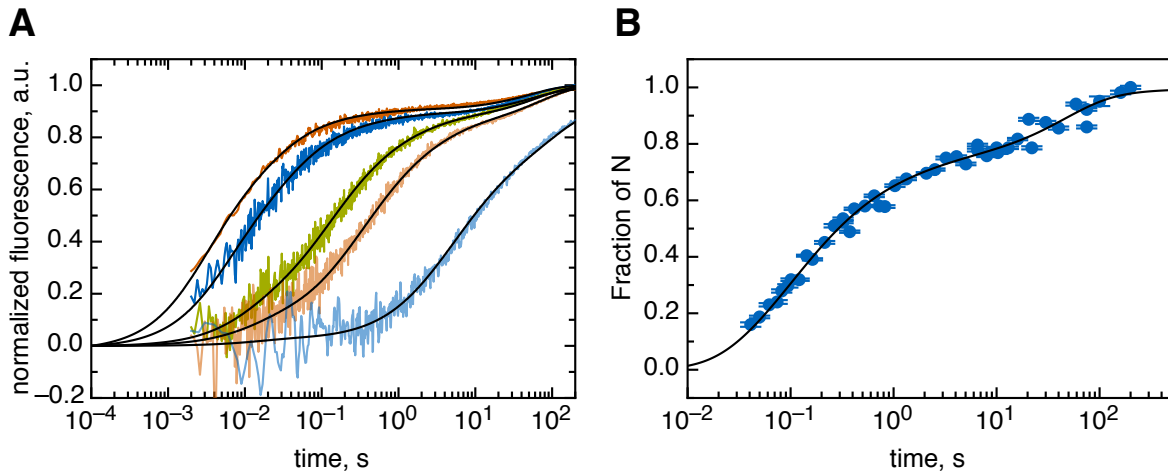


Figure 1.2.: A) Refolding kinetic traces of foldon wild-type at foldon monomer concentrations varying from $1 \mu\text{M}$ to $102 \mu\text{M}$. The black lines represent the fitting curves from the global fit. B) Interrupted refolding amplitudes of $10 \mu\text{M}$ foldon monomer concentration plotted against the time-span given for refolding. The black line represents the fitting curve of the global fit. Figures taken and modified from⁹⁵.

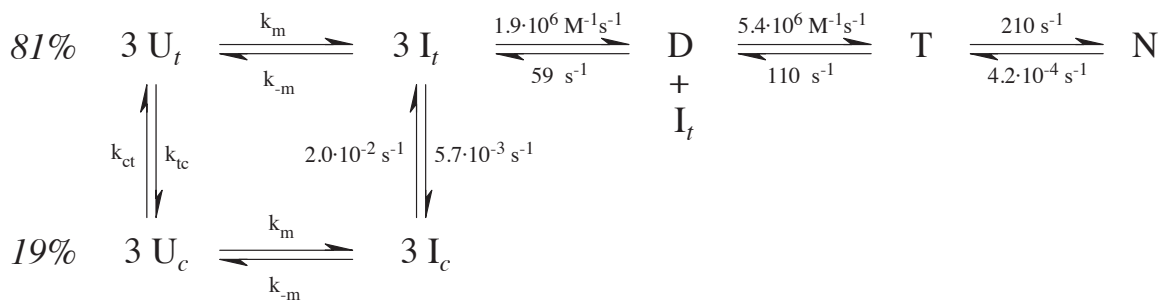


Figure 1.3.: The foldon assembly pathway with all determined rate constants in GdmCl. Figure taken from reference⁹⁵.

arises from one or both proline residues present in the foldon sequence. All fitting was done under the assumption that only foldon intermediates in the all-*trans* conformation of their proline residues are able to form dimers which can act productively in the association reaction⁹⁵. The association steps are close to the fastest known protein association reactions^{26,118}, yet the association reaction of foldon is not diffusion-limited: Foldon variants containing the triplet-excitable labels xanthonic acid as donor and naphthyl-alanine as acceptor at position 20 were used in triplet-triplet-energy-transfer (TTET) experiments to determine the bimolecular rate constant with $k = 5.3 \cdot 10^9 \text{ M}^{-1} \cdot \text{s}^{-1}$. This value is about 200fold higher than the association steps of foldon in GdmCl. Fluorescence spectra of foldon wild-type were determined in the native and unfolded state. Furthermore, the burst-phase intermediate fluorescence spectrum

As ionic interactions are not able to induce fast association of the foldon subunits, the question remains which interactions govern the association of the T4 phage fibrin foldon domain. The hydrophobic cluster determined in the foldon E5R NMR structure is a candidate for forming the initial contacts among the subunits, as the fluorescence spectra of the foldon wild-type burst-phase intermediate and the foldon E5R native state spectrum show a large similarity. Is the hydrophobic cluster formation essential for native trimer formation, and which role do the individual amino acids of the hydrophobic cluster play during this reaction? And if the hydrophobic cluster is indeed essential for the association, are the effects exerted by the different amino acids caused by a common mechanism? It is additionally of interest whether only one or both of the proline residues are responsible for the *cis-trans* isomerization reaction observed in refolding kinetics, and if folding can be enhanced if these prolines are removed from the foldon domain. Any observed effects can be caused by structural, thermodynamic or kinetic effects, or a combination of all of these. The aim of this thesis is to unravel these effects.

2. Aim of thesis

The foldon domain of T4 phage fibrin is a natural occurring trimerization domain. The isolated foldon domain, consisting of 27 amino acids, is a well-studied model system for the folding and association reaction of small multimeric proteins. However, the association mechanism is still poorly understood. The stability and folding kinetics of foldon were determined in GdmCl and urea, as was the NMR structure of the native trimer. Additionally, the NMR structure of the foldon mutant E5R was determined. This mutation renders foldon incapable of trimerization at low salt concentrations in solution. The isolated monomer revealed the existence of a hydrophobic cluster consisting of 6 amino acids.

The aim of this study is to evaluate the effect of mutations at the different position in the hydrophobic cluster of foldon. As this hydrophobic cluster could be important for the association reaction of the foldon monomers, the impact of each mutation on structure, stability and folding kinetics is measured. More than one amino acid was used to replace the wild-type amino acid at several positions to evaluate from the folding mechanism and reaction order if the mutations act in a comparable way on the association reaction or if one has to clearly distinguish between the different inserted amino acids and their position. Mutations of the two prolines located within the hydrophobic cluster allow the determination which proline is responsible for the observed *cis-trans* isomerization reaction. The two proline residues are located in a poly-proline type II structure in the N-terminal region of the foldon monomer. The function of this structure at the N-terminal end of foldon is also of interest in this study. Kinetic measurements and global fitting allow the determination of each individual rate constant, deciphering the function of the hydrophobic cluster during the assembly process, especially the effect on both dimerization and trimerization. Further we try to solve the native structure of the varying foldon variants by X-ray crystallography and want to see if any effects on the stability caused by the mutations are due to structural changes of the native state or changes of the equilibrium between the unfolded and folded state.

3. Material and methods

All chemicals used throughout this work were purchased either from Carl Roth (Karlsruhe, Germany) or Sigma Aldrich (St. Louis, MO, USA), if not stated otherwise. Fmoc-protected amino acids were purchased from either Novabiochem (Hohenbrunn, Germany) or Iris Biotech (Marktredwitz, Germany). Restriction enzymes, Pfu polymerase and all buffer stock solutions were purchased from Promega (Madison, WI, USA).

3.1. Peptide synthesis and purification

Peptide synthesis

All foldon peptides used in this study except GS-foldon A6S were synthesized using 9-fluorenylmethoxycarbonyl solid-phase peptide synthesis¹⁵². Synthesis was run using a TentaGel S PHB-Leu Fmoc resin from RAPP Polymere (Tübingen, Germany). Synthesis was carried out on a 433A Peptide Synthesizer by Applied Biosystems (Wilmington, DE, USA) using a batchwise approach at a 0.1 mmol binding capacity of the chosen resin. Fmoc removal is done with 20 % piperidine in N-Methyl-2-pyrrolidone (NMP) and controlled via UV feedback monitoring at 301 nm for the Fmoc leaving group. Deprotection was repeated until the absorbance value was less than 5 % of the initial absorbance value. Fmoc-protected amino acids were activated by O-benzotriazole-N,N,N',N'-tetramethyl-uronium-hexafluorophosphate (HBTU) or N-hydroxybenzotriazole (HOBt) / HBTU and then coupled to the unprotected N-terminal end of the growing peptide chain. For coupling, amino acids were used in a ten-fold excess above the resin capacity.

As aggregation of the peptides was observed for most synthesized foldon variants at position 19 (glutamate), a oxazolidine dipeptide (Fmoc-Leu-Ser($\Psi^{Me,Me}$ pro)-OH by Novabiochem (Hohenbrunn, Germany)) was inserted at position 23/24 to prevent secondary structure formation or formation of hydrophobic contacts and packing against other peptide chains or the resin^{153 154}. To prevent aspartimide formation which would lead to the formation of a

3. Material and methods

β -aspartyl peptide by cleavage with piperidine or basic compounds¹⁵⁵, the amino group of the following glycine is protected by either a 2,4-dimethoxybenzyl (Dmb)- or a 2-hydroxy-4-methoxybenzyl (Hmb)-group^{156 157}. Incorporation of this back-bone protected residue was achieved by manual coupling in N,N-dimethylformamide (DMF). The resin was washed 3 times with dichloromethane (DCM), followed by three washing steps in DMF to achieve swelling of the resin in DMF. 0.3 mmol of the desired amino acid, either Fmoc-(Dmb)Gly-OH or Fmoc-Asp(tBu)-(Hmb)Gly-OH, were incubated in 1.6 ml DMF with 0.3 mmol benzotriazol-1-yl-oxytripyrrolidinophosphonium hexafluorophosphate (PyBOB) for 5 min at room temperature. Addition of 130 μ l 44 % N-methylmorpholine (NMM) in DMF to this solution initiates the coupling reaction. The resin was incubated for 45 min under shaking at room temperature before washing with DMF and thoroughly washing and swelling in DCM before the synthesis was continued on the peptide synthesizer. Solid-phase peptide synthesis allows the incorporation of non-physiological amino acids. In total, three non-physiological amino acids have been incorporated into four different foldon variants throughout this study: 2-(L)-aminobutyric acid (Abu), (L)-norvaline (Nva), and 1-(L)-naphthylalanine (Nal). The chemical structure of these three amino acids is given in Figure 3.1.

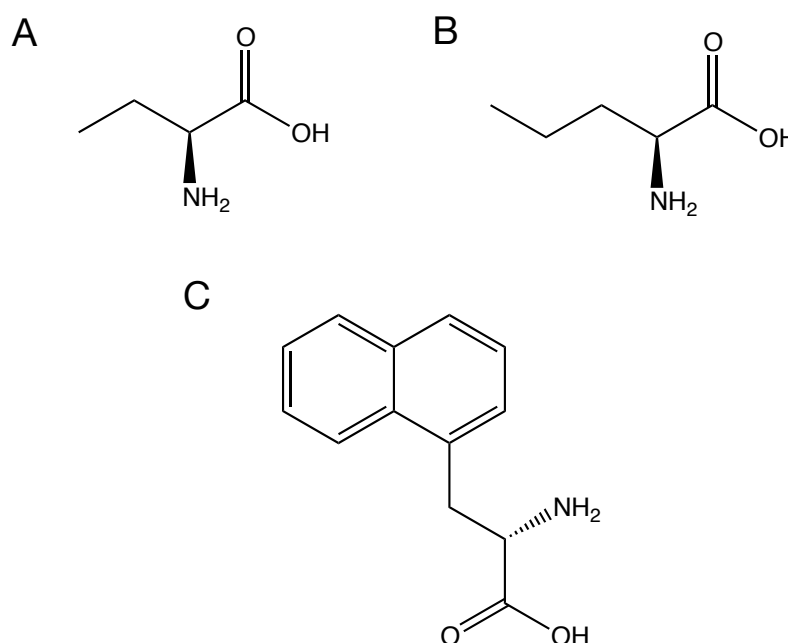


Figure 3.1.: Chemical structure of the non-physiological amino acids used in this study, A) 2-(L)-aminobutyric acid (Abu), B) (L)-norvaline (Nva), C) 1-(L)-naphthylalanine (Nal).

Peptide purification

All peptides synthesized as described in 3.1 were purified by preparative reverse-phase high pressure liquid chromatography (HPLC). For this, the peptides were cleaved off the resin by incubation in 92.5 % trifluoroacetic acid (TFA) with 2.5 % water and 5 % triisopropylsilane (TIPS) for 3 h at room temperature under an argon atmosphere. Addition of phenol to the cleavage solution was found to give no enhancing effect to the cleavage. The resin was filtered off and the peptide precipitated in approximately 40 ml ice-cold tert.-butyl-ethyl-ether, followed by centrifugation for at least 60 min with 4,500 rpm at 4 °C. The precipitated raw product was frozen in liquid nitrogen and dried under vacuum. As several protecting groups tend to stay bound to the peptide in high TFA concentrations, but are rapidly cleaved off at lower TFA concentrations, the raw product was resolubilized in approximately 10 ml 60 % TFA with water and incubated under shaking for 3 h at room temperature. 5 ml of this solution were then applied to a preparative reverse-phase HPLC column (Merck Hibar Lichrospher 100 RP-8 (5 μ M)) using a 10 ml loop. The run was carried out with a gradient from 30 to 50 % acetonitrile with 0.1 % TFA. The foldon peptides were eluted at acetonitrile concentrations of about 42 % to 44 %. Purity of all collected fractions was analyzed by an analytical reverse-phase HPLC (Phenomenex Jupiter 4u Proteo 90A 150 x 4.60 mm or Merck Lichrosorb RP-8 (5 μ M)) using similar run conditions as for the preparative run. The fractions with the desired product were identified by either matrix-assisted laser desorption/ionization-time of flight mass spectrometry (MALDI-TOF) on a ultraflex MALDI TOF/TOF Mass Spectrometer by Bruker Daltonik (Bremen, Germany) or electrospray ionization mass spectrometry (ESI-MS) on a micrOTOF II by Bruker Daltonik (Bremen, Germany). All fractions with a foldon content larger than 90 % were shock-frozen in liquid nitrogen and lyophilized. The dried protein was stored at -20 °C until further usage. Fractions with a purity higher than 99 % were directly used for experiments while all other fractions were subject to an additional purification step using a semi-preparative reverse-phase HPLC column (Phenomenex Jupiter 4u Proteo 90A 250 x 10.00 mm).

3.2. Expression and purification from *E.coli*

The foldon A6S variant was in part obtained by expression in *E.coli* as described before¹⁴³. Foldon wild-type was expressed and purified from *E.coli* using a pET-32a-c(+)-plasmide. The plasmid encodes a fusion protein of the foldon domain with thioredoxin and a His-Tag at

3. Material and methods

the N-terminal end⁹⁵. This construct had been inserted into a BamHI / EcoRI cleavage site. Foldon expressed in *E.coli* and cleaved by thrombin has two additional amino acids at the N-terminus, glycine and serine, which do not effect stability or folding¹⁵⁸. DNA primers were designed for quick-change mutagenesis polymerase chain reaction (PCR) using the programs Restriction Map and Primer X, both available on the Bioinformatics Organization homepage (www.bioinformatics.org). The primer sequences are given in Table 3.1, together with the original sequence on the plasmid.

Foldon variant	Primer sequence
GS-Foldon wt	5'-GATCC GGT TAC ATC CCG GAA GCT CCG CGT G-3'
GS-Foldon P4T (long)	5'-GATC G GGT TAC ATC ACG GAA GCT CCG CGT G-3'
GS-Foldon P4T (short)	5'-GATC G GGT TAC ATC ACG GAA GC-3'
GS-Foldon A6S	5'-GATC G GGT TAC ATC CCG GAA TCT CCG CG-3'

Table 3.1.: Sequences of the PCR primers for quick-change mutagenesis. Bold letters mark the locations of point mutations introduced into the primer.

The first mutation, C to G, is a silent mutation in the linker region between foldon and thioredoxin which removes a BamHI cleavage site within the first Gly-codon of GS-foldon. Upon successful quick-change mutagenesis, only one of the two original BamHI cleavage sites still exists, thus providing a possibility to test successful completion of the quick-change PCR by performing a BamHI cleavage. The second mutation, C to A or G to T, lead to the desired mutations. The primer pair for GS-foldon A6S was ordered from BioSpring (Frankfurt am Main, Germany). The primers were solubilized in MilliQ H₂O to yield a 200 μ M stock solution, which was then diluted to a 4 μ M solution for further usage. The mixture for the quick-change PCR is given in Table 3.2.

5 μ l of the mixture were taken prior to the PCR as a sample for later analysis. The PCR was performed on a PEQLAB Biotechnologie (Erlangen, Germany) Advanced Primus 25 PCR machine with the temperature program given in Table 3.3.

After completion of the PCR, a DpnI cleavage was carried out to digest all template DNA and thus prevent transformations of the foldon wild-type plasmid during heat-shock. For the DpnI cleavage, 2.5 μ l of a 10x restriction enzyme buffer (Promega Buffer B, 60 mM Tris-HCl (pH 7.5), 500 mM NaCl, 60 mM MgCl₂ and 10 mM DTT) and 1 μ l of a 10 U/ μ l DpnI solution

Volumes	Compound
14.5 μl	MilliQ H_2O
2.5 μl	10x Rpx-buffer
1.0 μl	template (unknown concentration)
2.5 μl	forward primer GS-foldon A6S (4 μM)
2.5 μl	reverse primer GS-foldon A6S (4 μM)
1.0 μl	dNTP mix (10 mM for each nucleotide)
1.0 μl	Pfu polymerase (2.5 U/ μl)
25.0 μl	total volume

Table 3.2.: Mixture for the quick-change mutagenesis PCR.

Temperature	Time	Cycles
95 °C	5 min	1x
50 °C	1 min	16x
68 °C	16 min	
95 °C	1 min	
50 °C	1 min	1x
68 °C	20 min	
8 °C	∞	

Table 3.3.: Temperature program using a PEQLAB Advanced Primus 25 for the quick-change mutagenesis PCR.

were added to the PCR solution and incubated for 2 h at 37 °C. 5 μl of the solution were taken as a sample for further analysis after the digestion. The PCR and subsequent restriction enzyme cleavages were analyzed on a 1.0 % agarose gel. A heat-shock transformation was carried out with the PCR product into *E.coli* cells of the XL1 strain. 250 μl cell solution were thawed on ice and incubated on ice for 30 min with 3 μl plasmid solution. The heat-shock was performed for 90 s at 42 °C, followed by 2 min on ice. 200 μl LB-medium were added and the cell solution incubated for 30 min at 37 °C in a water-bath. Cells were plated on a LB plate with 0.1 mg/ml ampicillin and incubated overnight at 37 °C. 4 colonies were picked off the Amp-plate, each was inoculated in 2.5 ml LB-medium containing 0.1 mg/ml ampicillin and incubated overnight at 37 °C under shaking. Plasmid purification of these overnight cultures was carried out using a QIAprep Miniprep Unit (250) by QIAGEN (Hilden, Germany) for plasmid purification. The success of the PCR was analyzed by a BamHI restriction assay, the pipetting scheme is given in Table 3.4.

3. Material and methods

Volumes	Compound
16.3 μl	MilliQ H ₂ O
2.0 μl	10x restriction enzyme buffer
0.2 μl	acetylated BSA (10 mg/ml)
1.0 μl	PCR product solution
0.5 μl	BamHI, 10 U per μl
20.0 μl	total volume

Table 3.4.: Pipetting scheme for the restriction enzyme test of the PCR product.

The BamHI-cleavage was analyzed by an 1.0 % agarose gel, and samples that did not show a decreased size were used for a heat-shock transformation into *E.coli* BL21 DE3, followed by a plasmid purification as described before. These purifications were again analyzed on a 1.0 % agarose gel and then four samples were chosen for sequencing. Sequencing was carried out by GATC Biotech (Konstanz, Germany) using primers for the T7 promotor. Using one of the successfully mutated vectors, a heat-shock transformation of chemocompetent *E.coli* BL21 DE3 cells was carried out, followed by incubation overnight in four 15 ml falcon tubes. These were used to perform a purification of the verified plasmid. A 50 ml overnight culture was inoculated with one colony from the same agarose plate. This solution was used to inoculate four two liter flasks each containing 500 ml LB-medium 0.1 mg/ml ampicillin. An optical density at 600 nm of $\text{OD}_{600} = 0.1$ was set as starting value. Incubation was done at 37 °C under shaking until an OD_{600} of 0.5 was reached, then expression was induced by the addition of 500 μl 1 M isopropyl β -D-1-thiogalactopyranoside (IPTG), followed by incubation for at least 2 more hours. The cells were harvested by centrifugation at 3,000 rpm, 4 °C using a Sorvall RC6 Plus centrifuge with a Sorvall SLA-3000 rotor. The cell pellets were frozen in liquid nitrogen and stored at -80 °C. For cell opening, the cells were thawed on ice and resuspended in 5 mM imidazole 0.2 M NaCl 20 mM Tris-HCl pH 7.9 buffer (binding buffer). Cell opening was performed using a Basic Model Z microfluidizer from Constant Systems (Kennesaw, GA, USA) and subsequent sonication using a SONOPLUS by Bandelin (Berlin, Germany) to disrupt all DNA strands. The solution was applied to a Ni²⁺ sepharose column with 5 ml column volume, using a Pump P-1 peristaltic pump by GE Healthcare (Chalfont St Giles, UK) with a flow rate of approximately 1 ml/min. The affinity chromatography was run on an ÄKTApurifier by GE Healthcare (Chalfont St Giles, UK) with a flow rate of 1 ml/min. The column was first washed with approximately two column volumes of binding buffer, followed by a wash with 3 % 1 M imidazole 0.2 M NaCl 20 mM Tris-HCl pH 7.9 buffer (elution buffer) for one column volume.

A gradient from 3 to 5 % elution buffer over 4 min was run, followed by washing with 5 % elution buffer for one column volume. Finally, a gradient from 5 % to 35 % elution buffer was performed over 1 h, followed by a column wash with 100 % elution buffer and regeneration in binding buffer. All peaks were analyzed by 16 % SDS-PAGE, and all fractions containing the fusion protein were pooled and dialyzed three times consecutively for a minimum of eight hours per dialysis in a Spectra/Por Dialysis Membrane from Spectrum Laboratories Inc. (Rancho Dominguez, CA, USA) with a 6-8,000 molecular weight cut-off against a 100-fold volume excess of 200 mM Tris-HCl pH 8.4, 150 mM NaCl buffer. After dialysis, the concentration of the fusion protein in solution was determined by molar extinction at 280 nm, using an extinction coefficient of $\epsilon = 22,500 \text{ l} \cdot \text{mol}^{-1} \cdot \text{cm}^{-1}$, which was determined using the ProtParam software on the ExPASy Proteomics Server¹⁵⁹. With a known molecular mass of 17,053.4 Da for the GS-foldon fusion protein, the total amount of protein could be calculated. For the thrombin cleavage, a CaCl_2 stock solution is added to the protein solution to yield a final CaCl_2 concentration of 2.5 mM. A 0.2 U/ml thrombin solution was added to the protein solution to yield a final activity of 0.5 U/mg of protein. The solution was incubated at room temperature under shaking and samples were taken at regular intervals for an analysis by 16 % SDS-PAGE to monitor the cleavage. The cleaved protein solution was again applied to a Ni^{2+} sepharose column with a flow rate of 1 ml/min. As the His-tag is connected to the thioredoxin domain of the fusion protein, the thioredoxin was bound to the Ni^{2+} sepharose column while the GS-foldon A6S was directly eluted while applying the protein solution and during the consecutive washing step with binding buffer. The column was then washed with 100 % elution buffer to elute the bound thioredoxin, followed by regeneration with binding buffer. All peaks were collected and analyzed by 16 % SDS-PAGE and MALDI-TOF. All fractions containing the desired GS-foldon A6S were finally pooled and dialyzed three times against a 100-fold excess of NH_4HCO_3 buffer at decreasing concentrations each time in 3,500 molecular weight cut-off Spectra/Por Dialysis Membrane dialysis bags from Spectrum Laboratories Inc. (Rancho Dominguez, CA, USA). Dialysis was followed by shock-freezing in liquid nitrogen and subsequent lyophilizing. The purified samples were stored at -20°C . Samples that still contained impurities were purified by a preparative HPLC run as described before in chapter 3.1.

3.3. Crystallization of foldon variants and X-ray structure determination

For the crystallization of the foldon variants, stock solutions of each variant with a concentration of 8 to 12 mg/ml in water were prepared. The foldon monomer concentration was thus set to 3 mM, given a mass of approximately 3,100 Da for the 12 used foldon variants. The concentration was determined by absorption at 280 nm of 100-fold diluted peptide solutions using the absorption coefficient $\epsilon_{280} = 8,250 \text{ l} \cdot \text{mol}^{-1} \cdot \text{cm}^{-1}$ ($\epsilon_{280} = 8,480 \text{ l} \cdot \text{mol}^{-1} \cdot \text{cm}^{-1}$ for the foldon W20Nal variant and $\epsilon_{280} = 2,980 \text{ l} \cdot \text{mol}^{-1} \cdot \text{cm}^{-1}$ for the foldon W20H variant). All extinction coefficients were determined using ProtParam¹⁵⁹. The pH-value of the peptide solutions was adjusted to pH 7.0, if possible, as several foldon variants tended to aggregate when exposed to neutral pH values and were only soluble at pH 2.0 or 10.0 at high peptide concentrations. See Table 3.5 for details.

pH 2.0	pH 7.0	pH 10.0
Pro7Nva	Pro4Thr	Pro7Phe
Pro4Thr Pro7Nva	Ala6Val	Pro7Val
Trp20Nal	Ala6Ser	Tyr13Phe
Leu22Phe	Ala6Abu	
	Trp20His	

Table 3.5.: pH values of the different stock solutions of the foldon variants when given to crystallization at a concentration of approximate 8 to 12 mg/ml.

Suitable crystallization conditions for the different foldon conditions were determined using commercially available kits from QIAGEN (Hilden, Germany) to screen for different crystal growth conditions. The kits used were Classics, Classics II, Magic I and Magic II for foldon P4T, pHClear Suite, pHClear Suite II and JCSG+ Suite for foldon L22F and foldon A6Abu, and pHClear Suite, pHClear Suite II, Classics, Classics II and JCSG+ Suite for all other foldon variants. Pipetting was performed by a Phoenix pipetting robot by Art Robbins Instruments (Sunnyvale, CA, USA) to 96-well-plates. Incubation was done at 20.0 °C for at least 24 h, or up to 4 weeks using the sitting-drop method. After this time, for all variants except foldon A6V and foldon P7V, crystals suitable for X-ray diffraction measurements were available. The crystallization conditions of the crystal chosen for X-ray diffraction measurement are given

3.3. Crystallization of foldon variants and X-ray structure determination

in Table 3.6.

foldon variant	crystallization conditions		
Pro4Thr	0.1 M HEPES sodium salt	0.8 M Sodium phosphate 0.8 M Potassium phosphate	pH 7.5
Ala6Abu	2.9 M sodium malonate		pH 5.0
Ala6Ser	0.1 M citric acid	1.0 M sodium chloride	pH 4.0
Pro7Phe	0.1 M Tris-HCl	2.0 M ammonium sulfate	pH 8.5
Pro7Nva	2 M ammonium sulfate	0.1 M sodium acetate	pH 4.6
Pro4Thr Pro7Nva	0.1 M HEPES	10 % (w/v) PEG 6000 5 % (v/v) MPD	pH 7.5
Tyr13Phe	0.1 M sodium acetate	2.0 M ammonium sulfate	pH 4.6
Trp20His	0.2 M Ammonium sulfate 0.1 M Sodium acetate	22 % (w/v) PEG 4000	
Trp20Nal	1.4 m Na/K phosphate		pH 5.0
Leu22Phe	3 M sodium chloride	0.1 M Bis-Tris	pH 5.5
	0.1 M Bicine	2.4 M ammonium sulfate	pH 8.0
	0.1 Tris	20 % (w/v) PEG 6000	pH 8.0
	2.4 M sodium malonate		pH 7.0
	0.1 M citric acid	2.0 M sodium chloride	pH 4.0

Table 3.6.: Crystallization conditions of the crystals chosen for X-ray diffraction measurements of all foldon variants.

X-ray diffraction measurements were carried out in a stream of cold nitrogen to prevent thawing of the crystal, with either 1 M Li₂SO₄ buffer in water or a 20 % v/v glycerol mixture with the crystallization buffer used as a cryo-protectant. Measurements were carried out on a Bruker X-ray diffractometer X8 using a CuK_α rotating anode with a Platinum¹³⁵ detector. The unit cell was determined for each crystal before start of the actual measurement, and the angles and length of radiation for a complete data set were calculated depending on the quality and number of reflexes. As both a NMR- and a X-ray structure of foldon wild-type are known, the molecular replacement method can be applied to solve the phase problem. All data evaluation steps were performed using the CCP4 software compilation^{160,161}. The *R*-factors are determined using Refmac5, and molrep was used to fit the electron density with an already existing foldon wild-type structure. Any faults and missing or truncated amino acids were fixed manually using the Coot program of CCP4, yielding the final pdb-file. Further evaluation of the resulting crystal structure was done using the SWISS Pdb-Viewer¹⁶², while image generation was done using MacPyMOL¹⁶³.

3.4. Fluorescence spectra determination

The fluorescence spectra of the native and the unfolded state at 5 μM monomer concentration of the respective foldon variants were determined in 0.58 M guanidinium hydrochloride (GdmCl) 20 mM sodium cacodylic acid pH 7.0 and 6.4 M GdmCl 20 mM sodium cacodylic acid pH 2.08, respectively. The spectra were determined using a Aminco Bowman Series 2 Luminescence Spectrometer by American Instrument Company (Silver Spring, MD, USA) at 20.0 °C. The excitation wavelength was set to 280 nm with a band-pass width of 1 nm. The emission spectrum was measured from 280 to 430 nm in 1 nm steps with 1 s per step using a 16 nm band-pass width. The fluorescence spectra of the different buffers were determined in the same way and these spectra subtracted from the peptide traces. The obtained traces are then normalized to the trace of foldon wild-type *U*-state.

During refolding of foldon wild-type, a burst-phase intermediate had been observed⁹⁵. Due to its transient nature, the spectrum of this intermediate was determined from stopped-flow refolding kinetic experiments using a SX20 Stopped-Flow Spectrometer by Applied Photophysics (Surrey, UK). The stopped-flow device was set up to enable a 10:1 dilution of the foldon stock solutions upon mixing with 20 mM sodium cacodylic acid pH 11.4 buffer. The refolding reaction was initiated by dilution to refolding conditions with 0.58 M GdmCl pH 7.0. Syringes used were 2.5 ml and 250 μl with the 250 μl syringe connected to the lower inlet of the mixing chamber to avoid mixing artifacts. This can occur if the solution with the higher density is inserted from the upper inlet of the mixing chamber. A 2.5 ml stop syringe was used and the volume set to approximately 110 μl . A 55 μM foldon stock solution in 6.4 M GdmCl 20 mM sodium cacodylic acid pH 2.08 was prepared as described before and diluted against 20 mM cacodylate pH 11.6. The photo-multiplier voltage was set automatically and then lowered by hand to the nearest decade. Traces were recorded in a logarithmic setting for 10 s with 1,000 data points. An excitation wavelength of 280 nm with a band-pass width of 4 nm was used, while the emission wavelength was varied between 280 and 430 nm with the band-pass width set to 12 nm, as can be seen in Table 3.7. At least 5 traces were measured for each emission wavelength to gain a better signal-to-noise ratio. Additionally, the spectrum of the buffer was measured in a continuous scan from 280 to 430 nm in 1 nm steps, measuring for 5 s at each emission wavelength.

emission wavelengths, nm									
280	285	290	295	300	303	306	309	312	315
318	320	322	324	326	328	330	332	334	336
338	340	342	344	346	348	350	352	354	356
358	360	362	365	368	371	374	377	380	385
390	395	400	410	420	430				

Table 3.7.: Emission wavelengths used to determine the fluorescence spectrum of the foldon burst-phase intermediate

The stopped-flow device does not possess an instrumental correction file. In order to correct and normalize the burst-phase intermediate spectrum obtained in the stopped-flow to the spectra obtained in the fluorimeter, the emission spectra of a 10 μM N-acetyl-(L)-tryptophanamide (NATA) solution was determined both in the stopped-flow and the fluorimeter. The concentration of NATA was determined by absorption at 280 nm using an extinction coefficient of $5,960 \text{ l} \cdot \text{mol}^{-1} \cdot \text{cm}^{-1}$ ¹⁶⁴. The stopped-flow spectrum of NATA could be used to calculate a correction file for the stopped-flow, which allows the correction of the burst-phase intermediate.

3.5. Denaturant transitions

Guanidinium chloride transitions

A stock solution of the chosen foldon variant in 20 mM sodium cacodylic acid, pH 7.0 was prepared and its concentration determined by absorption using the same absorption coefficients as in 3.3. Solutions with the same peptide concentration in varying GdmCl concentrations in 0.2 M steps were prepared by mixing of a 8 M and a 0 M GdmCl solution in 20 mM sodium cacodylic acid pH 7.0.

For each sample, the fluorescence emission at 320 nm upon excitation at 280 or 295 nm, depending on the monomer concentration, was measured and averaged over 60 s, using an Aminco Bowman Series 2 Luminescence Spectrometer by American Instrument Company (Silver Spring, MD, USA). An excitation wavelength of 273 nm and an emission wavelength of 303 nm were used for the foldon variants W20Nal and W20H. The measurements were carried out at 20.0 °C under stirring in a quartz crystal cuvette with a 10 mm path length at a right angle to the excitation beam. The bandpass for excitation and emission were 4 nm

3. Material and methods

and 8 nm, respectively. The voltage of the photo-multiplier was set by auto-detection at 60 % of the maximum value of the first sample with 0 M denaturant concentration. Along with fluorescence, the CD absorption for each sample was detected at 228 nm over 60 s, using a Circular Dichroism Spectrometer Model 410 by Aviv Biomedical Inc. (Lakewood, NJ, USA) at 20.0 °C. These measurements were carried out in a 2 mm or 5 mm quartz crystal cuvette, depending on the concentration, and all data points averaged.

The exact GdmCl concentration of each solution was determined by measuring the refractive indexes using a Reichert Inc. (Depew, NY, USA) AR7 Series Automatic Refractometer. With equation 3.1, the final GdmCl concentration of each sample was calculated:

$$y = (x - a) \cdot 57.15 + (x - a)^2 \cdot 38.68 - (x - a)^3 \cdot 91.60 \quad (3.1)$$

where x is the measured refractive index for a given probe, y is the corrected GdmCl concentration and a is the refractive index of the solution without denaturant³⁶.

The averaged data points of both the fluorescence and CD measurements are plotted against the final GdmCl concentrations and the resulting curves analyzed assuming a two-state transition between the unfolded monomers U and the native trimeric state N . The following derivation is taken from⁹⁵:



According to the law of mass action, this gives:

$$K_{eq} = \frac{[U]_{eq}^3}{[N]_{eq}} \quad (3.3)$$

The total monomer concentration $[M]_0$ is at each time-point the sum of the unfolded monomer U and three times the native trimer N concentration.

$$[M]_0 = [U] + 3[N] \quad (3.4)$$

The fractions of foldon monomers in both the native and the unfolded state, f_N and f_U , are given by the equations:

$$f_U = \frac{[U]}{[U] + 3[N]} = \frac{[U]}{[M]_0} \quad (3.5)$$

$$f_U = \frac{3[N]}{[U] + 3[N]} = \frac{3[N]}{[M]_0} = 1 - f_U \quad (3.6)$$

By rearranging and inserting these equations into equation 3.3 one obtains:

$$K_{eq} = \frac{U^3}{N} = \frac{3f_U^3[M]_0^2}{1 - f_U} \quad (3.7)$$

This accumulates in an equation for the monomer-trimer transition similar to the two-state transition fit for monomeric proteins proposed by Santoro and Bolen⁴⁰:

$$S(x) = S_N(x) - \frac{S_N(x) - S_U(x)}{3[M]_0} \cdot \left(\sqrt[3]{\exp \frac{-\Delta G^0(H_2O) + m_{eq}x}{RT} \left(\frac{9}{2}[M]_0 + \sqrt{D} \right)} + \sqrt[3]{\exp \frac{-\Delta G^0(H_2O) + m_{eq}x}{RT} \left(\frac{9}{2}[M]_0 - \sqrt{D} \right)} \right) \quad (3.8)$$

with

$$D = \left(\frac{9}{2} \right)^2 [M]_0^2 + \exp \frac{-\Delta G^0(H_2O) + m_{eq}x}{RT} \quad (3.9)$$

All spectroscopic data obtained either by circular dichroism absorption or fluorescence for one foldon variant were globally fitted to equation 3.8. The native and unfolded baselines are fitted to each trace individually, but $\Delta G^0(H_2O)$ and m_{eq} were fitted globally. For each foldon variant the GdmCl transition was measured at at least two different foldon monomer concentrations to achieve a higher fidelity of the fit. Additionally, this allows a fit even if one baseline is only weakly defined. All fitting was performed using the pro Fit program by Quantum Soft (Uetikon am See, Switzerland). All traces were normalized to the fraction of native foldon f_N .

Urea transitions

For urea transitions, a similar procedure as described for the GdmCl transitions was carried out. The chosen foldon variants were resolubilized in a 10 mM sodium cacodylic acid buffer

3. Material and methods

pH 7.0, with the foldon concentration determined by absorption with absorption coefficients given in 3.3. The stock solution was mixed in an appropriate ratio with a 10 M urea 10 mM sodium cacodylic acid pH 7.0 buffer to yield urea concentrations from 0 M to 9.25 M in 0.25 M steps. The exact urea concentration y was determined by measuring the refractive indexes, using the equation:

$$y = (x - a) \cdot 117.66 + (x - a)^2 \cdot 29.75 - (x - a)^3 \cdot 185.56 \quad (3.10)$$

where x represents the measured refractive index for a given probe, y the corrected urea concentration, and a the refractive index of the solution without denaturant³⁶. The data obtained from the measurements of both the changes in tryptophan fluorescence or CD absorption were treated as described before for GdmCl. Ideally, the values for $\Delta G^0(H_2O)$ for both the GdmCl and urea transition should be identical while the m_{eq} -value for the GdmCl transition should be larger than that of the urea transition due to the stronger denaturing effect of GdmCl.

Global denaturant transitions

Due to the low stability of the majority of foldon variants, the native baselines of both the GdmCl and urea denaturant transitions were not well-defined. Thus a global fit of both denaturant transitions was performed, resulting in a global folding stability $\Delta G^0(H_2O)$ for the respective foldon variant and m_{eq} -values for GdmCl and urea ($m_{eq}(\text{urea})$ and $m_{eq}(\text{GdmCl})$). The fitting procedure is carried out as described before, and the traces were normalized with the newly obtained fitting coefficients.

Denaturant transitions in the presence of sulfate anions

The foldon variants W20Nal and W20H are too unstable under the conditions used before for GdmCl and urea denaturant transitions, thus transitions of both variants were not possible under the previously described buffer conditions. Sulfate is known to have a stabilizing effect on the $\Delta G^0(H_2O)$ of proteins, therefore the transitions were carried out in the presence of 0.5 M sodium sulfate. To be able to compare these data from the variants with foldon wild-type, a GdmCl transition of foldon wild-type in the presence of 0.5 M sodium sulfate was performed as well. A foldon wild-type transition using urea was not possible under the stabilizing conditions as urea is not strong enough as denaturing agent.

Sodium sulfate lowers the solubility of both GdmCl and urea, thus the transitions could only be performed at final GdmCl concentrations lower than 6 M and urea concentrations lower than 8 M. As foldon wild-type is very stable under the previously described buffer conditions and its stability is even further increased by the addition of sodium sulfate, GdmCl transitions of foldon wild-type could only be carried out at 2 and 4 μM foldon monomer concentration to ensure that a *U*-state baseline could still be detected. Measurements and evaluations of the transitions were carried out as described before. Additionally, global fits of both the GdmCl and urea data were carried out as well, as described previously.

3.6. Equilibrium *N*-test of foldon in guanidinium chloride

Denaturant transitions of all destabilized foldon variants in GdmCl and urea yielded differing free folding enthalpies $\Delta G^0(H_2O)$. To check whether this finding is an artifact due to ill-defined native baselines in the GdmCl transitions or a real result, an equilibrium *N*-test was carried out. As the destabilized foldon variants do not reach 100 % *N* at foldon monomer concentrations below 200 to 500 μM , the studied foldon variants were first solubilized in 500 mM sodium sulfate 20 mM sodium cacodylic acid pH 7.0 to ensure that all foldon monomers are present in the native state. The foldon monomer concentration was determined by absorption at 280 nm using the respective extinction coefficient of the foldon variants given in 3.3. The foldon solution in sodium sulfate was unfolded by 6fold dilution with 7 M GdmCl 20 mM sodium cacodylic acid pH 7.0 using a stopped-flow device. Unfolding was carried out at 20.0 °C. Intrinsic tryptophan fluorescence of the foldon variants, which was used to monitor the unfolding reaction, was excited at 280 nm with a band-pass width of 4 nm. Fluorescence emission was detected at 320 nm with a band-pass width of 12 nm. After the amplitude of 100 % native foldon unfolding was determined, the same was done for the same foldon monomer concentration in 20 mM sodium cacodylic acid, pH 7.0, with GdmCl concentrations reaching from 0.0 to 2.0 M GdmCl in 0.2 M steps, and additionally at 2.4, 2.8, and 4.0 M GdmCl. Unfolding of the native foldon was initiated by 6fold dilution with 7 M GdmCl 100 mM sodium sulfate 20 mM sodium cacodylic acid pH 7.0. The sodium sulfate in the unfolding buffer ensured identical unfolding conditions for both foldon solution with and without sodium sulfate. Variation of the final GdmCl concentrations due to the varying starting conditions did not have an impact on the unfolding reaction. The unfolding traces were fitted in

pro Fit using a single-exponential equation, and the determined amplitudes normalized with the amplitude of the foldon unfolding experiment with 500 mM sodium sulfate. The normalized fractions of natively folded foldon were finally plotted against the GdmCl concentration and allowed direct comparison with the GdmCl transition determined by intrinsic tryptophan fluorescence or circular dichroism absorption under equilibrium conditions. The exact GdmCl concentration was determined by measuring the refractive indexes of all solutions, followed by calculation of the exact GdmCl concentration using equation 3.1, as described in detail in Chapter 3.5.

3.7. Refolding kinetics

Guanidinium chloride refolding kinetics

For guanidinium chloride refolding experiments of the various foldon variants, stock solutions in 6.4 M GdmCl 20 mM cacodylic acid pH 2.08 were prepared with concentrations varying from 1.1 mM to 11 μ M. The concentrations of the solutions were determined by absorption at 280 nm, using an extinction coefficient of $\epsilon = 8,250 \text{ l} \cdot \text{mol}^{-1} \cdot \text{cm}^{-1}$. The solutions were incubated overnight at room temperature to ensure complete unfolding of the samples. Refolding kinetics were determined either by stopped-flow measurements on a SX20 Stopped-Flow Spectrometer by Applied Photophysics (Surrey, UK) or by hand-mixing and subsequent fluorescence change measurement using an Aminco Bowman Series 2 Luminescence Spectrometer by American Instrument Company (Silver Spring, MD, USA). Time traces were obtained by detecting the change in intrinsic tryptophan fluorescence at 320 nm after excitation of the sample at 280 nm, or 295 nm if the foldon monomer concentration exceeds 10 μ M to avoid inner filter effects. The stopped-flow device was set up to enable a 10:1 dilution of the foldon stock solutions upon mixing with 20 mM sodium cacodylic acid pH 11.4 buffer. The refolding reaction was initiated by dilution to refolding conditions with 0.58 M GdmCl pH 7.0. Syringes used were 2.5 ml and 250 μ l with the 250 μ l syringe connected to the lower inlet of the mixing chamber to avoid mixing artifacts. This can occur if the solution with the higher density is inserted from the upper inlet of the mixing chamber. A 2.5 ml stop syringe was used and the volume set to approximately 110 μ l. The mixed solution was excited at 280 nm or 295 nm, as stated above. The band-pass of the monochromatic excitation light was set to 4 nm. Single-wavelength detection of the tryptophan fluorescence was carried out at 320 nm with the band-pass set to 12 nm. Measurements were carried out in logarithmic detection mode, the

photo-multiplier voltage set automatically and then lowered by hand to the nearest decade. Six or more traces were averaged for each concentration to gain a better signal-to-noise ratio. Detection times were set to 5 times the half-life of the reaction which was detected from the first trace. It ranged from 200 to 1,000 s in the stopped-flow, with 200 s being the lower limit due to the proline *cis-trans* isomerization reaction, while 1,000 s is the longest time suitable for the stopped-flow. Otherwise diffusion effects and photo-bleaching can produce artifacts. For reaction times longer than 1000 s, measurements were carried out by hand-mixing and subsequent fluorescence measurement in a fluorimeter. Excitation and emission wavelengths were set up as for the stopped-flow fluorescence measurements. Band-pass width were set to 1 nm for excitation and 8 nm for emission. The photomultiplier voltage was set to 60 % of the maximum signal by using a previously mixed sample for auto-detecting. A 10 mm quartz-crystal cuvette was used for measurement. 100 μ l of foldon stock solution in 6.4 M GdmCl 20 mM sodium cacodylic acid pH 2.08 are mixed in the cuvette with 1,000 μ l of refolding buffer (20 mM sodium cacodylic acid pH 11.4) to yield the same refolding conditions as in the stopped-flow device. Refolding was measured under constant stirring. While the fluorescence change was measured constantly for the first 600 s, the fluorescence intensity was only measured for 10 s at different time points during refolding. The different intervals chosen are given in Table 3.8. Measurements were carried out until equilibrium was reached, which was indicated by a constant fluorescence signal.

Intervals	First and last data point
Constant	0 to 10 min
2 min	10 min to 22 min
3 min	22 min to 25 min
5 min	25 min to 40 min
10 min	40 min to 100 min
20 min	100 min to 180 min
30 min	180 min to ∞ min

Table 3.8.: Time intervals chosen for the measurement of the fluorescence intensity for 10 s of the foldon refolding kinetics.

As the signal-to-noise ratio in the fluorimeter is far better than in the stopped-flow device, only one or two measurements were required to get signal-to-noise ratios good enough for further analysis. The curves of both the stopped-flow and the hand-mixing refolding kinetics measurement were averaged and normalized by fitting the curves to a three- or four-exponential

3. Material and methods

equation. The obtained amplitudes were used to normalize the starting fluorescence to 0 and the end-value to 1.

Urea refolding kinetics

Refolding kinetics in urea were initiated by a 6-fold dilution of a foldon stock solution in 6 M urea, 10 mM glycine, pH 2.08 against a 5-fold excess of 10 mM sodium cacodylic acid, pH 11.6. Glycine was chosen as buffer substance due to its lower pK_a -value of 2.34 of its carboxyl group. Thus a lower amount of acid was needed to adjust the pH-value of the solution, resulting in a lower ionic strength of the solution. Mixing and measurements were carried out as described before for GdmCl. A 7 M urea solution was prepared and its final urea concentration determined by measuring the refractive index as described in chapter 3.5. The solution was then adjusted to 6 M urea by dilution with 10 mM glycine, pH 2.08 buffer. The stopped-flow setup was similar to that described GdmCl refolding experiments, but using a 500 μ l syringe for the foldon stock solution, which was again connected to the lower inlet of the mixing chamber to avoid mixing artifacts.

3.8. Reaction order determination

Reaction order determination by measuring the half-life times

The reaction order of the association reaction was determined from half-life times of the refolding traces. This method allows the determination of reaction orders of reactions with an unknown reaction type, even broken reaction orders. Additionally, in foldon it allows the determination of rate-limiting concentration-independent folding steps by deviation from the linear slope, especially for higher peptide concentrations⁹⁵. After normalization of the refolding traces as described in Chapter 3.7, the half-life time of the concentration-dependent phases was determined. If the proline *cis-trans* isomerization reaction was detectable, the amplitude change caused by this reaction was excluded from the analysis, as the *cis-trans* isomerization is concentration-independent and would thus distort the obtained reaction order.

The common equation for the half-life time for reactions orders larger than one is²⁸:

$$t_{1/2} = \frac{1}{k[M]_0^{n-1}} \quad (3.11)$$

By rearranging and logarithmizing this equation, one gets:

$$\log t_{1/2} = -\log k - (n - 1) \log [M]_0 \quad (3.12)$$

The common logarithm of the obtained half-life times was plotted against the foldon monomer concentration and the graph was fitted using equation 3.12 with n being the reaction order of the reaction.

Reaction order determination from the initial slope

The reaction order for all refolding kinetics in both denaturants was also determined from the initial slope, additionally to the reaction order determined from the half-life times. To gain a higher resolution for the first seconds of the refolding traces, denaturant refolding experiments were carried out in the stopped-flow. The refolding was measured for 1 s for refolding experiments which had previously been performed in the stopped-flow, or for 10 s for refolding experiments previously performed by hand-mixing in the fluorimeter. The refolding traces had to be adjusted to the final foldon monomer concentration which was able to form native foldon. The stability of the respective foldon variant at the used refolding buffer condition $\Delta G^0(denat)$ was first calculated from the determined global folding stability $\Delta G^0(H_2O)$ and the m_{eq} -values for GdmCl or urea. This value was then used in a modified version of equation 3.8, which is:

$$y = 1 - \frac{1}{3x} \cdot \left(\sqrt[3]{\exp \frac{-\Delta G^0(denat)}{RT} \left(\frac{9}{2}x + \sqrt{D} \right)} + \sqrt[3]{\exp \frac{-\Delta G^0(denat)}{RT} \left(\frac{9}{2}x - \sqrt{D} \right)} \right) \quad (3.13)$$

with

$$D = \left(\frac{9}{2} \right)^2 x^2 + \exp \frac{-\Delta G^0(denat)}{RT} \quad (3.14)$$

with the fraction of foldon monomer forming native trimer y , the total foldon monomer concentration x , and the values for $S_N(x)$ and $S_U(x)$ set to 1 and 0, respectively. The refolding traces were adjusted to the final concentration of foldon monomer which is able to fold into

3. Material and methods

native trimer under the given refolding conditions. The linear region of the short refolding traces at the beginning of the refolding reaction was fitted to a linear equation. The slope of the linear fit was adjusted for the spectroscopic parameter of the dimer, which was determined from the global fit of all data. The natural logarithm of the adjusted slope was plotted against the common logarithm of the foldon monomer concentration. A linear fit of this plot yields the reaction order near the time-point zero as the slope and the natural logarithm of the dimerization rate constant as the y-intercept.

3.9. Unfolding kinetics

Guanidinium chloride unfolding kinetics

Unfolding of the various foldon variants in GdmCl was carried out using the stopped-flow apparatus or fluorimeter mentioned before. For foldon wild-type, two different stock solutions (30 μM and 60 μM) were prepared in 0.58 M GdmCl 20 mM sodium cacodylic acid pH 7.0 to check for a concentration-dependence in the unfolding reaction and to measure the GdmCl-dependence of the unfolding rate. The stock solutions for all other foldon variants had a concentration of 300 μM in 0.58 M GdmCl 20 mM cacodylate pH 7.0. The stopped-flow was set up to allow a 6-fold dilution of the foldon-containing solution against GdmCl buffer 20mM sodium cacodylic acid pH 7.0 with varying GdmCl concentrations, ranging from 2 M to 8 M GdmCl; the final concentrations of foldon wild-type were thus 5 μM and 10 μM , while the final concentration of all other foldon variants was 50 μM , except for the double-mutant foldon P4T P7Nva, where the final foldon concentration was 25 μM . Measurements were carried out at 20.0 °C with an excitation wavelength of 280 nm for foldon wild-type and 295 nm for all other foldon variants to avoid inner-filter effects, using a band-pass of 4 nm. The emission wavelength was set to 320 nm with a band-pass of 12 nm. The fluorescence decrease upon unfolding was detected until equilibrium was reached, which could vary from 10 s to 1000 s, depending on the foldon variant and the final GdmCl concentration. Measurements were carried out in a logarithmic mode with 2000 data points per measurements if the reaction lasted longer than 500 s; otherwise 1000 data points were used. At least 10 traces were averaged for each GdmCl concentration. For unfolding experiments requiring longer unfolding times than 1,000 s, measurements were carried out in the fluorimeter. The solutions were mixed manually and the course of the reactions followed by the change in tryptophan fluorescence, using an excitation wavelength of 295 nm and a band-pass of 4 nm, while the emission wave-

length was set to 320 nm with a band-pass of 8 nm. The photo-multiplier voltage was set to 60 % of the maximum signal. All measurements were carried out at 20.0 °C until equilibrium was reached. The course of the reaction was followed using the same time intervals as for the foldon refolding experiments, given in Table 3.8.

Urea unfolding kinetics

Urea unfolding was carried out corresponding to the procedure described in Chapter 3.9. The foldon stock solutions were prepared in 1 M urea 10 mM sodium cacodylic acid pH 7.0 and mixed with a 6-fold excess of buffer containing varying urea concentrations from 3 M to 9 M and 10 mM sodium cacodylic acid pH 7.0. All foldon stock solutions had a foldon monomer concentration of 300 μM , resulting in a final foldon monomer concentration of 50 μM after mixture. As unfolding in urea was seen to be slower than in GdmCl, urea unfolding experiments of some foldon variants (foldon A6Abu, foldon L22F) needed to be performed by hand-mixing in a fluorimeter. These measurements were carried out as described before for GdmCl. Measuring the apparent rate constant k_{app} by rapid dilution of denatured protein into varying final concentrations of denaturant and plotting the common logarithm of the individual rate constants against the corresponding denaturant concentrations yields a V-shaped curve, often termed chevron plot^{68,69}, for some foldon variants. The two arms of this plot correspond to the refolding and unfolding reaction of the protein. The linear slope implies that not only the thermodynamic stability of a protein, ΔG^0 , but also the rate constants k_f and k_u depend linearly on the denaturant concentration. By a linear fit of the arms and extrapolation to the ordinate, the rate constants of the folding and unfolding reaction at zero molar denaturant can be determined. Thus, it is possible to define the following equations for both the folding and unfolding rate constant⁵⁷:

$$\ln k_f = \ln k_f^{H_2O} - \frac{m_{k_f}}{RT} \cdot [x] \quad (3.15)$$

$$\ln k_u = \ln k_u^{H_2O} + \frac{m_{k_u}}{RT} \cdot [x] \quad (3.16)$$

Combining the two equations 3.15 and 3.16 yields:

$$\ln k_{app} = \ln (k_f + k_u) = \ln \left[k_f^{H_2O} \exp \left(\frac{-m_{k_f}}{RT} \cdot [x] \right) + k_u^{H_2O} \exp \left(\frac{m_{k_u}}{RT} \cdot [x] \right) \right] \quad (3.17)$$

3.10. Rate constant determination

To determine the changes in the rate constants of the association reaction of the different foldon variants and compare these with the wild-type data published before⁹⁵, the obtained kinetic traces were globally fitted using KinTek (KinTek Corporation, Austin, TX, USA). It allows model building, followed by numerical fitting^{165,166}. Before global fitting could be performed, a consecutive procedure of 2 normalization steps had to be carried out using KinTek. This was necessary as the first rough fitting using a three- or four-exponential equation described before in Chapter 3.7 was not yielding exact results, as the normalization could have led to fitting artifacts and a three- or four-exponential equation is only a rough approximation of the signal course. As the first step, the original data of the foldon refolding experiments in both guanidinium chloride and urea were stored for each foldon monomer concentration separately as a .txt-file, after ensuring that the offset of the refolding trace was clearly positive. The sigma-value for each trace was determined by fitting each trace to a three-exponential equation with an offset to allow future error calculation. The traces were then globally fitted to the corresponding association model of the individual foldon variant. All models used are shown in Figure 3.2.

Five different folding models were used for rate constant determination to describe the data of the different foldon variants, yielding the rate constants for each reaction step. The model used for foldon wild-type⁹⁵ to evaluate the folding and association in guanidinium chloride was used in two different variants: For faster refolding experiments, where the concentration-independent rearrangement step from the trimer in all-*trans* conformation of the Xaa-Pro peptide bonds to the native state is observable, this step was included into fitting, while this step was omitted for slower refolding variants and conditions. The second model used, referred to as the "new" model throughout this study, allows formation of dimers and trimers even if one or more of the monomers has Xaa-Pro peptide bonds in *cis*-conformation. The last rearrangement step from the trimer to the native state *N* is only possible from the trimer in all-*trans* conformation. This model was used before for foldon wild-type refolding studied in urea¹⁵⁰, and was again used in two different variants, once with and once without the last rearrangement step, depending on the refolding speed. In general, the last rearrangement step is needed for refolding in guanidinium chloride, while refolding experiments in urea are too slow to allow detection of this folding step. The linear model finally was used for foldon mutants at position 7. No *cis-trans* isomerisation reaction is needed in this model, and as refolding of

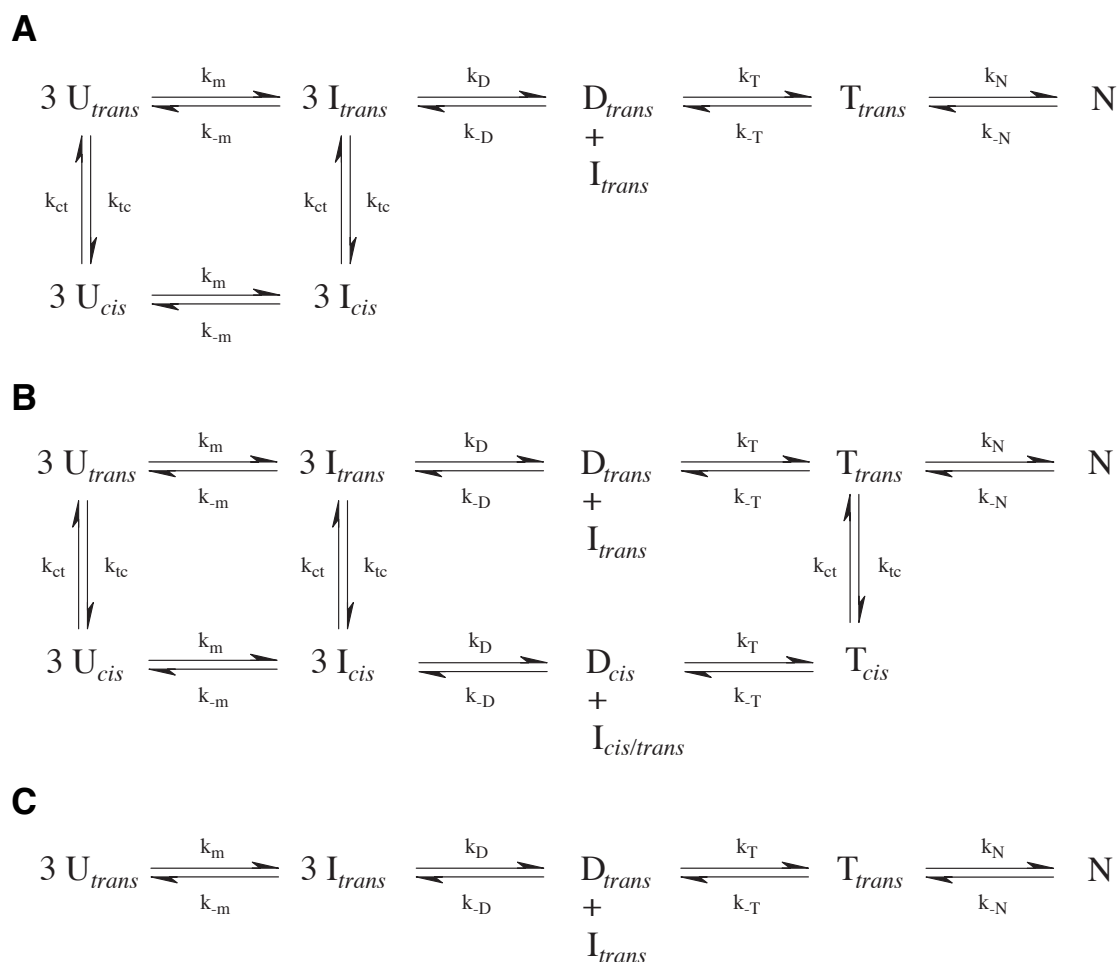


Figure 3.2.: The foldon association mechanisms with all observable reactions steps used for global fitting of the refolding kinetic traces. The folding step from the trimer in all *trans* conformation of its Xaa-Pro peptide bonds is always included, although this step is omitted in several fits. A) The foldon association model used before to evaluate the refolding reaction in GdmCl⁹⁵. B) The "new" folding model which allows productive dimer and trimer formation of foldon monomers with a *cis* conformation of a Xaa-Pro peptide bond. C) The "linear" model which does not include any *cis-trans* isomerization steps of Xaa-Pro peptide bonds.

3. Material and methods

mutants at this position is extremely slow, the last rearrangement could not be detected either. The linear model hence only consists of two consecutive dimerization reactions. Independent fluorescence parameters were assigned to all non-related species: Monomers M , dimers D , trimers T , and the native state N . There was no discrimination between species with different *cis-trans* conformations of the Xaa-Pro peptide bonds. The final signal course of each fluorescence refolding trace is constituted from the sum of the contribution of each individual species at the varying time points. Each individual contribution is itself dependent on both the concentration of the species at the given time point and the fluorescence parameter of the species. The original fluorescence traces were imported into KinTek. In the first fitting step, the fit traces were constituted from the sum of the different species contributions, multiplied with a correction factor, and an offset was added to the trace. Both the correction factor and the offset were independent for each foldon monomer concentration. Dimerization and trimerization steps were considered to have the same rate constant, regardless if the Xaa-Pro peptide bonds of the educts were in the *cis* or *trans* conformation.

A first global fit was carried out, which already gave a hint on the later results, as the fit was only usable for adjustments of the fluorescence signal traces if all parameters were already close to the final results. Rate constants for individual folding and association steps known from previous experiments and evaluations were inserted and held constant. This were the dimerization rate constant k_D and the rate constant for the unfolding of the native state, k_{-f} . The rate constant for Xaa-Pro *cis-trans* isomerization was determined from the last concentration-independent phase in refolding experiments, if possible. The rate of this phase corresponds to k_{ct} . The known fraction of *cis*-proline residues⁸¹, depending on the N-terminal residue, allowed the calculation of k_{tc} . The rate constant k_{ct} was also used for the reaction step from T_{cis} to T_{trans} , while the reverse reaction was set to 0 in this case. The favored interactions of the N-terminus in all-*trans* conformation were considered to prevent rearrangement of the Xaa-Pro peptide bond into the *cis*-conformation. T_{trans} can thus only fold into the native state N or dissociate into one monomer and one dimer, both in the all-*trans* conformation. The offset determined for each trace individually was subtracted from each corresponding original refolding trace. A second global fit was carried out with these modified traces, and the fluorescence signal determined after the reaction has reached its equilibrium. This final baseline was adjusted for each foldon monomer concentration to represent the concentration of monomer which folds into the native state at the given monomer and denaturant concentration, as described in Chapter 3.8. This inserted the folding stability into the analysis, allowing the correct

determination of the rate constants. These corrected fluorescence traces could finally be used for global fitting, yielding the final rate constants. As the fitting so far was only done using set values for the rate constants mentioned above, the last global fitting step was repeated with all rate constants run freely to minimize possible errors inserted from the previous evaluations. The individual traces were normalized and the fitting traces exported to allow a comparison with other foldon variants.

3.11. Interrupted refolding kinetics

Interrupted refolding experiments were carried out on a stopped-flow device set up for double-jump experiments. A 1.1 mM foldon monomer stock solution in 6.0 M urea 10 mM sodium cacodylic acid buffer pH 2.04 was prepared on the day before experiments; the concentration was determined by absorption at 280 nm, using the extinction coefficient of the respective foldon variant given in 3.3. The peptide solution was incubated over night at room temperature. The interrupted refolding experiment was started as previously described by six-fold dilution of the foldon solution in 6 M urea. After varying time steps, reaching from 30 ms to 999 s, folding was interrupted by rapid mixture in a 1:6 ratio of the refolding buffer with 8 M urea 10 mM sodium cacodylic acid pH 7.0, which caused an unfolding of all species except the monomer. Experiments were carried out at 20.0 °C, the excitation wavelength was set to 295 nm with a band-pass width of 4 nm. The emission wavelength was 320 nm, using a band-path width of 12 nm. At least 10 traces were taken for each refolding time span, with approximately 40 different time spans evenly distributed over the complete range for the refolding experiments. Evaluation of the obtained unfolding traces was carried out as described before in Chapter 3.9. The amplitudes of all time spans were plotted against the refolding time, and the graph fitted with a two-exponential equation. The final baseline was taken as the final amplitude of the reaction. Alternatively, it would have been possible to take the amplitude of the longest refolding reaction, but the described method was chosen for higher fidelity as the amplitude determined for the longest refolding time did not show the largest amplitude of all interrupted refolding experiments and the amplitudes for longer time-spans showed a high scattering of the amplitude. All amplitudes were divided by the final amplitude value of the reaction determined from the fit, and these values again plotted against the refolding time span, giving the fraction of formed native trimer at each chosen time point.

4. Results

4.1. Design of foldon variants

The hydrophobic cluster of foldon consists of the amino acids proline at position 4 and 7, alanine at position 6, tyrosine at position 13, tryptophan at position 20 and leucine at position 22. The impact of each amino acid in the hydrophobic cluster on the association reaction of foldon is probed by inserting at least one mutation at each position. Evaluation of the position, the interactions of the natural amino acid and possible impacts of the mutation are considered for each amino acid exchange. In total, 11 different single-point mutations and one double-mutant are studied. An overview is given in Table 4.1. The rationality for each amino acid exchange is discussed below.

position 4	position 6	position 7	position 13	position 20	position 22
P4T	A6S	P7F	Y13F	W20His	L22F
	A6V	P7V		W20Nal	
	A6Abu	P7Nva			
double mutant	P4T P7Nva				

Table 4.1.: All foldon variants studied throughout this work.

For the proline at position 4, threonine is chosen as replacement. The proline at this position interacts chiefly with the hydrophobic cluster by a hydrogen bond formed between the backbone carbonyl-group of residue 4 and the N₁ atom of the indole group of tryptophan at position 20. The backbone is almost completely solvent-accessible due to the extended poly-proline type II helix. The exchange of the proline to threonine is thus proposed to have two consequences: First, by the breaking of the heterocyclic side chain, the backbone gains conformational freedom. This allows for a direct evaluation of the impact of the backbone conformational freedom on assembly and stability. As *cis-trans* isomerization of at least one Xaa-Pro peptide bond was shown by Güthe et al. to be a rate-limiting step in foldon assem-

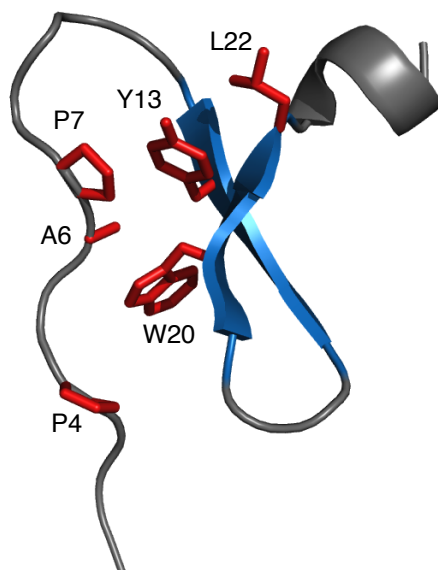


Figure 4.1.: A foldon monomer in the native state fold in a ribbon representation with all amino acids constituting the hydrophobic cluster depicted in red as a stick representation and the amino acid name and position given in single-letter code. The backbone atoms constituting the β -strands are colored in blue, while all other backbone atoms are colored in grey. The figure was produced using MacPyMOL¹⁶³ from the foldon E5R NMR structure (PDB accession number 2KBL)¹⁵¹

bly⁹⁵, the question arose which of the two prolines is responsible for this deceleration, if not both contribute to the isomerization reaction. Second, by the introduction of a hydroxyl group, the solubility of the side chain should get enhanced in contrast to the non-polar pyrrolidine ring of proline. This could result in a higher stability, as the adverse effect of the non-polar side chain pointing towards the solvent is removed.

Three different mutations are synthesized for the alanine at position six. The alanine interacts strongly with the surrounding amino acids and is almost completely buried within the hydrophobic core of the native foldon. Thus only very conservative mutations are considered at this position. The canonical amino acids serine and valine are chosen to study the effect caused by an introduction of a single hydroxyl- or two methyl-groups. However, as the introduction of a hydroxyl-group within a hydrophobic cluster should cause an adverse effect both by steric hindrance and unfavorable interactions of the hydroxyl-group with its non-polar environment, the non-natural amino acid 2-(L)-aminobutyric acid is incorporated as well. Non-physiological amino acids broaden the spectrum of side chain mutants and isoforms, which is an advantage if even the most conservative mutations lead to severe effects on both stability

and folding kinetics. As both serine and aminobutyric acid should have a similar effect on the interactions with neighboring side chains, the impact of the introduced hydroxyl group can be probed by comparison of the effect on both stability and folding kinetics of the two mutants.

Mutations of proline 7 are intended to disrupt the rigid backbone conformation introduced by the pyrrolidine ring without effecting the hydrophobic interaction with the rest of the hydrophobic cluster. For this reason, phenylalanine and valine are chosen as substitutions. The impact on the stability of these natural amino acid exchanges is large. Thus, the non-natural amino acid norvaline is chosen, as it has the same side chain conformation as proline with the difference that the C_δ is not bound to the backbone nitrogen. This variant can have almost native-like side chain interactions with the hydrophobic cluster, leading to a higher stability of the native state.

To study the assembly of foldon without proline *cis-trans* isomerization, a double mutant is prepared with the proline 4 mutated to a threonine, while the proline at position 7 is replaced by a norvaline. This mutation can additionally give a hint whether or not the poly-proline II structure is only stable in the presence of the two proline residues, and if the effects observed for the single-mutant foldon variants at these positions are influencing each other or act independently.

The most conservative of all amino acid exchanges is performed for the tyrosine at position 13, which is replaced by a phenylalanine. The exchange should not have any consequences on the interactions of the hydrophobic cluster, as all native non-polar interactions are still prevailed in the mutant. Disruption of the hydrogen-bonding network of the tyrosine's hydroxyl group with the first type II-turn is the proposed consequence of this mutation, which can be directly probed and thus can give information on the strength of a buried hydrogen bond in proteins.

Tryptophan 20 is replaced by the non-natural amino acid 1-(L)-naphthylalanine (Nal). Naphthylalanine has an almost identical side chain structure as tryptophan and should be an extremely conservative mutation. All observed effects of this mutation should thus occur due to the disruption of the hydrogen bond between the indole N_1 and the carbonyl group of the proline residue at position 4, making it another excellent probe for a buried hydrogen bond. To observe if any effects caused by the disruption of this hydrogen bond can be reversed, histidine is used as a second replacement at position 20.

The mutation of the leucine at position 22 incorporates phenylalanine. The leucine residue packs onto the phenyl group of the tyrosine at position 13 and interacts by van der Waals interactions. This interactions should be prevailed or even enhanced by an exchange to phenylalanine, as the contact surface of the two phenyl groups and their dislocated π -electrons is enlarged.

All foldon variants are synthesized using Fmoc-chemistry, except foldon A6S, which is obtained by expression as a fusion protein in *E.coli*. All used probes showed > 99 % purity on analytic HPLC runs before performing experiments. The experiments carried out for the individual foldon variants are discussed for each foldon variant in detail.

4.2. Foldon P4T

Amino acid exchanges can have a large impact on the native state structure, especially when each mutation is present three times, as it is the case in foldon. To detect if all foldon mutants are still able to fold into the native structure, crystallization of each foldon variant was tried to determine the crystal structure. The crystal structure of foldon P4T does not show any significant deviations from the foldon wild-type crystal structure if comparing both the backbone root mean square deviations (r.m.s.d.s) and the configuration of the changed amino acid at position 4. The configuration of the amino acid backbone of foldon P4T is slightly deviated to foldon wild-type at position four, as seen from the ϕ - and ψ -angles of both wild-type and the threonine residue in the mutant at this position. The backbone angles are $\phi = -75.28^\circ$ and $\psi = 163.42^\circ$ for foldon wild-type, while in the P4T variant the angles are $\phi = -85.68^\circ$ and $\psi = 167.87^\circ$. This nevertheless results in an almost identical conformation, and especially the *trans* conformation of the peptide bond is prevailed. The prominent hydrogen bond between the backbone carbonyl group of residue four and the N_1 in the indole of tryptophan 20 is not disrupted. This is expected as all other contributions stabilizing the conformation of the N-terminus of foldon are still present in the mutant. The poly-proline II conformation of the N-terminus hence seems not to dependent on the presence of the proline residues in the native state structure, as it is stabilized by both intra- and intermolecular interactions of all N-terminal amino acids. The loss of only one structure-defining amino acid is not sufficient to change the overall N-terminal backbone conformation.

The fluorescence spectrum of the foldon P4T native state N is almost identical to foldon wild-type, both in intensity and in shape. The unfolded state U of foldon P4T has a decreased fluorescence maximum of approximately 20 %, but still shows the characteristic shape of the foldon wild-type unfolded state. The burst-phase intermediate I of foldon P4T has a strongly decreased maximum at 330 nm, although the spectra of the intermediates can be superimposed for the wavelengths below 310 and above 370 nm. This could indicate a less compact intermediate structure, caused by the higher flexibility of the last N-terminal residues due to the lost backbone constraint resulting from the exchange of the proline 4 residue.

The GdmCl transition of foldon P4T results in a folding stability of $\Delta G^0(\text{H}_2\text{O}) = (-101.5 \pm 3.0) \text{ kJ} \cdot \text{mol}^{-1}$ and an m_{eq} -value of $m_{eq} = (11.4 \pm 0.7) \text{ kJ} \cdot \text{mol}^{-1} \cdot \text{M}^{-1}$. Foldon P4T is the only foldon variant that is even more stable than foldon wild-type, with a difference of $\Delta\Delta G^0(\text{H}_2\text{O}) = -12.3 \text{ kJ} \cdot \text{mol}^{-1}$. A comparison of all free folding enthalpies is given in Tables 4.3 and 4.4 in Chapter 4.14. Due to the decreased free folding enthalpy, a determination of the free folding enthalpy by an urea transition is not possible, as the denaturant strength of urea is not sufficient to unfold foldon P4T even at low foldon monomer concentrations (data not shown). This high stability against denaturation is also the reason denaturant-dependent unfolding kinetics to determine k_{-N} are not possible either. Refolding experiments of foldon P4T however can be carried out in urea, as unfolding of all foldon variants is achieved by both high denaturant concentrations and lowering the pH-value to pH 2.08. Complete unfolding of foldon P4T in urea at the highest foldon monomer concentration is ensured by an unfolding experiment of the stock solution, which does not show any fluorescence change upon 20fold dilution with 6.4 M GdmCl pH 2.0 buffer (data not shown).

Foldon P4T shows the highest resemblance to foldon wild-type with regard to the refolding kinetics. The half-life times for concentration-dependent refolding steps in both denaturants results in a reaction order of three for low foldon monomer concentrations, while the half-life times deviate from the linear behavior for higher foldon monomer concentrations when the concentration-independent reaction step from the intermediate trimeric state to the native state becomes rate-limiting (Figure A.26A and A.28A in Appendix A.7). A summary of all reaction orders determined from the half-life times is given in Tables 4.13 and 4.14 in Chapter 4.17 for GdmCl and urea refolding experiments, respectively. The initial slope of all studied foldon variants is used to determine both the reaction order at the start of the reaction and the

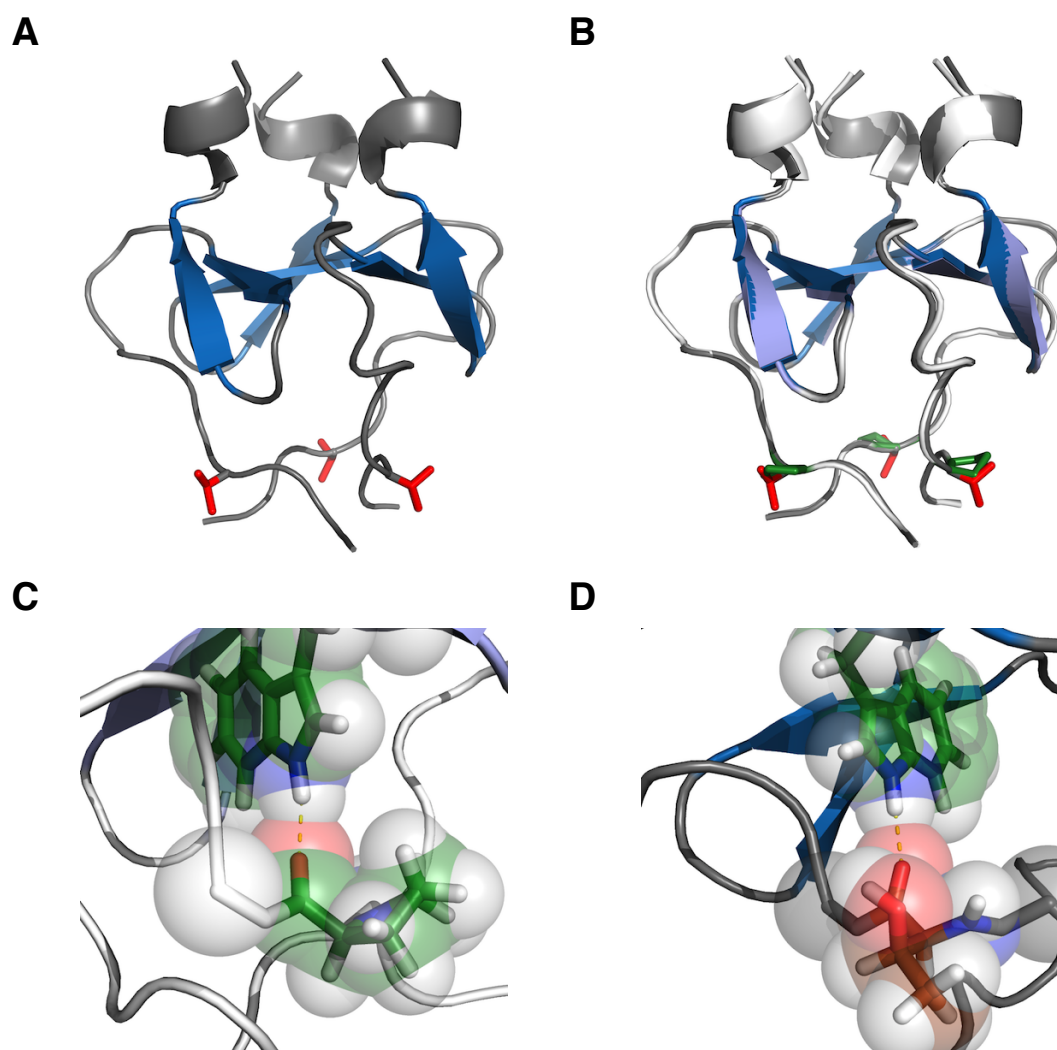


Figure 4.2.: Crystal structure of foldon P4T. A) Crystal structure of foldon P4T in a ribbon representation, with the β -strands highlighted in blue, while the 3_{10} -helix and the other regions of the structure are colored in grey. The side chain of the mutated residue 4 is shown in stick representation and the side chain is colored in red. B) Backbone alignment of foldon P4T and foldon wild-type. Foldon wild-type, also shown in ribbon representation, is colored in white, with the exception of the β -strands, which are colored in light blue. The proline 4 residue is shown in stick representation and colored in green. The foldon P4T crystal structure is shown and colored as in A). C) Foldon wild-type stick representation of the residues 4 and 20 with atoms colored in green for carbon, red for oxygen, blue for nitrogen and white for hydrogen atoms. The van der Waals radii of the different atoms are shown as semi-transparent spheres. The hydrogen bond between the residue 4 backbone oxygen and the residue 20 side chain nitrogen is highlighted in yellow. D) Foldon P4T stick representation of the residues 4 and 20, using the same color scheme and representation as in C), with the exception that the carbon atoms of the residue 4 are colored in brown. The illustration shows that the hydrogen bond between the residue 4 backbone oxygen and the residue 20 side chain nitrogen is formed in the variant as well.

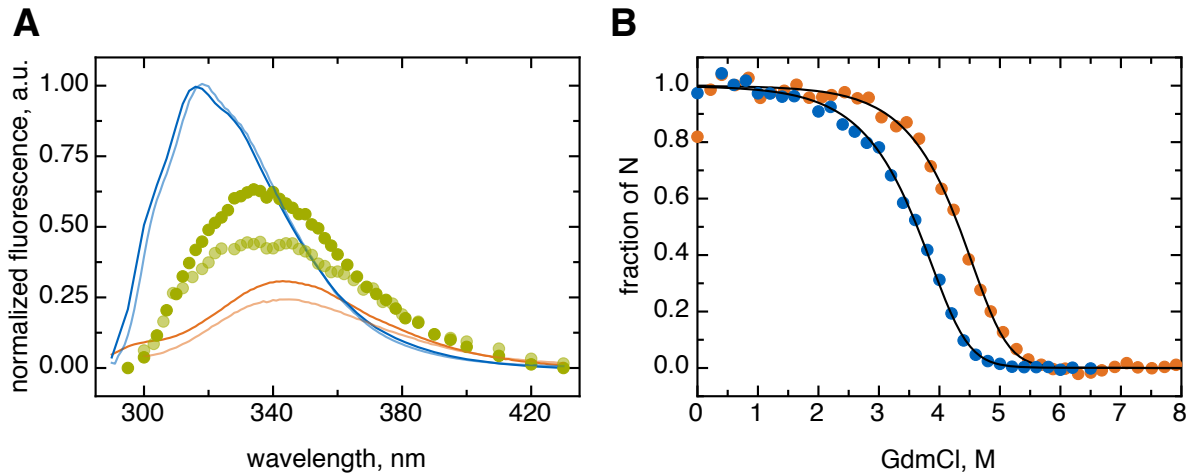


Figure 4.3.: A) Fluorescence spectra of foldon wild-type and foldon P4T. The native state spectra are colored in blue, the ones of the unfolded state in orange and the burst-phase intermediate spectra in green. The fluorescence spectra of foldon P4T are lighter colored than the corresponding foldon wild-type spectra. B) Comparison of the foldon wild-type (blue) and foldon P4T (orange) GdmCl transitions at $30 \mu\text{M}$ foldon monomer concentration.

dimerization rate constant k_D . The reaction order is close to 2 in both GdmCl and urea for all other foldon variants. This is determined from the slope of the common logarithm of the initial slope plotted as a function of the common logarithm of the foldon monomer concentration. The ordinate intercept yields the dimerization rate constant k_D of foldon P4T, with $k_D = (7.24 \pm 0.05) \cdot 10^5 \text{ M}^{-1} \cdot \text{s}^{-1}$ and $k_D = (4.80 \pm 0.06) \cdot 10^5 \text{ M}^{-1} \cdot \text{s}^{-1}$ in GdmCl and urea, respectively.

Foldon P4T has a proline isomerization step with a similar rate constant and amplitude compared to foldon wild-type. The impact of this mutation is only marginal on the refolding speed. The effect on the stability, about $4.1 \text{ kJ} \cdot \text{mol}^{-1}$ per monomer, is likely caused by the gained entropy of the N-terminal region, due to the loss of the rigidifying Ile3-Pro4 peptide bond. This is supported by NMR-data of foldon E5R, which shows that 10 to 14 % of this peptide bond are in the *cis* conformation in the folded monomeric state¹⁵¹. In contrast, the Ala6-Pro7 peptide bond is completely in the *trans* conformation in the foldon E5R NMR structure. Foldon P4T has the same amplitude of the *cis-trans* isomerization reaction as foldon wild-type of approximately 20 % although the *cis* content of the Ala6-Pro7 peptide bond is only 7 %⁸¹. This leads to the conclusion that the Ile3-Pro4 peptide bond does not contribute to the overall *cis-trans* isomerization reaction and that foldon monomers with a *cis* conformation of the

4. Results

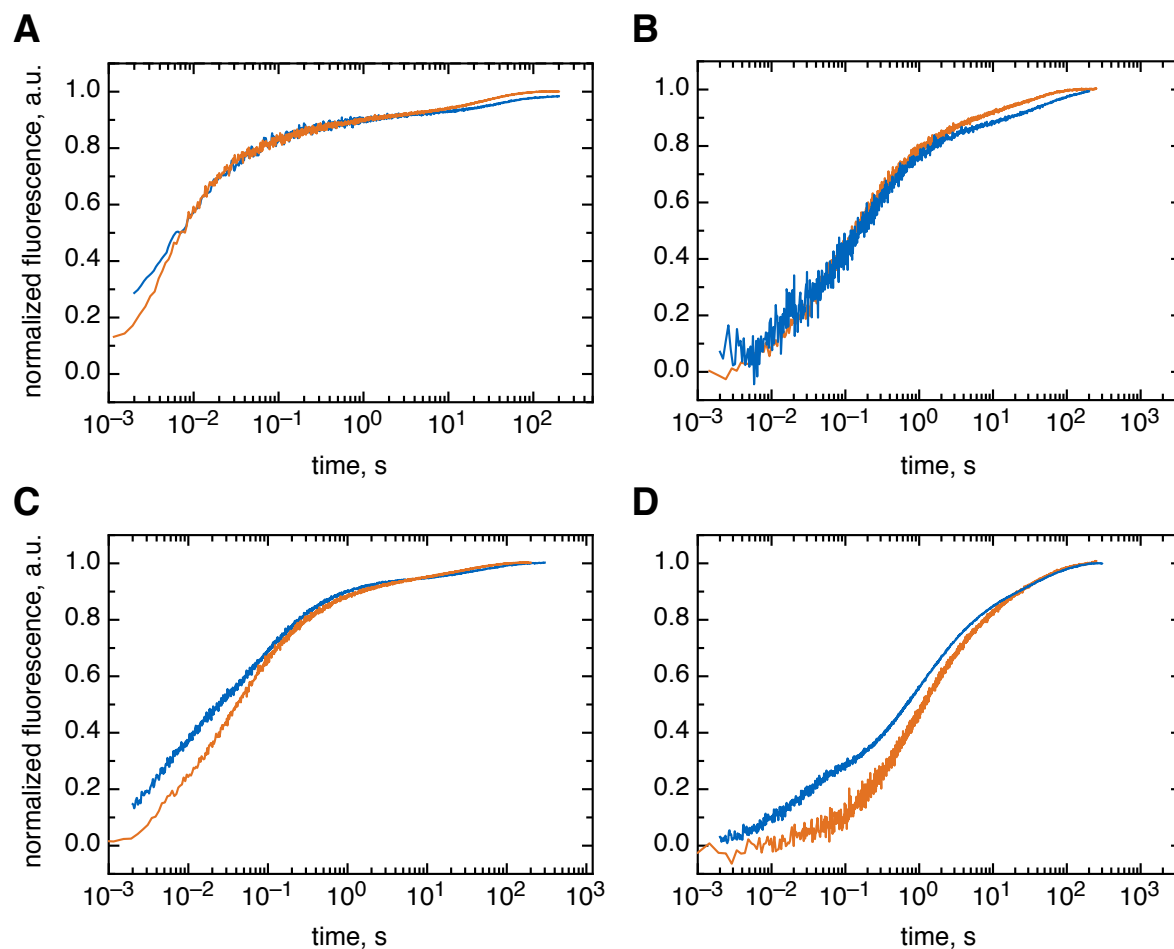


Figure 4.4.: Comparison of foldon wild-type (blue) and foldon P4T (orange) refolding traces. A) 100 μM foldon monomer refolding traces in 0.58 M GdmCl. B) 10 μM foldon monomer refolding traces in 0.58 M GdmCl. C) 100 μM foldon monomer refolding traces in 1.0 M urea. D) 10 μM foldon monomer refolding traces in 1.0 M urea.

Ala6-Pro7 peptide bond can form dimers and trimers. The model used for rate constant determination is adjusted to represent this productive dimer- and trimer-formation of *cis* monomers. The rate constants determined for the foldon P4T association show a high similarity to foldon wild-type (see Table 4.15). The dissociation rate constant k_{-T} can be determined with high accuracy from the fit. Fitting with a fixed dimerization rate constant k_D results in significant deviations of the fit curves from the measured refolding traces, especially for high monomer concentrations in the first 10 to 50 ms of the reaction. The fit is therefore run freely without any restrictions, which results in an increased dimerization rate of $k_D = 1.5 \cdot 10^6 \text{ M}^{-1} \cdot \text{s}^{-1}$. This value is twice as large as the one determined from the initial slope fit, but still within the

error range of this value. The refolding traces of foldon P4T in urea show a larger resemblance to the foldon wild-type refolding traces in GdmCl than to the ones in urea. Comparison of the rate constants of foldon P4T and foldon wild-type in urea show that the dimer association rate constant is decreased by a factor of two for foldon P4T. The stabilization of the dimeric intermediate found for foldon wild-type is lost. Separation of the dimer and trimer formation phase is further perturbed by the fact that the trimerization rate constant is slightly larger than the dimerization rate constant, which is contrary to the results determined for foldon wild-type. In foldon wild-type in urea, dimerization happens with a faster rate than trimerization. The dissociation rate constant of the native state is two orders of magnitude lower than in foldon wild-type, caused by the higher stability of foldon P4T compared to foldon wild-type.

4. Results

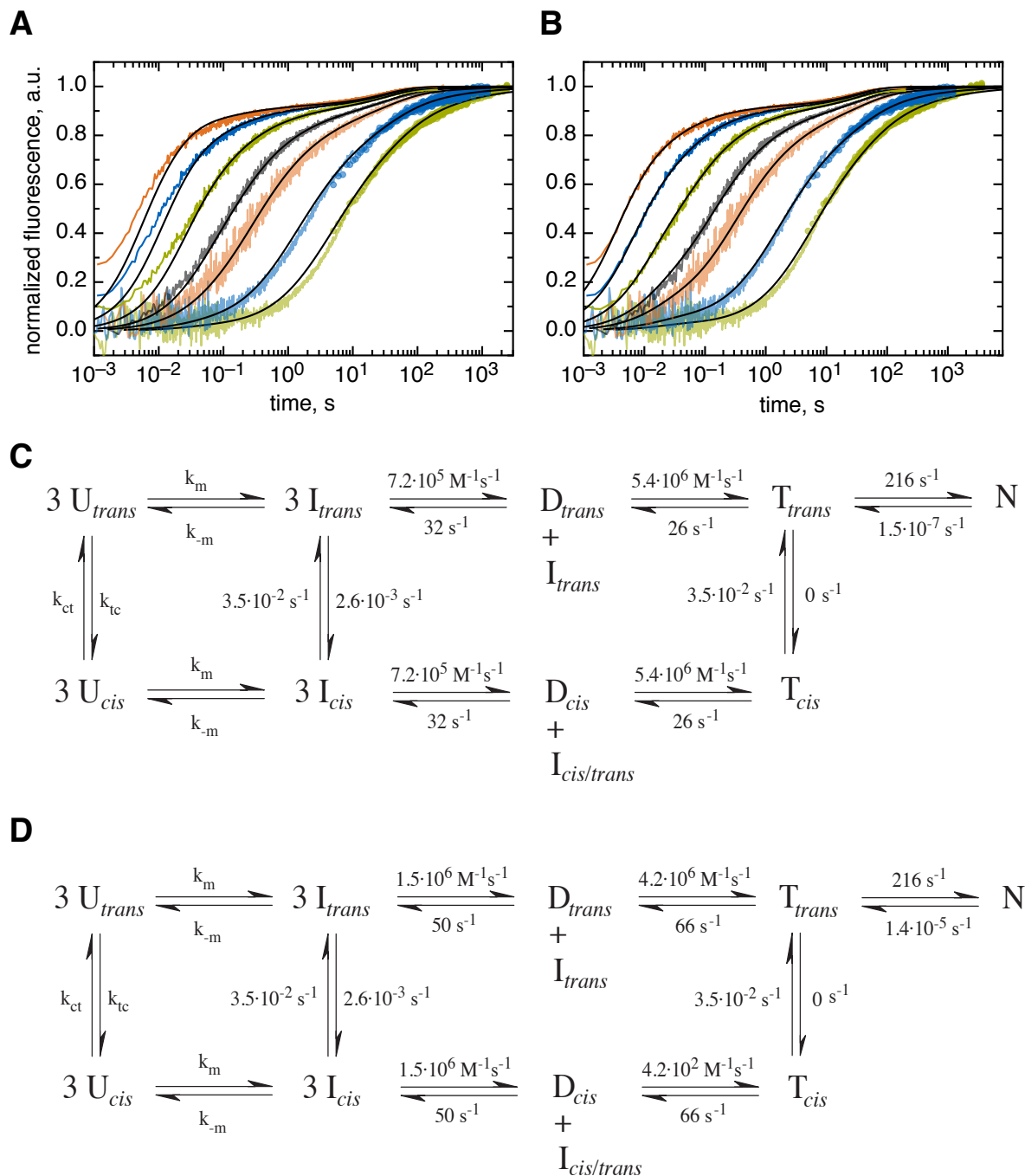


Figure 4.5.: Normalized global fit traces of foldon P4T GdmCl refolding kinetics. A) Normalized GdmCl refolding traces after fitting with fixed dimerisation rate constant k_D . B) Normalized GdmCl refolding traces after fitting with all parameters run freely. C) Fitting model with all determined rate constants with fixed dimerisation rate constant k_D . D) Fitting model with all determined rate constants with all parameters run freely.

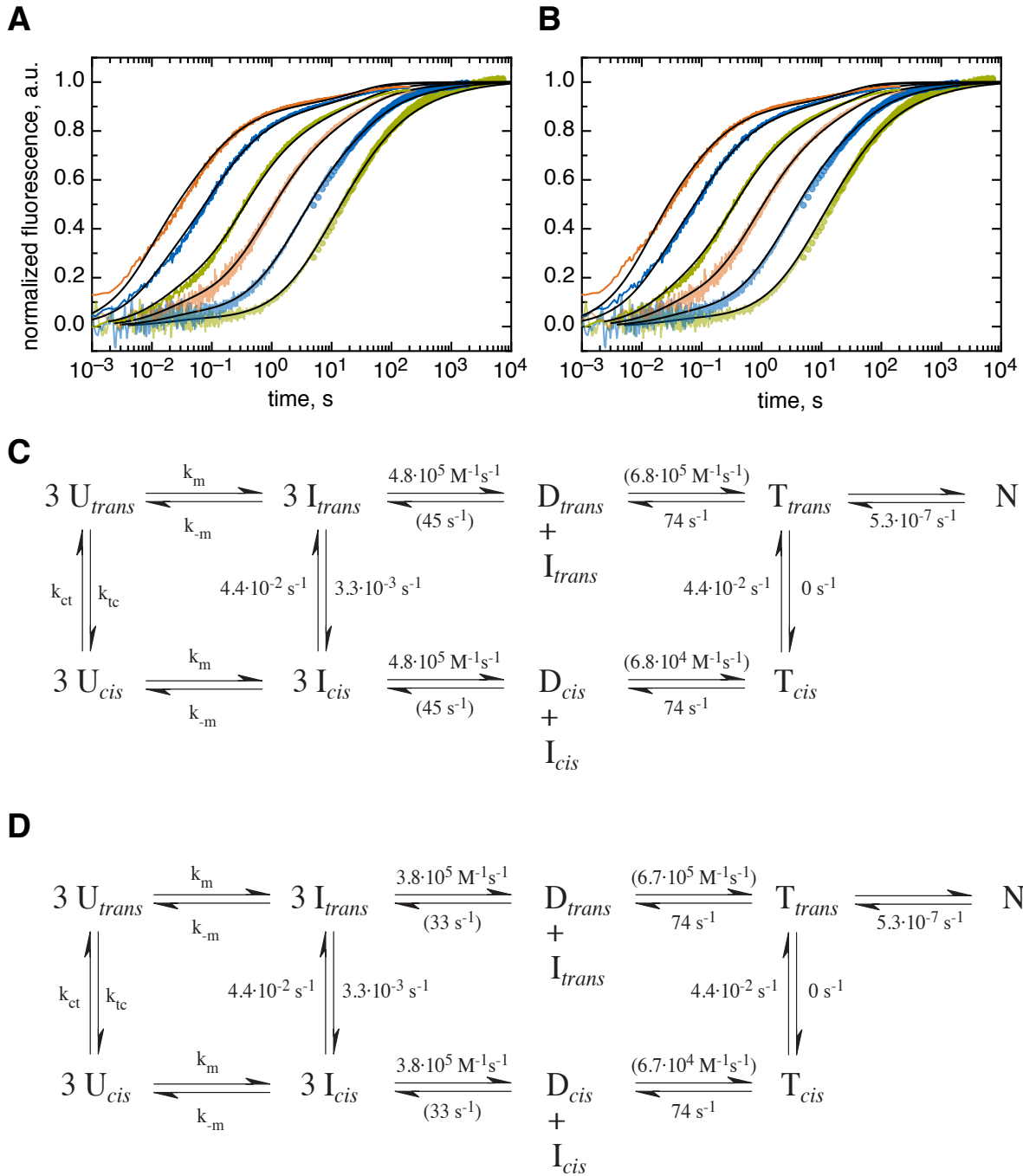


Figure 4.6.: Normalized global fit traces of foldon P4T urea refolding kinetics. A) Normalized urea refolding traces after fitting with fixed dimerisation rate constant k_D . B) Normalized urea refolding traces after fitting with all parameters run freely. C) Fitting model with all determined rate constants with fixed dimerisation rate constant k_D . D) Fitting model with all determined rate constants with all parameters run freely.

4.3. Foldon A6Abu

In the crystal structure of foldon A6Abu, the additional methyl-group of the 2-(L)-aminobutyric acid points towards the arginine 15 of the neighboring chain. This causes a steric clash with the C_γ of this arginine side chain. An increased contact area to none of the other amino acids of the hydrophobic cluster can be seen. If any observed effects of this variant are due to impacts on the native state, they are likely a consequence of the steric hindrance with residue 15. However, the crystal structure of foldon A6Abu highly resembles the foldon wild-type structure, thus any structural effects are unlikely.

The shape of the native state fluorescence spectrum of the foldon A6Abu variant has an almost identical shape to the one of foldon wild-type, except that its maximum is lowered by about 10 %. It is unclear whether this decrease is caused by a change in the fluorescence properties of the tryptophan due to the changed local environment of the fluorophore or just due to a lower stability and thus fewer formed native trimer of foldon A6Abu. The maximum is also decreased by about 10 % for the unfolded state of foldon A6Abu, but otherwise no deviations are detected in the shape of the spectrum. The intermediate spectrum of foldon A6Abu matches the spectrum of foldon wild-type *I*, despite a slight increase of the fluorescence maximum. Foldon A6Abu thus seems highly comparable in all states to foldon wild-type. The free folding enthalpy of foldon A6Abu is determined by both GdmCl and urea transitions. The stabilities are $\Delta G^0(H_2O) = (-73.7 \pm 1.4) \text{ kJ} \cdot \text{mol}^{-1}$ with $m_{eq} = (11.0 \pm 0.5) \text{ kJ} \cdot \text{mol}^{-1} \cdot \text{M}^{-1}$ in GdmCl and $\Delta G^0(H_2O) = (-69.3 \pm 0.7) \text{ kJ} \cdot \text{mol}^{-1}$ with $m_{eq} = (5.3 \pm 0.1) \text{ kJ} \cdot \text{mol}^{-1} \cdot \text{M}^{-1}$ in urea. The difference in the determined stability can be caused by the ill-defined native baseline for the native state in the GdmCl transitions. A global fit resulting an overall free folding enthalpy is thus performed, resulting $\Delta G^0(H_2O) = (-70.7 \pm 0.8) \text{ kJ} \cdot \text{mol}^{-1}$ with $m_{eq}(\text{GdmCl}) = (9.7 \pm 0.3) \text{ kJ} \cdot \text{mol}^{-1} \cdot \text{M}^{-1}$ and $m_{eq}(\text{urea}) = (5.5 \pm 0.2) \text{ kJ} \cdot \text{mol}^{-1} \cdot \text{M}^{-1}$. This procedure had also to be carried out for several other foldon variants. The results are summarized in Table 4.6 in Chapter 4.14, where the observed differences in the free folding enthalpies detected in GdmCl and urea are discussed in detail.

Unfolding of foldon A6Abu in 0.58 M GdmCl 20 mM sodium cacodylic acid pH 7.0 is achieved by 6fold dilution with GdmCl stock solutions of varying concentrations (3 M to 8 M) with pH 7.0. Unfolding is detected by stopped-flow measurements at all used GdmCl concentrations, and traces are fitted to a double-exponential equation without significant systematic

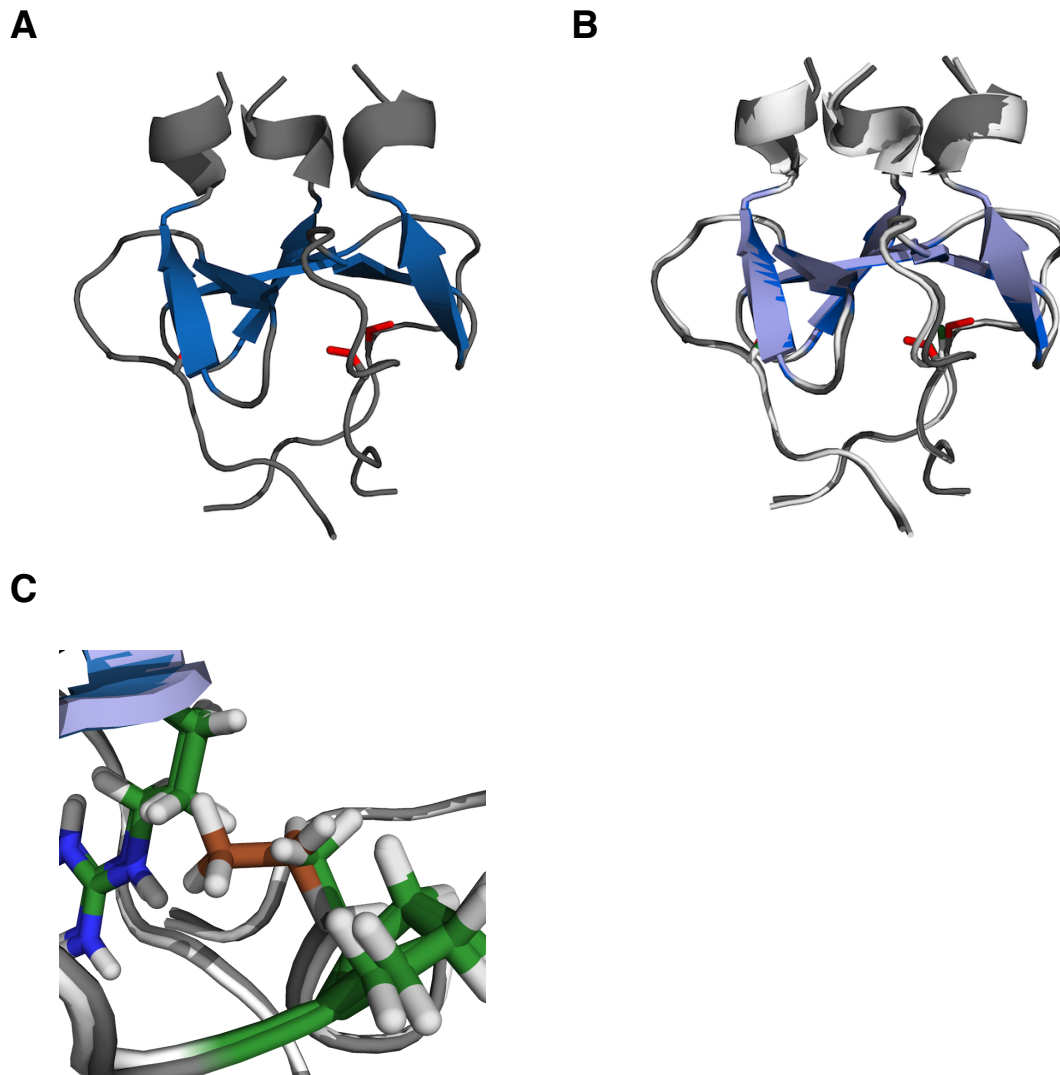


Figure 4.7.: Crystal structure of foldon A6Abu. A) Crystal structure of foldon A6Abu in a ribbon representation, with the β -strands highlighted in blue, while the 3_{10} -helix and the other regions of the structure are colored in grey. The side chain of the mutated residue 6 is shown in stick representation and the side chain colored in red. B) Backbone alignment of foldon A6Abu and foldon wild-type. Foldon wild-type, also shown in ribbon representation, is colored in white, with the exception of the β -strands, which are colored in light blue. The alanine residue at position 6 is shown in stick representation and colored in green. The foldon A6Abu crystal structure is shown and colored as in A). C) Overlay of the residue 6 side chains of both foldon wild-type and foldon A6Abu, together with the neighboring residues Pro7 and Arg15. All three side chains are shown in stick representation, while the rest of the proteins is colored as in A). The carbon atoms are shown in green (with the exception of Abu6, where the carbon atoms are shown in brown), the nitrogen atoms in blue and the hydrogen atoms in white.

4. Results

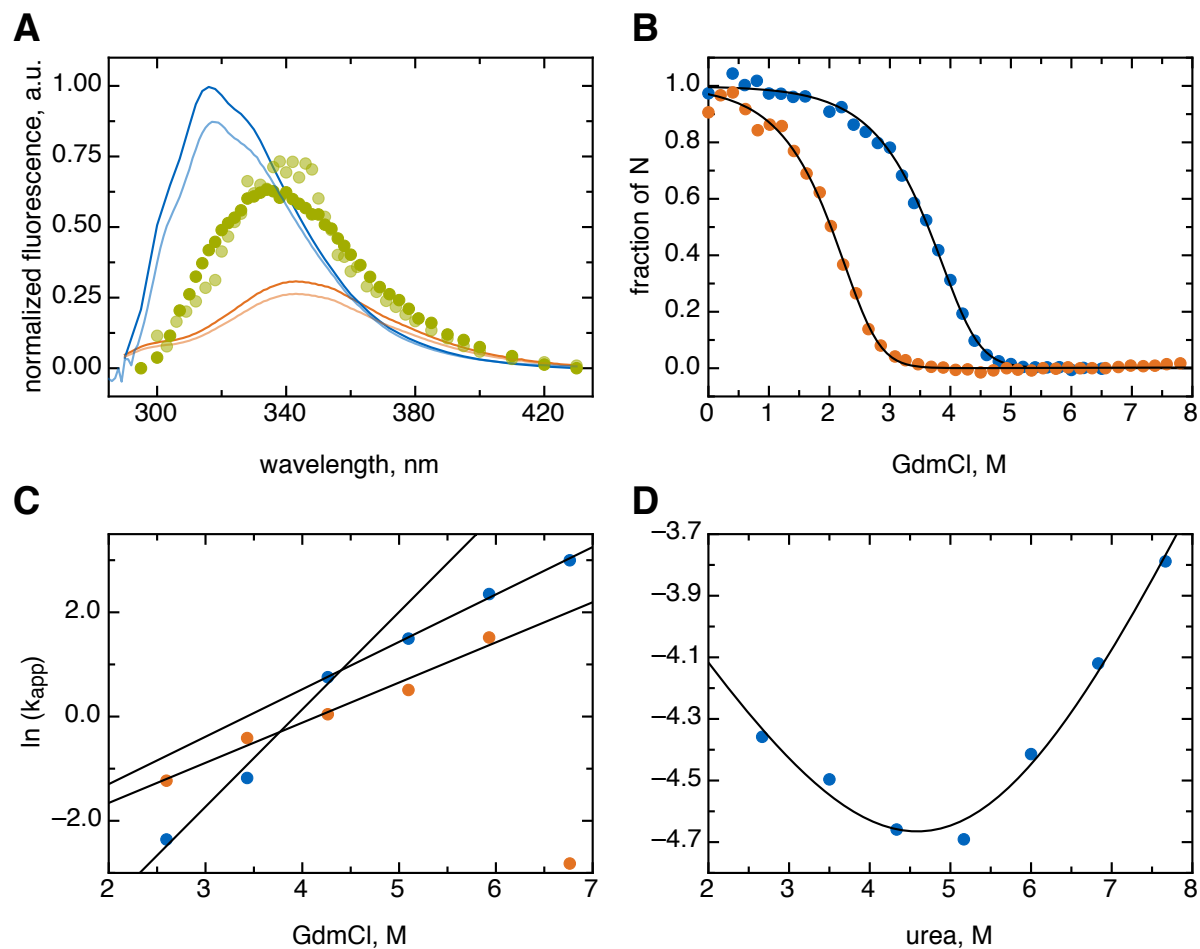


Figure 4.8.: A) Fluorescence spectra of foldon wild-type and foldon A6Abu. The native state spectra are colored in blue, the spectra of the unfolded state in orange and the burst-phase intermediate spectra in green. The fluorescence spectra of foldon A6Abu are lighter colored than the corresponding foldon wild-type spectra. B) Comparison of the foldon wild-type (blue) and foldon A6Abu (orange) GdmCl transitions at 30 μ M foldon monomer concentration. C) Linear plot of the two phases of the foldon A6Abu GdmCl unfolding. The common logarithm of the apparent rate constants are plotted against the corresponding GdmCl concentrations. D) Chevron plots for foldon A6Abu urea unfolding. The common logarithm of the apparent unfolding rate constants is plotted against the varying urea concentrations, and the data points are fitted using equation 3.17, derived in Chapter 3.9.

deviation of the residuals, as shown by the fit and residuals in Figure A.11A. The common logarithm of the two apparent rate constants are plotted against the final GdmCl concentrations after dilution and should yield a linear plot of the unfolding limb seen in chevron plots (Figure 4.8C). The apparent rate constants are grouped according to their contribution to the total amplitude of the reaction, which is necessary, as the rates of the largest amplitude change show a kink in their plot and are for low GdmCl concentrations slower than the rate constants of the lower amplitude, although this assignment is not robust for the unfolding traces with 4 and 5 M GdmCl. For higher GdmCl concentrations, however, the major phase is also the fastest phase, showing a kink or roll-over in the unfolding trace for this reaction. Only one phase is detectable for the highest GdmCl concentration at 8 M. This behavior is also seen for all other residues in the N-terminal region of foldon, as shown below. The slowest unfolding rate constant determined is always considered to correspond to k_{-T} , which for foldon A6Abu results in $k_{-T} = (6.52 \pm 0.81) \cdot 10^{-4} \text{ s}^{-1}$ with an $m_U = (4.55 \pm 0.64) \text{ kJ} \cdot \text{mol}^{-1} \cdot \text{M}^{-1}$. A summary of all GdmCl unfolding experiments is shown in Table 4.7 in Chapter 4.15.

Urea unfolding traces of foldon A6Abu are fitted with a one-exponential equation, though some systematic deviations are observed for $t > 100 \text{ s}$. However, these deviations can be neglected due to the small amplitude. The deviations are likely caused by photo-bleaching of the tryptophan. The typical V-shape of a chevron plot is observed when the common logarithms of the apparent rate constants are plotted against the final urea concentrations (Figure 4.8D). Fitting yields $k_f = (3.4 \pm 1.1) \cdot 10^{-2} \text{ s}^{-1}$ and $k_u = (3.2 \pm 2.8) \cdot 10^{-4} \text{ s}^{-1}$; the kinetic m -values are $m_f = (975 \pm 328) \text{ kJ} \cdot \text{mol}^{-1} \cdot \text{M}^{-1}$ and $m_u = (1.3 \pm 0.3) \cdot 10^3 \text{ kJ} \cdot \text{mol}^{-1} \cdot \text{M}^{-1}$, respectively. While it is clear that the unfolding rate constant k_u is equivalent with k_{-N} , k_f cannot be clearly assigned. The same holds true for the m_f -value, especially as the sum of both kinetic m -values does not yield the m_{eq} -value, but falls short by $3.0 \text{ kJ} \cdot \text{mol}^{-1} \cdot \text{M}^{-1}$. A possible explanation is that the folding reaction probed in the chevron plot does not represent the complete assembly reaction, but only the step from an assembly intermediate to the native trimer. Whether this intermediate is the trimeric intermediate T or the dimer D cannot be determined. k_f , though apparent first-order, could thus be a combination of the first order rearrangement step and previous concentration-dependent second-order association steps. A summary of all urea unfolding experiments is shown in Table 4.8 in Chapter 4.15.

The reaction order for GdmCl refolding of foldon A6Abu is determined as 2.96 ± 0.26 from the concentration-dependent refolding half-life times. While the deviation from the linear

4. Results

behavior is still observable for higher foldon monomer concentrations, the effect is not as pronounced as for foldon wild-type (Figure A.26B). The initial slope results in a reaction order of 2.04 ± 0.05 , yielding $k_D = (3.32 \pm 0.04) \cdot 10^5 \text{ M}^{-1} \cdot \text{s}^{-1}$. In urea, a reaction order of 2.67 ± 0.03 is determined from the half-life times. This broken reaction order could be a consequence of the strongly decelerated refolding reaction. The initial slope of the urea refolding traces yields a reaction order of 2.09 ± 0.13 , resulting in $k_D = (5.75 \pm 0.13) \cdot 10^4 \text{ M}^{-1} \cdot \text{s}^{-1}$.

Fitting of the GdmCl refolding data was accompanied by severe difficulties for all foldon variants except foldon P4T and foldon P4T P7Nva. The refolding curves need to be adjusted to the fraction of natively folded protein at both the monomer concentration and the stability at the refolding conditions. If using the fraction of N for each foldon monomer concentration determined in this manner, fitting of the adjusted refolding curves is not possible. The fitting curves are not able to globally represent the reaction course, or lead to non-physiological association or dissociation rate constants. When the stability of N is however enlarged, resulting in higher fractions of folded protein, fitting of the data yields well-defined results with little or no deviation from the determined data. The fitting of the refolding traces after adjustment to the respective fraction of N is however possible in urea for all studied foldon variants. A wrong model is thus unlikely, as a change of the denaturant should not have an effect on the reaction mechanism. There is no consistent explanation for the needed corrections of the native state stability in GdmCl, and unfortunately the adjustment necessary to gain a suitable fit varies for different foldon variants. The complete process to determine the least error-prone free folding enthalpy is an iterative process and accompanied by a large uncertainty. The obtained results have thus to be interpreted with extreme caution.

Despite the adjustments, and especially in urea, deviations of the fits from the refolding curves are observed. There are several possible explanations for this behavior: First, an exact determination of the foldon monomer concentration is not possible, especially at low foldon monomer concentration and in buffers with high denaturant concentrations. A deviation of the concentration from the determined one by light absorption has large impacts on the fit, as all association steps are highly sensible to the monomer concentration. Additionally, when using an erroneous concentration, the calculation of the formed fraction of N cannot yield correct results. This step is additionally error-prone as the determined stability of the foldon variant is itself defective, leading to a propagation into the rate constant determination. This is the most likely source for all deviations: The low stability results in a strong shift of the equi-

librium towards the unfolded state, and the strong concentration dependence of the reaction yields fractions of N at all refolding conditions which are significantly lower than 1. When considering all these effects, the observed deviations of around 5 % are rather small. Foldon wild-type and P4T however are so stable that the fraction of native trimers is close to 1 under all chosen refolding conditions, thereby eliminating this error source. This explains the better fits observed with these two variants.

Every mutation at the position six results in an increase of the *cis* content of the Xaa6-Pro7 peptide bond to approximately 10 %. This *cis* content is found both for serine and valine⁸¹, and is also expected for 2-(L)-aminobutyric acid. This should lead to an increase of the concentration-independent phase in all mutations at position 6 caused by the *cis-trans* isomerization. Comparison of the normalized refolding curves shows that not an increase, but a decrease of the amplitude of this phase is observed. The slow phase only contributes about 10 % to the overall amplitude change of the reaction, in contrast to the 30 % that would be expected using the modified folding model. Global fitting of the refolding data for each foldon variant with a mutation at position 6 is however not possible when using the modified folding model, even for single traces. Instead, the model used before to fit the foldon wild-type refolding traces in GdmCl is used, where only foldon monomers in the all-*trans* conformation are able to form dimers that can actively contribute to the association of the native trimer. This together with the lower amplitude of *cis-trans* isomerization points to the fact that for foldon A6Abu as well as all other foldon variants at position 6 (shown below), dimerization of monomers with a *cis* conformation of the Xaa6-Pro7 peptide bond is not possible any more. Results of the fit together with all rate constants for foldon A6Abu in GdmCl are shown in Figure 4.10. The stability used for the adjustment of the foldon A6Abu refolding traces is $\Delta G^0 = 73.7 \text{ kJ} \cdot \text{mol}^{-1}$.

The dissociation rate constants of the dimeric and trimeric intermediate state k_{-D} and k_{-T} show a large deviation from foldon wild-type, while the association rate constants k_D and k_T are only decreased by a factor of two for the fit with all rate constants run freely. The increase of the dissociation rate constant of the native trimer is expected from the lower overall stability of the foldon A6Abu variant, although a robust fit can only be achieved if the dissociation rate constant is increased by almost one order of magnitude compared to the value determined from the unfolding fits. The fit traces show a good agreement with the detected data points, and the close resemblance with the curves of foldon wild-type indicates the minor changes of

4. Results

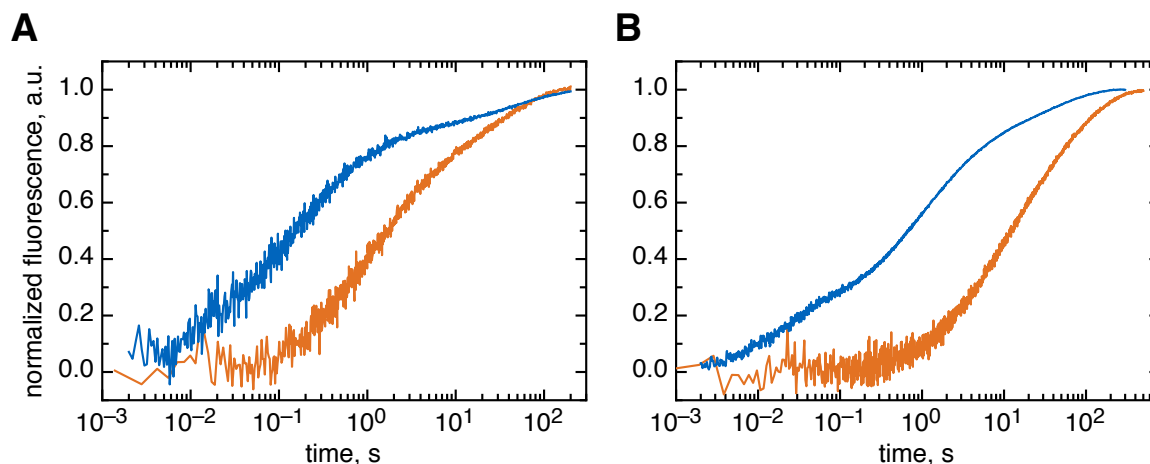


Figure 4.9.: Comparison of foldon wild-type (blue) and foldon A6Abu (orange) refolding traces. A) 10 μ M foldon monomer refolding traces in 0.58 M GdmCl. B) 10 μ M foldon monomer refolding traces in 1.0 M urea.

all rate constants compared to foldon wild-type. The most tremendous effect of the A6Abu mutation is the fact that 10 % of all monomers are not able to productively form dimers and trimer any more but have to undergo *cis-trans* isomerization of the Xaa6-Pro7 peptide bond before any further association reaction.

An overview over the normalized urea refolding traces after fitting and the resulting rate constants for each individual step are given in Figure 4.11. The dimerization rate constant is decreased by a factor of 20 compared to foldon wild-type, while the trimerization rate constant is enlarged by a factor of 2. The dissociation rate constants of the native and intermediate trimeric state are enlarged as well, as expected from the increased free folding enthalpy. The difference between the fixed and the free run are only marginal.

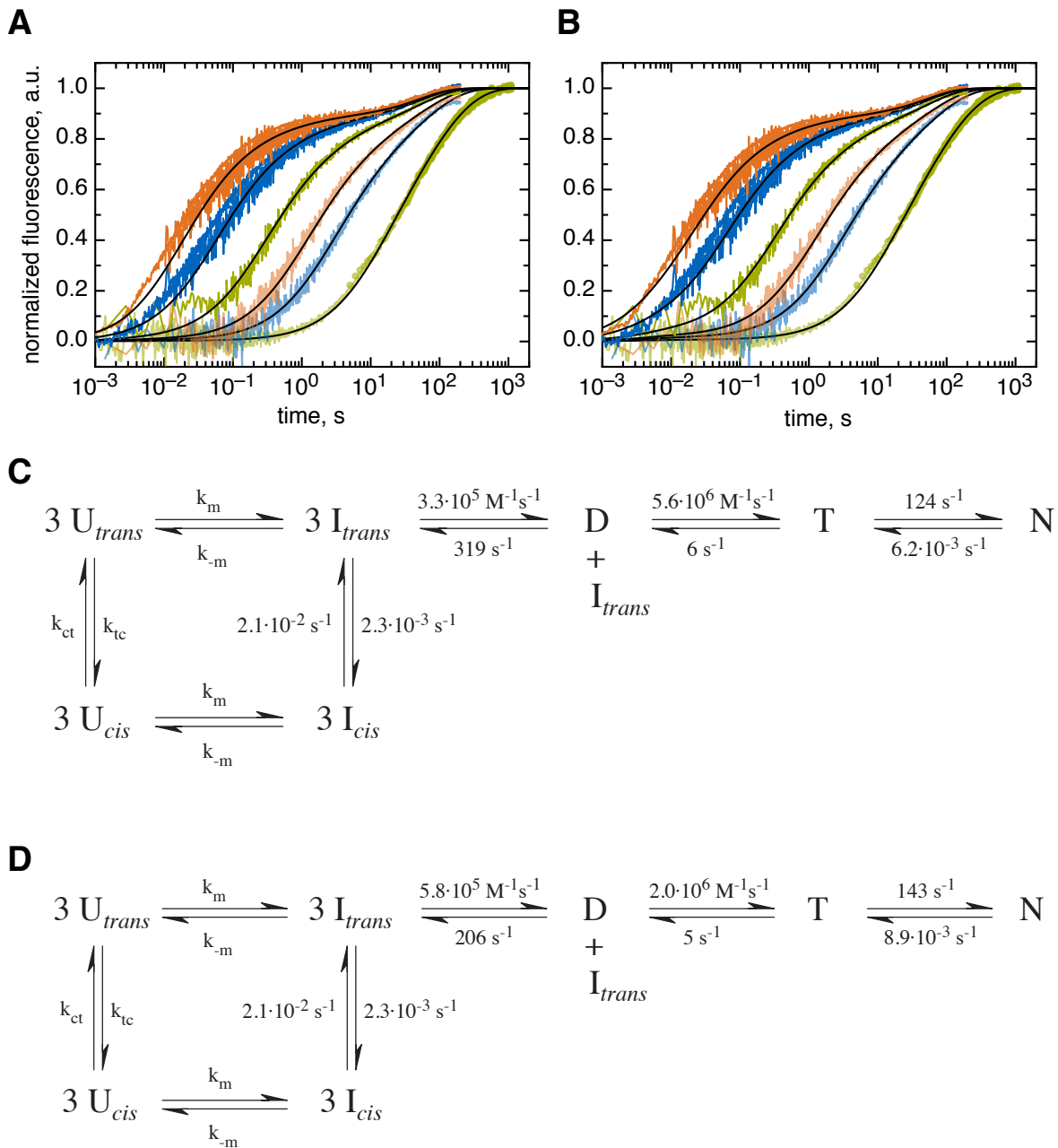


Figure 4.10.: Normalized global fit traces of foldon A6Abu GdmCl refolding kinetics. A) Normalized GdmCl refolding traces after fitting with fixed dimerization rate constant k_D . B) Normalized GdmCl refolding traces after fitting with all parameters run freely. C) Fitting model with all determined rate constants with fixed dimerization rate constant k_D . D) Fitting model with all determined rate constants with all parameters run freely.

4. Results

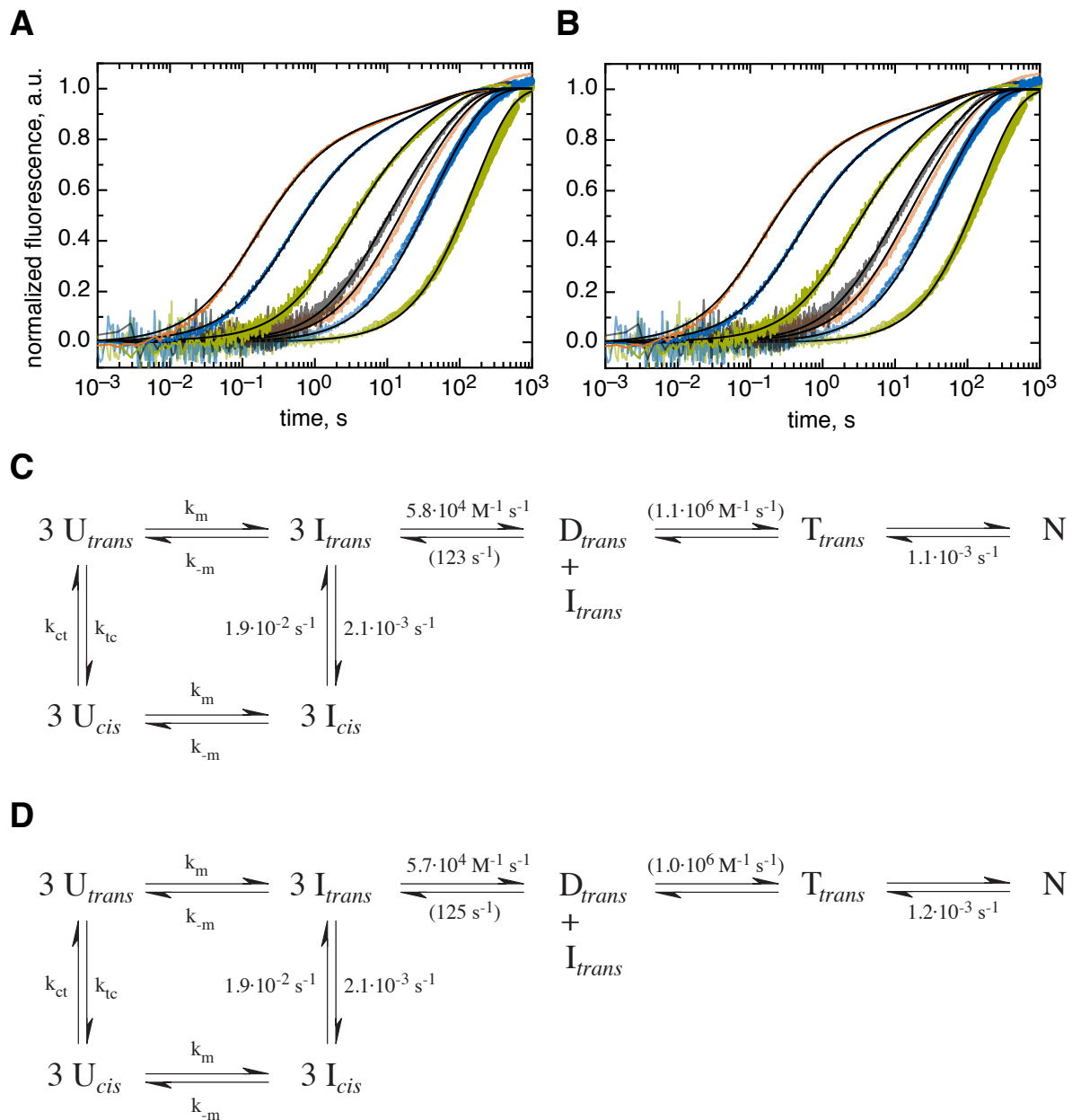


Figure 4.11.: Normalized global fit traces of foldon A6Abu urea refolding kinetics. A) Normalized urea refolding traces after fitting with fixed dimerization rate constant k_D . B) Normalized urea refolding traces after fitting with all parameters run freely. C) Fitting model with all determined rate constants with fixed dimerization rate constant k_D . D) Fitting model with all determined rate constants with all parameters run freely.

4.4. Foldon A6S

Similar to the A6Abu structure, the X-ray structure of foldon A6S shows that the hydroxyl-group of the serine side chain at position 6 points away from the tryptophan residue at position 20 and towards the arginine at position 15 of the neighboring chain, resulting in a steric hindrance of the oxygen of the side chain with the C_γ of Arg15. The side chain of this serine is not oriented in a fashion that allows direct solvent interaction of the hydroxyl group. It is probably energetically more favorable to bury the hydroxyl group within the hydrophobic cluster than a distortion of the backbone to orient the side chain towards the solvent. The only surface accessible region of the residue 6 is the backbone, which has a solvent-accessible surface of only 210.45 \AA^2 , or about 2.2 % of the total surface area of the native foldon trimer, calculated using MacPyMOL¹⁶³.

For foldon A6S, almost the same characteristics of the fluorescence spectra are observed as for foldon A6Abu. The maximum of the native state spectrum is decreased to approximately 60 % of the wild-type spectrum with no significant deviation of the shape. The spectrum of the unfolded state is decreased by approximately 20 % at the maximum. The spectrum of the burst-phase intermediate is red-shifted by 10 nm and has a slightly lower maximum, but the overall shape highly resembles the foldon wild-type intermediate spectrum. The stabilities for foldon A6S determined by denaturant transitions are $\Delta G^0(\text{H}_2\text{O}) = (-70.2 \pm 0.3) \text{ kJ} \cdot \text{mol}^{-1}$ with $m_{eq} = (14.2 \pm 0.1) \text{ kJ} \cdot \text{mol}^{-1} \cdot \text{M}^{-1}$ in GdmCl and $\Delta G^0(\text{H}_2\text{O}) = (-65.3 \pm 0.7) \text{ kJ} \cdot \text{mol}^{-1}$ with $m_{eq} = (5.8 \pm 0.2) \text{ kJ} \cdot \text{mol}^{-1} \cdot \text{M}^{-1}$ in urea. The global fit results in $\Delta G^0(\text{H}_2\text{O}) = (-68.1 \pm 0.5) \text{ kJ} \cdot \text{mol}^{-1}$ with $m_{eq}(\text{GdmCl}) = (13.1 \pm 0.3) \text{ kJ} \cdot \text{mol}^{-1} \cdot \text{M}^{-1}$ and $m_{eq}(\text{urea}) = (6.4 \pm 0.1) \text{ kJ} \cdot \text{mol}^{-1} \cdot \text{M}^{-1}$.

Unfolding traces of foldon A6S can, as for A6Abu, be fitted with a single or double-exponential equation without significant deviations of the residuals (Figure A.12A in Appendix A.4). One-exponential fits are possible for GdmCl concentrations higher than 4 M. The second phase observed for GdmCl concentrations lower than 4 M contributes up to 20 % to the overall amplitude, while the rate constants are similar to the ones of the major amplitude change. Omitting of this second phase however leads to a strong systematic deviation of the fit. Plotting of the common logarithms of the apparent rate constants of the major phase against the final GdmCl concentrations yields a plot which can be fitted with two linear equations, one for the lower and one for the higher GdmCl concentrations (Figure 4.13C). The dissociation con-

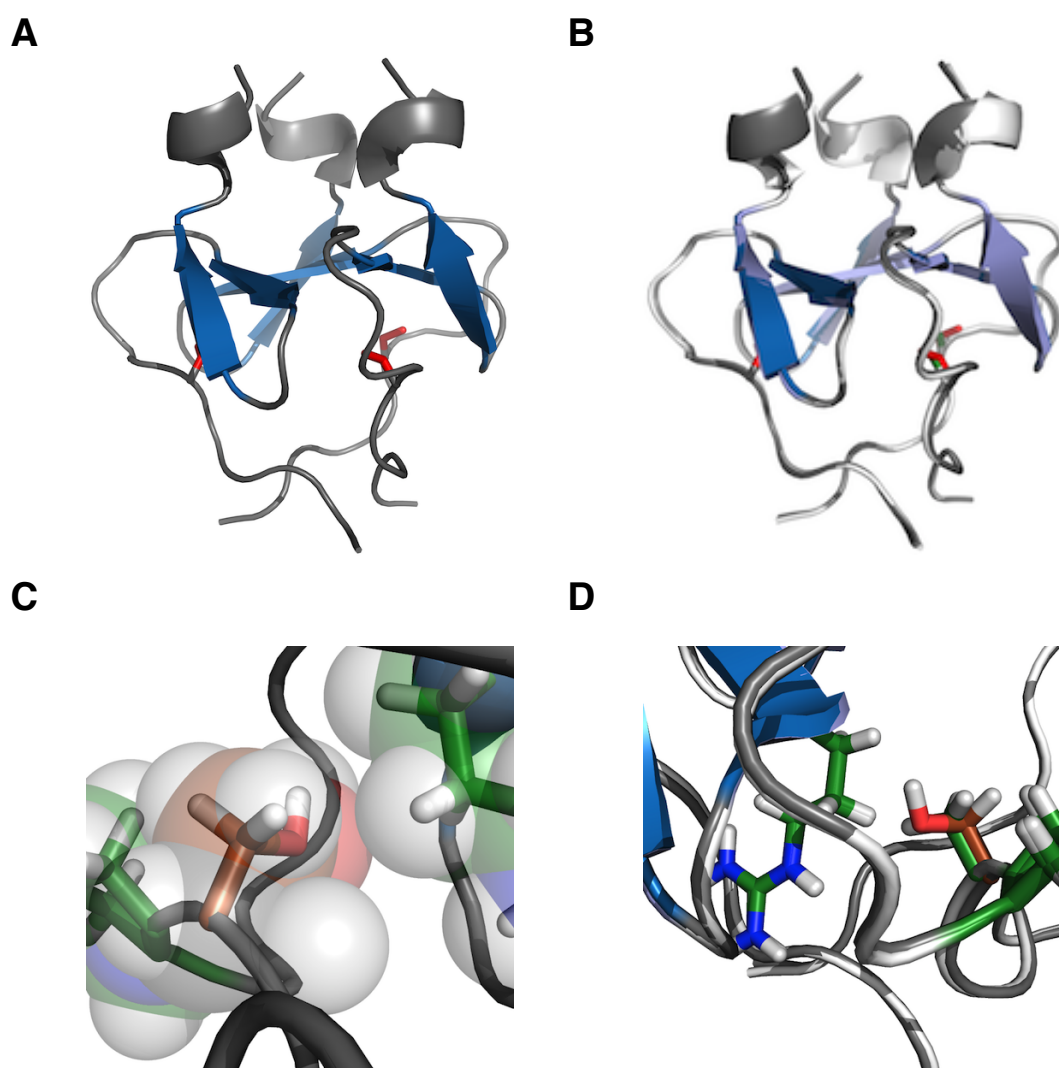


Figure 4.12.: Crystal structure of foldon A6S. A) Crystal structure of foldon A6S in ribbon representation, with the β -strands highlighted in blue, while the 3_{10} -helix and the other regions of the structure are colored in grey. The side chain of the mutated residue 6 is shown in stick representation and the side chain colored in red. B) Backbone alignment of foldon A6S and foldon wild-type. Foldon wild-type, also shown in ribbon representation, is colored in white, with the exception of the β -strands, which are colored in light blue. The alanine 6 residue is shown in stick representation and colored in green. The foldon A6S crystal structure is shown and colored as in A). C) Stick representation of the residues Ser6, Pro7 and Arg15 colored as described in Figure 4.2. The van der Waals-radii for all atoms are shown to determine whether or not steric hindrance occurs due to the inserted hydroxyl-group of the serine side chain. D) Overlay of the residue 6 side chains of both foldon wild-type and foldon A6Abu, together with the neighboring residues Pro7 and Arg15. The stick representation of these amino acids shows all carbon atoms in green (with the exception of Abu6, where the carbon atoms are shown in brown), the nitrogen atoms in blue and the hydrogen atoms in white.

stant is determined as $k_{-T} = (3.52 \pm 0.01) \text{ s}^{-1}$, with a $m_u = (4.57 \pm 0.01) \text{ kJ} \cdot \text{mol}^{-1} \cdot \text{M}^{-1}$. The dissociation of the native trimer is approximately 20fold faster than in foldon wild-type, which is very likely caused by the reduced stability of foldon A6S.

For foldon A6S urea unfolding experiments, a two-exponential equation is sufficient to yield a good fit, although some systematic deviation from the fit is seen in the residuals for $t > 10 \text{ s}$. This can be neglected as the amplitude is too low to be fitted with an additional phase, as the fitting would rather produce fitting artifacts than follow this phase. Additionally, as observed before for foldon A6Abu, the phase is very likely due to a photo-bleaching artifact. The equilibrium is shifted by 2 molar to lower urea concentrations compared to the foldon A6Abu chevron plot in urea, making the refolding limb almost undefined. This shift in the equilibrium must be due to a kinetic effect of the mutation, resulting in a change of the equilibrium constant of one of the reaction steps, as the free folding enthalpy of foldon A6S and foldon A6Abu is almost identical. Fitting of equation 3.17 to the data points yields $k_f = (8.6 \pm 4.3) \cdot 10^{-2} \text{ s}^{-1}$ and $k_u = (1.9 \pm 0.3) \cdot 10^{-3} \text{ s}^{-1}$; the kinetic m -values are $m_f = (2.0 \pm 0.5) \cdot 10^3 \text{ kJ} \cdot \text{mol}^{-1} \cdot \text{M}^{-1}$ and $m_u = (1.2 \pm 0.1) \cdot 10^3 \text{ kJ} \cdot \text{mol}^{-1} \cdot \text{M}^{-1}$, respectively. Comparison between the unfolding rate constants k_u of foldon A6Abu and A6S reveals a major difference between these two variants, as unfolding of foldon A6S is accelerated by almost one order of magnitude. The m_u -values of both foldon variants at position 6 are almost identical, in contrast to the m_f -values, which deviate by a factor of two. The reaction orders of A6S refolding experiments show broken reactions for both GdmCl and urea refolding traces with (2.57 ± 0.02) and (2.65 ± 0.08) for GdmCl and urea, respectively. The k_D 's determined from the initial slope in GdmCl and urea are $(2.06 \pm 0.13) \cdot 10^5 \text{ M}^{-1} \cdot \text{s}^{-1}$ and $(5.25 \pm 0.08) \cdot 10^4 \text{ M}^{-1} \cdot \text{s}^{-1}$, respectively, while the reaction orders are 2.06 ± 0.14 and 1.93 ± 0.08 .

For foldon A6S, the stability finally used to globally fit the refolding traces is $\Delta G^0(\text{H}_2\text{O}) = -70.2 \pm \text{kJ} \cdot \text{mol}^{-1}$, the stability at 0 M GdmCl determined by the GdmCl transition. The deviations of the "free" fit in GdmCl are small compared to the fixed fit with constrained rate constants k_D and k_{-T} . The dimerization and trimerization rate constants are decreased by about one order of magnitude compared to the foldon wild-type rates, while the dissociation rate constant of the dimer is five-fold lower compared to A6Abu. This finding shows that despite the almost identical steric hindrance of the mutations, the hydroxyl group seems to have a slightly stabilizing effect on the dimer.

4. Results

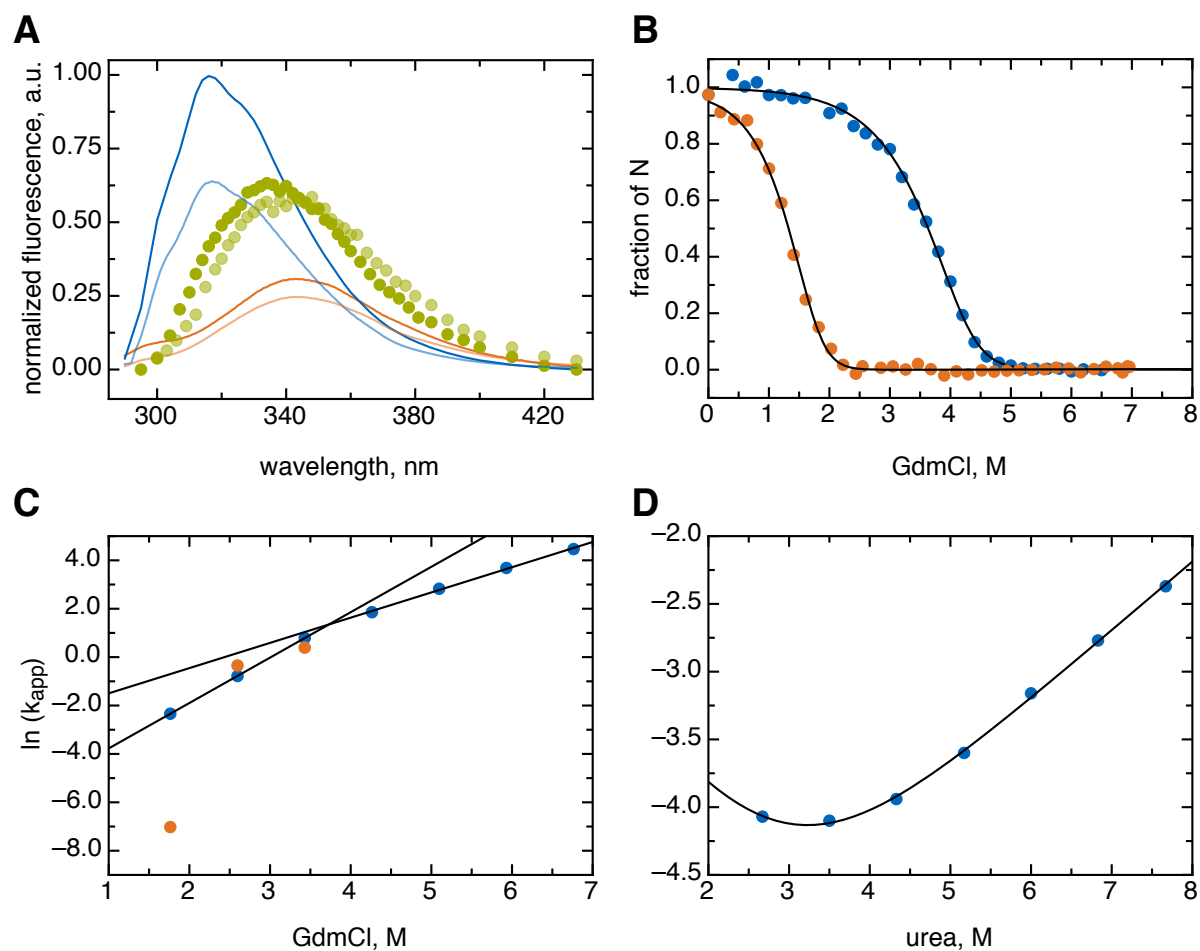


Figure 4.13.: A) Fluorescence spectra of foldon wild-type and foldon A6S. The native state spectra are colored in blue, the spectra of the unfolded state in orange and the burst-phase intermediate spectra in green. The fluorescence spectra of foldon A6S are lighter colored than the corresponding foldon wild-type spectra. B) Comparison of the foldon wild-type (blue) and foldon A6S (orange) GdmCl transitions at $30 \mu\text{M}$ foldon monomer concentration. C) Linear unfolding plot of the two phases of the foldon A6S GdmCl unfolding kinetics against the corresponding GdmCl concentrations. D) Chevron plots for foldon A6S urea unfolding. The common logarithm of the apparent unfolding rate constants is plotted against the varying urea concentrations, and the data points are fitted using equation 3.17, derived in Chapter 3.9.

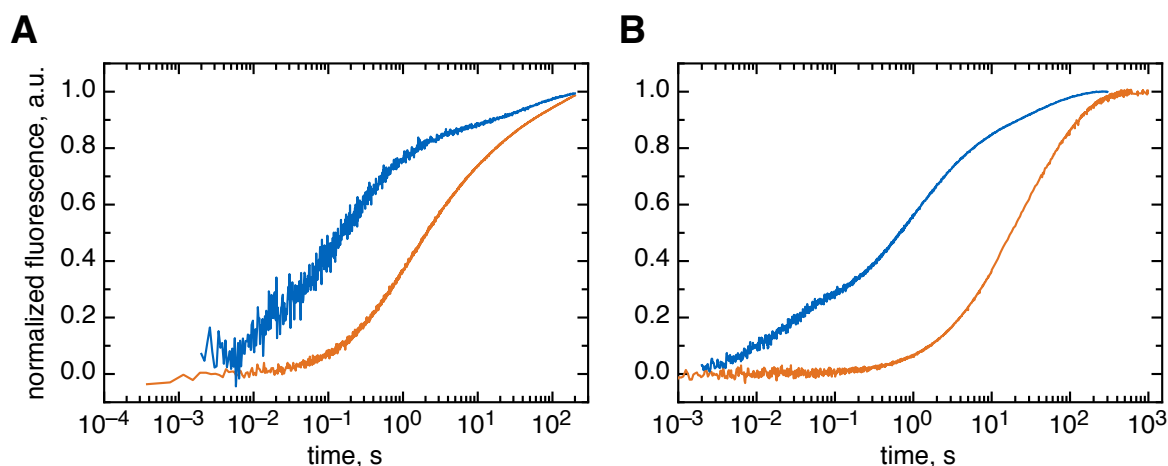


Figure 4.14.: Comparison of foldon wild-type (blue) and foldon A6S (orange) refolding traces. A) 10 μM foldon monomer refolding traces in 0.58 M GdmCl. B) 10 μM foldon monomer refolding traces in 1.0 M urea.

The fitting of the urea refolding traces of foldon A6S yields only minor statistical deviation of the fit from the data points for the lower foldon monomer concentrations. As in GdmCl, the dimerization rate constant is decreased significantly in urea. The decrease is almost identical to the one seen for foldon A6Abu, as is the trimerization rate constant. In contrast to both foldon wild-type and foldon A6Abu, the dissociation rate constants are increased for both the native state and the trimeric intermediate k_{-N} and k_{-T} . The effect of the hydroxyl group is thus not seen during association, but in a faster dissociation. The effect is however too small to have a significant impact on the free folding enthalpy, as otherwise a larger difference to foldon A6Abu would have been observed.

4. Results

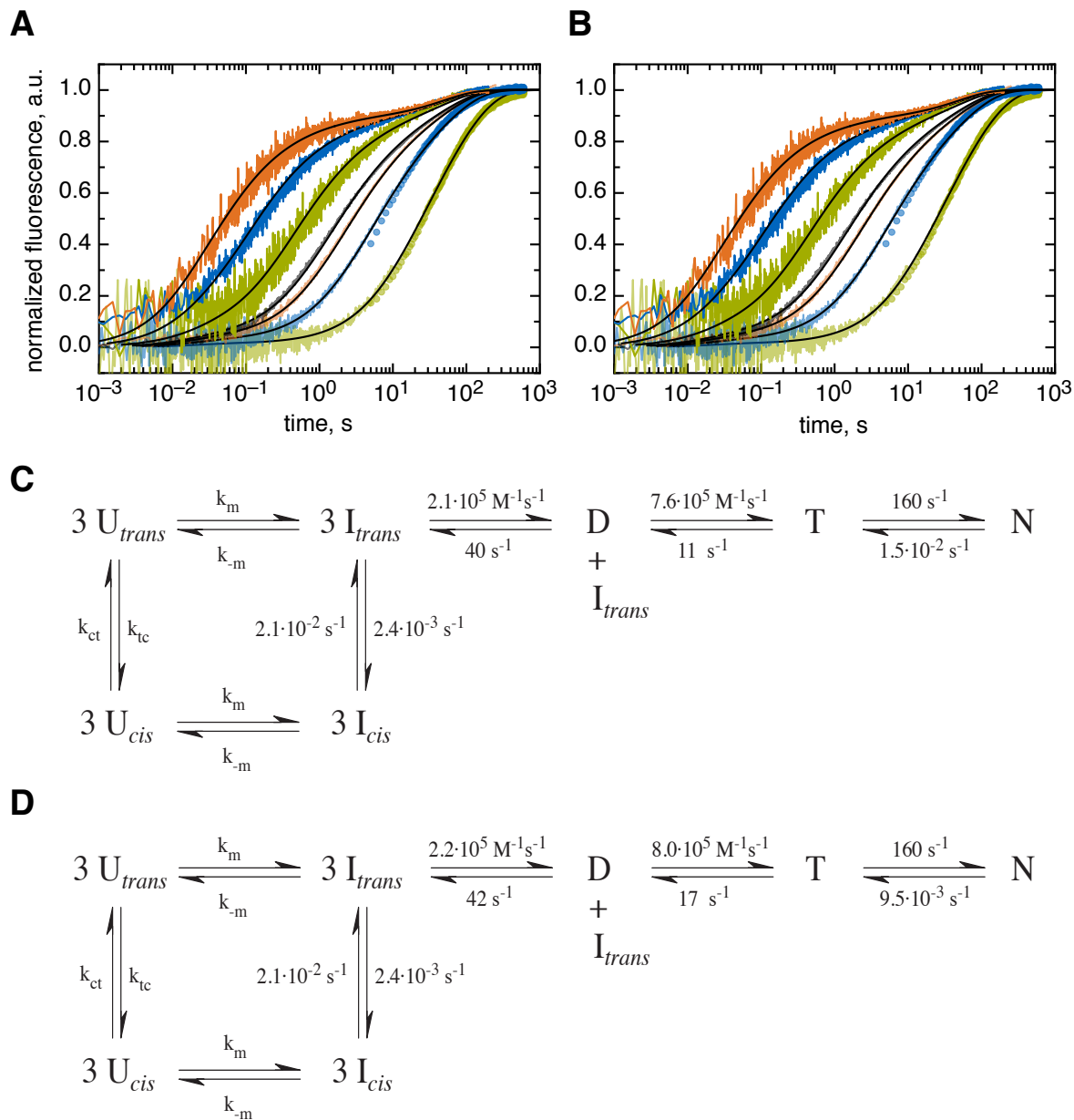


Figure 4.15.: Normalized global fit traces of foldon A6S GdmCl refolding kinetics. A) Normalized GdmCl refolding traces after fitting with set dimerization rate constant k_D . B) Normalized GdmCl refolding traces after fitting with all parameters run freely. C) Fitting model with all determined rate constants with set dimerization rate constant k_D . D) Fitting model with all determined rate constants with all parameters run freely.

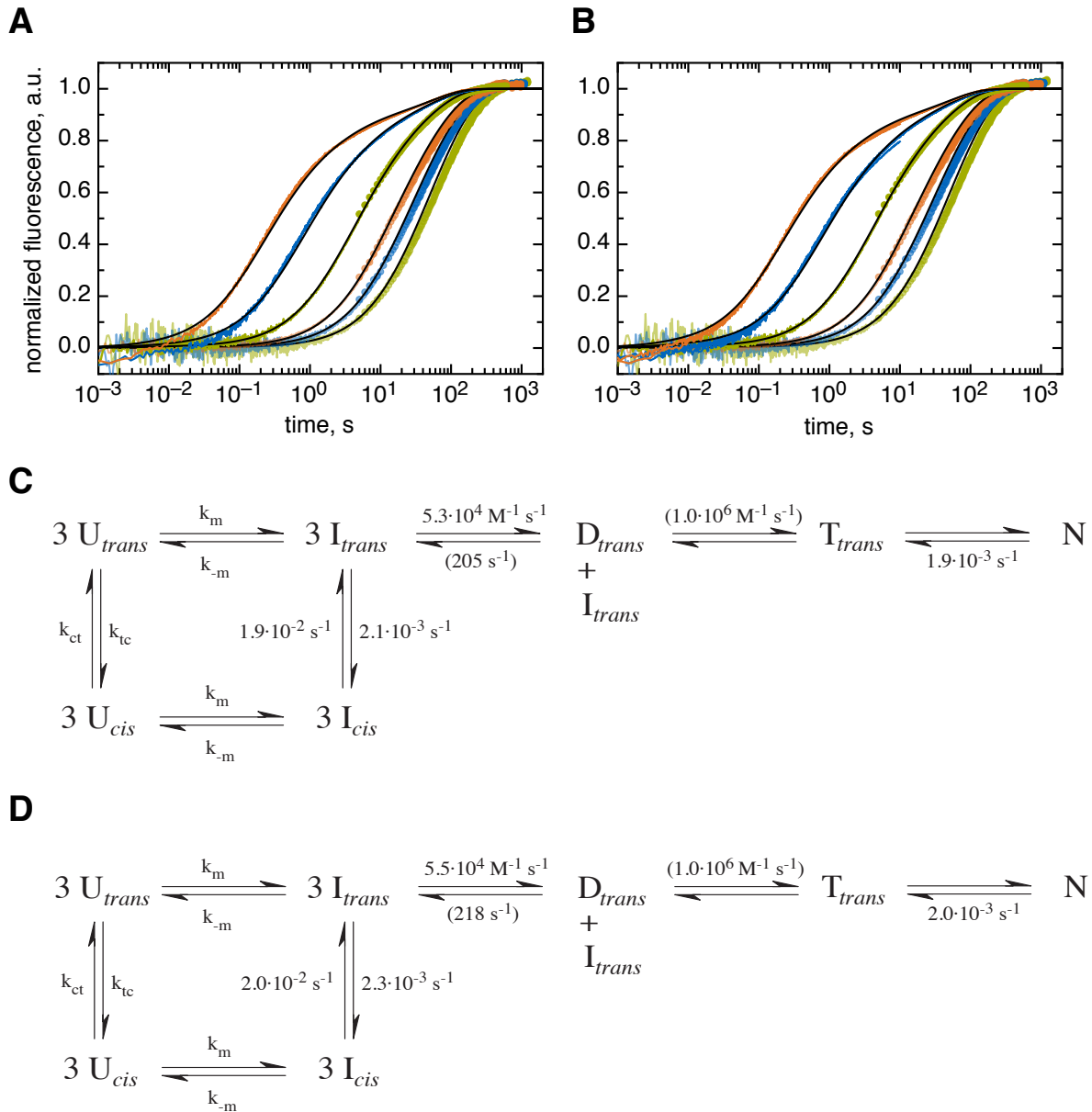


Figure 4.16.: Normalized global fit traces of foldon A6S urea refolding kinetics. A) Normalized urea refolding traces after fitting with fixed dimerization rate constant k_D . B) Normalized urea refolding traces after fitting with all parameters run freely. C) Fitting model with all determined rate constants with fixed dimerization rate constant k_D . D) Fitting model with all determined rate constants with all parameters run freely.

4.5. Foldon A6V

Foldon A6V is one of the two foldon variants that did not form crystals that allowed a structure determination by X-ray diffraction. Nevertheless the stability and kinetics of foldon A6V can be determined in GdmCl. The native state fluorescence spectrum of foldon A6V does not show any characteristics of the foldon wild-type native state fluorescence spectrum any more. In fact, it strongly resembles the spectrum of the unfolded state of foldon wild-type. The only difference is a slight blue-shift of the maximum of the native state and the absence of the characteristic shoulder with a maximum around 303 nm, indicative for tyrosine fluorescence. In the native state, tyrosine fluorescence in foldon is quenched by fluorescence resonance energy transfer (FRET) to the tryptophan. The missing of this shoulder indicates that FRET between the tyrosine residues and the tryptophan is still possible. Thus the native state spectrum does not show completely unfolded foldon, seconded by the fact that the unfolded state fluorescence spectrum shows an even lower fluorescence intensity. The burst-phase intermediate of foldon A6V shows the highest fluorescence intensity of all foldon A6V fluorescence spectra, though still lower than the intermediate spectrum of foldon wild-type. The fluorescence intensities at 320 nm of the native and intermediate spectrum are almost identical, thus it has to be shown that kinetic measurements at this wavelength are possible. The intermediate spectrum of foldon A6V has a maximum at 348 nm, which is red-shifted by almost 10 nm with respect to the foldon wild-type spectrum. Additionally, the shape of the spectrum is changed from a smooth curve to a more pronounced peak.

The stability determined for foldon A6V in GdmCl is $\Delta G^0(\text{H}_2\text{O}) = (-62.6 \pm 0.3) \text{ kJ} \cdot \text{mol}^{-1}$ with $m_{eq} = (11.9 \pm 0.1) \text{ kJ} \cdot \text{mol}^{-1} \cdot \text{M}^{-1}$. The insertion of an additional methyl group by the valine results in an additional shift in the $N \rightleftharpoons U$ equilibrium towards U by two orders of magnitude compared to foldon A6Abu or foldon A6S, while the insertion of either the first methyl- or the hydroxyl group shifts the equilibrium by almost three orders of magnitude for a 1 M foldon monomer concentration. Foldon A6V is directly comparable to foldon A6S in its unfolding behavior. The kink in the unfolding trace is again seen at approximately 4 M GdmCl concentration. Fitting of the unfolding traces is achieved by fitting with a single-exponential equation for all chosen GdmCl concentrations, the dissociation rate constant of the native trimer is determined as $k_{-T} = (3.98 \pm 0.11) \cdot 10^{-3} \text{ s}^{-1}$. Due to the low stability and the missing structure data, no experiments were carried out in urea.

GdmCl refolding experiments show a clear increase of the tryptophan fluorescence intensity upon refolding in the stopped flow. The reaction order determined from the half-life times of the GdmCl refolding experiments is shifted to 2.05 ± 0.09 , indicating that one of the association rate constants can govern the reaction order if the rate constants become slow enough. The reaction order determined from the initial slope is 2.15 ± 0.15 , with a $k_D = (1.45 \pm 0.19) \cdot 10^4 \text{ M}^{-1} \cdot \text{s}^{-1}$.

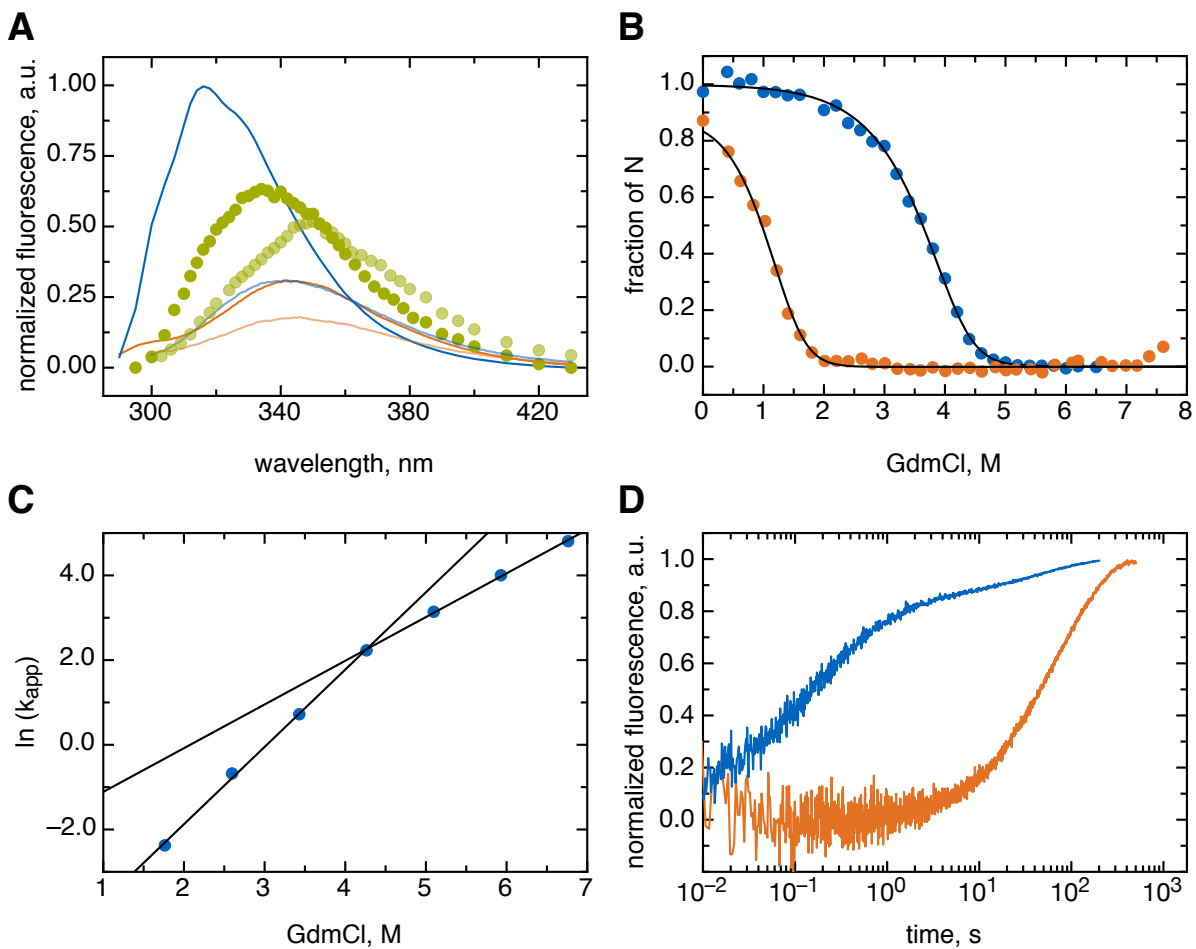


Figure 4.17.: A) Fluorescence spectra of foldon wild-type and foldon A6V. The native state spectra are colored in blue, the unfolded state in orange and the burst-phase intermediate in green. The fluorescence spectra of foldon A6V are lighter colored than the corresponding foldon wild-type spectra. B) Comparison of the foldon wild-type (blue) and foldon A6V (orange) GdmCl transitions at $30 \mu\text{M}$ monomer concentration. C) Linear plot of the common logarithm of the apparent rate constant of the foldon A6V GdmCl unfolding experiments against the corresponding GdmCl concentrations. D) Comparison of $10 \mu\text{M}$ monomer concentration foldon wild-type (blue) and foldon A6V (orange) refolding traces in 0.58 M GdmCl.

4. Results

The fits and rate constants in GdmCl for foldon A6V are given in Figure 4.18. The assumed stability used for the adjustment of the refolding traces is $61.6 \text{ kJ} \cdot \text{mol}^{-1}$. Despite this correction, the fits of foldon A6V are the least-defined of all global GdmCl refolding fits. Both association rate constants are decreased compared to foldon wild-type and the dimerization rate constant is the slowest rate constant of all GdmCl refolding fits with a 260fold decrease compared to foldon wild-type. The dissociation rate constant of the dimer k_{-D} on the other hand is the lowest determined for all GdmCl refolding experiments. Any formed dimer has thus a higher probability to bind a third monomer rather than to dissociate. The decrease of the trimerization rate constant by two orders of magnitude is larger than for most other foldon variants, with the exception of foldon P7Nva, discussed below.

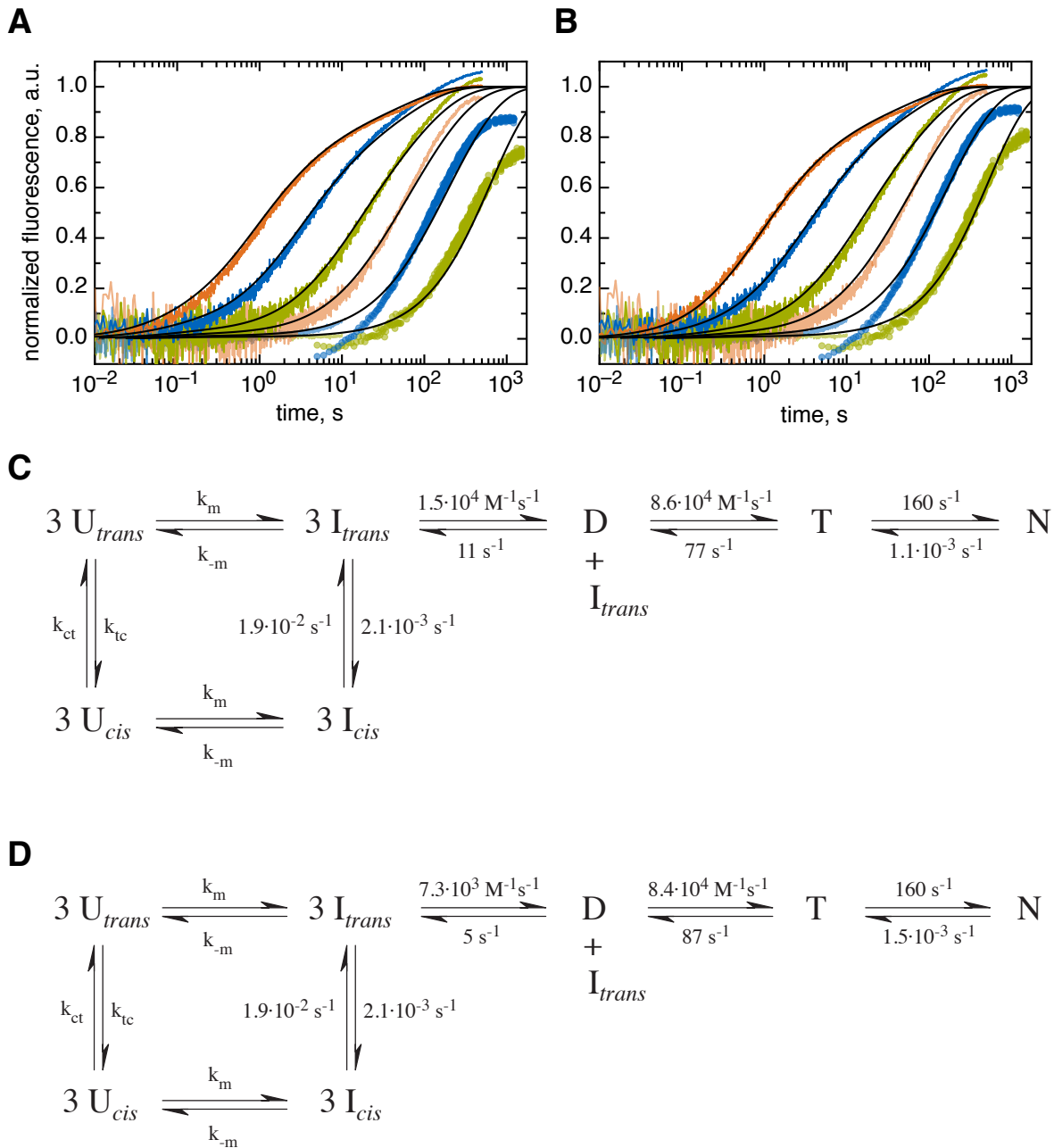


Figure 4.18.: Normalized global fit traces of foldon A6V GdmCl refolding kinetics. A) Normalized GdmCl refolding traces after fitting with fixed dimerization rate constant k_D . B) Normalized GdmCl refolding traces after fitting with all parameters run freely. C) Fitting model with all determined rate constants with fixed dimerization rate constant k_D . D) Fitting model with all determined rate constants with all parameters run freely.

4.6. Foldon P7F

The X-ray structure of foldon P7F shows that the conformation of the backbone is slightly distorted compared to foldon wild-type, with ϕ - and ψ -angles of foldon wild-type of $\phi = -54.22^\circ$ and $\psi = 144.84^\circ$, and $\phi = -50.37^\circ$ and $\psi = 142.37^\circ$ in foldon P7F. This is very likely induced by the side chain, which points into the solvent and not towards the hydrophobic cluster, in contrast to the proline or norvaline side chain (shown below). The norvaline side chain has enough conformational freedom to prevent steric hindrances and clashes with the tyrosine side chain at position 13 while still maintaining the native-like backbone conformation. The opposing effect is observed for the phenyl side chain. A steric hindrance of the phenylalanine side chain with the tyrosine side chain at position 13 changes the backbone conformation. This leads to a loss of direct interaction with the tryptophan at position 20 of the hydrophobic cluster. The gap formed between the phenylalanine at position 7 and the tryptophan at position 20 is only slightly covered by the phenylalanine side chain. The steric hindrance with Tyr13 additionally causes a tilt of the tyrosine 13 side chain. This distortion leads to a further destabilization of the hydrophobic cluster. Foldon P7F shows the lowest stability determined in GdmCl with $\Delta G^0(\text{H}_2\text{O}) = (-46.5 \pm 11.3) \text{ kJ} \cdot \text{mol}^{-1}$ and $m_{eq} = (13.2 \pm 4.2) \text{ kJ} \cdot \text{mol}^{-1} \cdot \text{M}^{-1}$. This low stability would render most refolding experiments impossible due to the low fraction of natively formed trimer at the refolding conditions, thus no further experiments were performed.

4.7. Foldon P7V

Next to foldon A6V, foldon P7V is the only foldon variant in this study which did not form any suitable crystals for X-ray structure determination. The stability of foldon P7V is determined by GdmCl transitions, resulting in a free folding enthalpy of $\Delta G^0(\text{H}_2\text{O}) = (-53.0 \pm 6.0) \text{ kJ} \cdot \text{mol}^{-1}$ with $m_{eq} = (10.7 \pm 2.4) \text{ kJ} \cdot \text{mol}^{-1} \cdot \text{M}^{-1}$. Due to the low stability and the lack of structural data, no further experiments were carried out using this foldon variant.

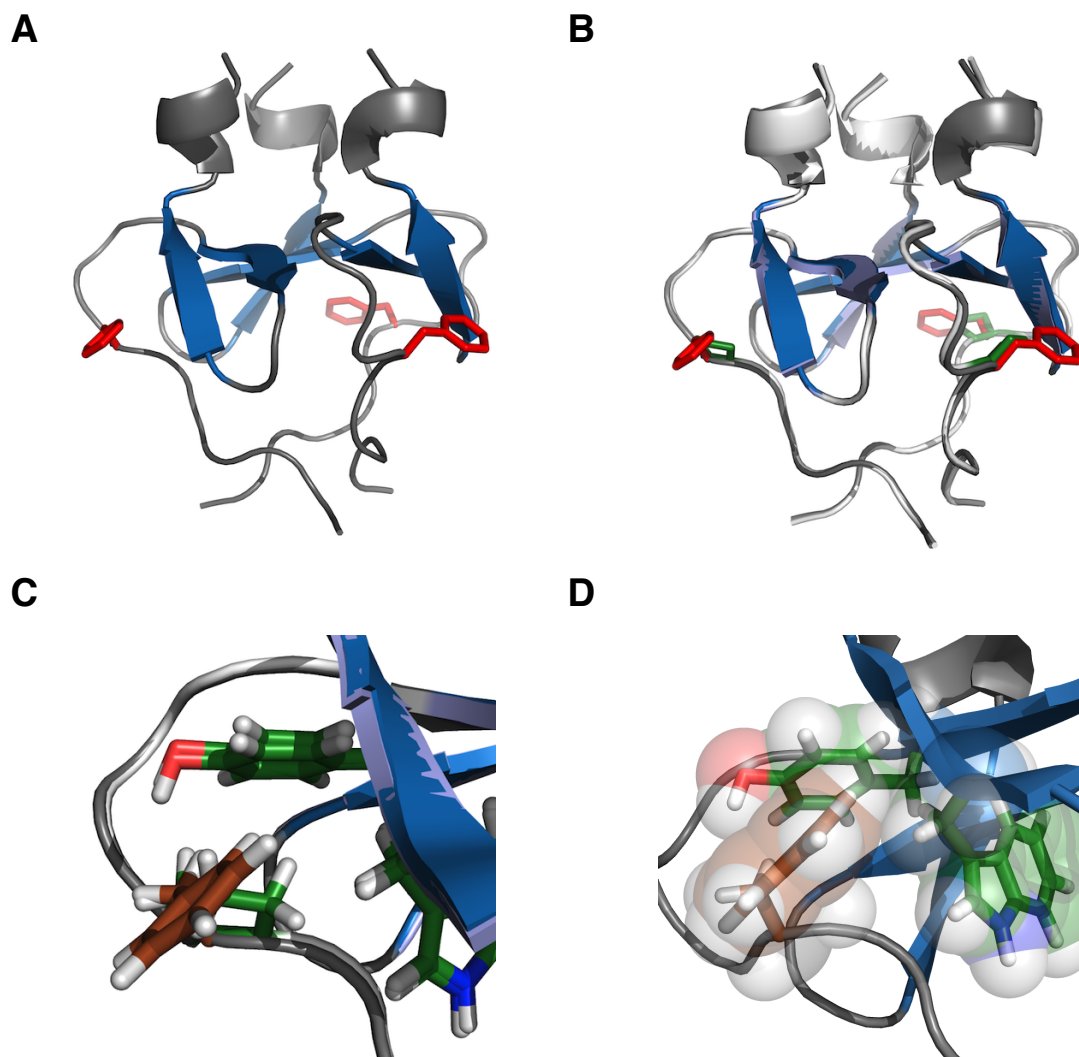


Figure 4.19.: Crystal structure of foldon P7F. A) Crystal structure of foldon P7F in ribbon representation, with the β -strands highlighted in blue, while the 3_{10} -helix and the other regions of the structure are colored in grey. The side chain of the mutated residue 7 is shown in stick representation and the side chain colored in red. B) Backbone alignment of foldon P7F and foldon wild-type. Foldon wild-type, also shown in ribbon representation, is colored in white, with the exception of the β -strands, which are colored in light blue. The proline 7 residue is shown in stick representation and colored in green. The foldon P7F crystal structure is shown and colored as in A). C) Overlay of the residue 7 side chains of both foldon wild-type and foldon P7F, together with the neighboring residues Tyr13 and Trp20. The all-atom representation shows all carbon atoms in green (with the exception of Phe7, where the carbon atoms are shown in brown), the nitrogen atoms in blue and the hydrogen atoms in white. The tilt of the tyrosine 13 side chain due to the insertion of the phenyl-group at position 7 is shown by the overlay of the two crystal structures. D) Stick representation of the residues Pro7, Tyr13 and Trp20 colored as described in Figure 4.2. The van der Waals-radii for all atoms are shown to highlight the steric clash of the phenylalanine and the tyrosine side chains and the gap between the phenyl-ring and the tryptophan 20.

4. Results

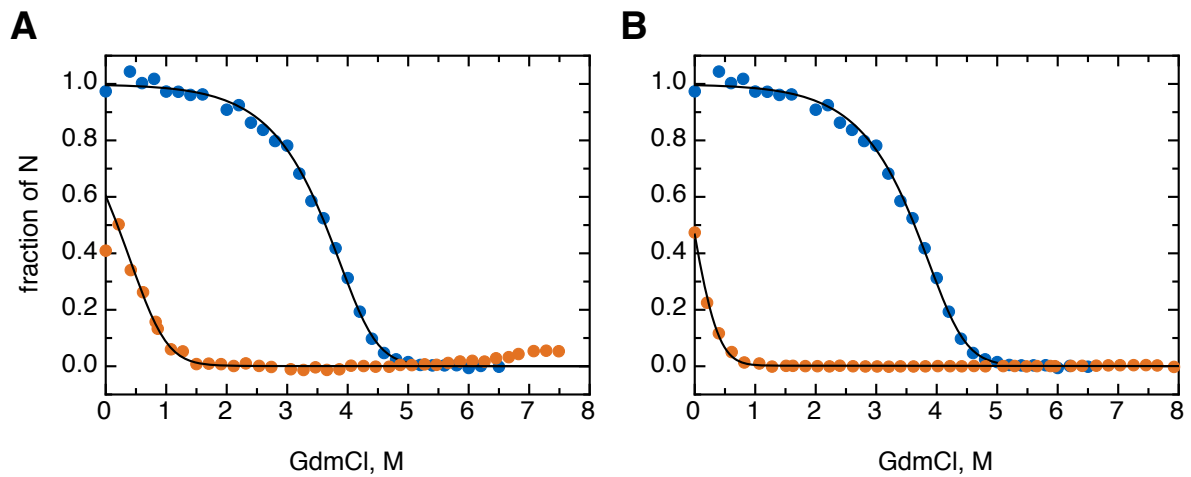


Figure 4.20.: GdmCl transitions of foldon P7F and foldon P7V. A) Comparison of the foldon wild-type (blue) and foldon P7F (orange) GdmCl transitions at 30 μ M monomer concentration. B) Comparison of the foldon wild-type (blue) and foldon P7V (orange) GdmCl transitions at 30 μ M monomer concentration.

4.8. Foldon P7Nva

The backbone conformation of the X-ray structure of foldon P7Nva at position 7 is in *trans* conformation and superimposes completely with the foldon wild-type conformation although the constraint of this conformation by the proline is missing. This can be seen from the backbone angles, which are $\phi = -54.22^\circ$ and $\psi = 144.84^\circ$ for foldon wild-type and $\phi = -55.72^\circ$ and $\psi = 147.62^\circ$ in foldon P7Nva. The norvaline side chain is not constrained in the same conformation as the proline side chain, but points outwards into the solvent. This results in a gap between the residues 7 and 20, which is closed by the pyrrolidine ring in foldon wild-type. The resulting loss of van der Waals contact and the higher accessibility of the hydrophobic core towards the solvent and thus to denaturants is likely to cause the decreased stability and higher m_{eq} -value determined for this foldon variant. Additionally, the loss of the restraining pyrrolidine ring leads to a higher entropy of the backbone, hence energy is needed to constrain the N-terminal in its native conformation. This is necessary to fix the N-terminal amino acids in the correct orientation packed against the β -turn.

The native spectrum of foldon P7Nva has a high resemblance to the spectrum of the foldon A6V native state, and it is as discussed before not determinable whether this effect is due to a change in the fluorescence properties or just a stability effect. However, the destabilization is not as profound as for foldon A6V. The unfolded state has the same characteristics as observed for all previously discussed foldon variants, as the maximum is decreased and the intensity is reduced for the lower wavelength while no significant deviation is observed for the higher wavelengths. The intermediate state spectrum is red-shifted in its maximum by 10 nm and the fluorescence intensity is also slightly increased compared to foldon wild-type. This indicates that the hydrophobic core of the burst-phase intermediate has a different conformation which influences the shape of the emission spectrum while the quantum yield is not effected. The stability determined in GdmCl is $\Delta G^0(\text{H}_2\text{O}) = (-58.7 \pm 1.3) \text{ kJ} \cdot \text{mol}^{-1}$ with $m_{eq} = (12.4 \pm 0.7) \text{ kJ} \cdot \text{mol}^{-1} \cdot \text{M}^{-1}$. Foldon P7Nva is the only variant at position 7 which was chosen for studies in urea, and the free folding enthalpy in urea is $\Delta G^0(\text{H}_2\text{O}) = (-54.6 \pm 0.6) \text{ kJ} \cdot \text{mol}^{-1}$, with a $m_{eq} = (6.2 \pm 0.2) \text{ kJ} \cdot \text{mol}^{-1} \cdot \text{M}^{-1}$. The global fit of both transitions results $\Delta G^0(\text{H}_2\text{O}) = (-57.8 \pm 0.4) \text{ kJ} \cdot \text{mol}^{-1}$ with $m_{eq}(\text{GdmCl}) = (11.8 \pm 0.2) \text{ kJ} \cdot \text{mol}^{-1} \cdot \text{M}^{-1}$ and $m_{eq}(\text{urea}) = (7.2 \pm 0.2) \text{ kJ} \cdot \text{mol}^{-1} \cdot \text{M}^{-1}$.

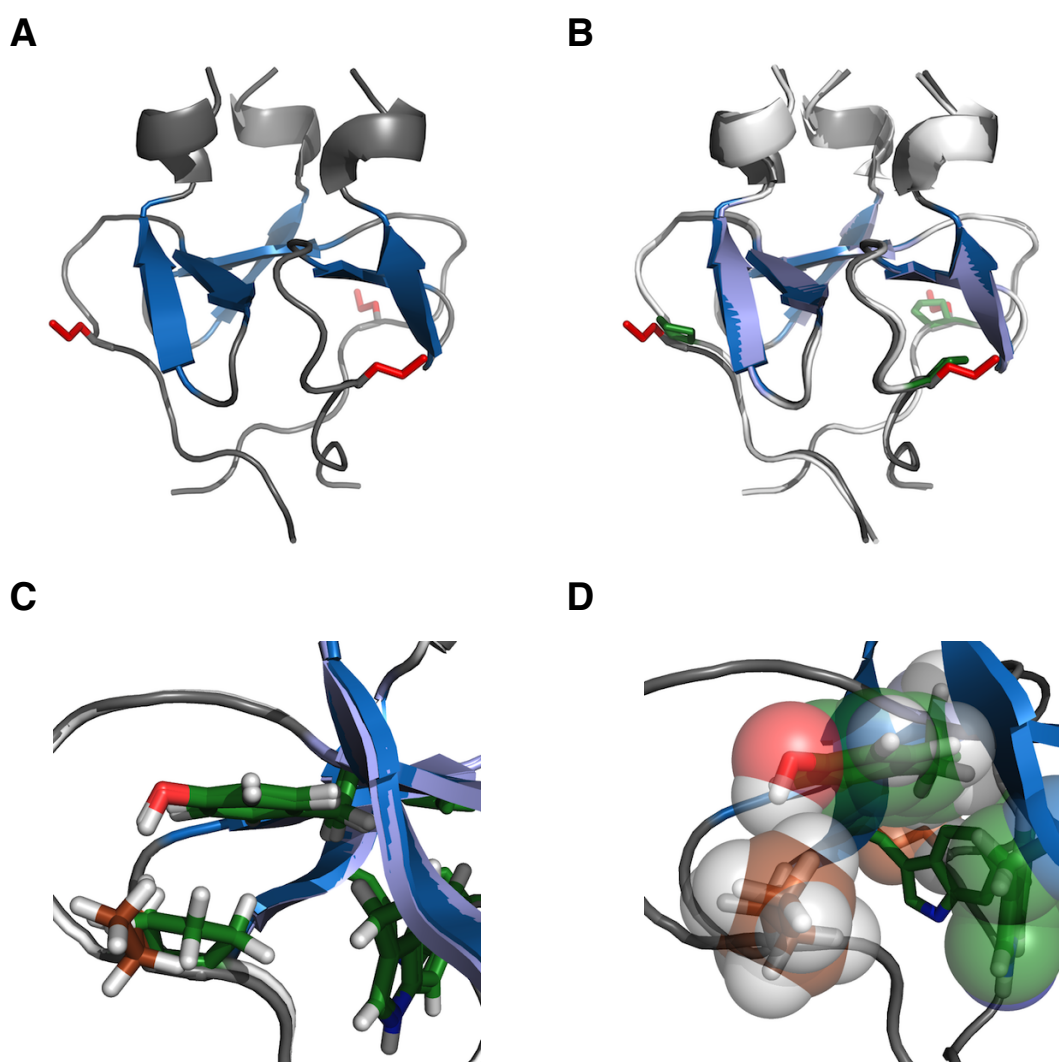


Figure 4.21.: Crystal structure of foldon P7Nva. A) Crystal structure of foldon P7Nva in ribbon representation, with the β -strands highlighted in blue, while the 3_{10} -helix and the other regions of the structure are colored in grey. The side chain of the mutated residue 7 is shown in stick representation and the side chain colored in red. B) Backbone alignment of foldon P7Nva and foldon wild-type. Foldon wild-type, also shown in ribbon representation, is colored in white, with the exception of the β -strands, which are colored in light blue. The proline 7 residue is shown in stick representation and colored in green. The foldon P7Nva crystal structure is shown and colored as in A). C) Overlay of the residue 7 side chains of both foldon wild-type and foldon P7Nva, together with the neighboring residues Tyr13 and Trp20. The all-atom representation shows all carbon atoms in green (with the exception of Nva7, where the carbon atoms are shown in brown), the nitrogen atoms in blue and the hydrogen atoms in white. D) Stick representation of the residues Pro7, Tyr13 and Trp20 colored as described in Figure 4.2. The van der Waals-radii for all atoms are shown to highlight the missing interaction of the norvaline side chain with the tryptophan side chain of residue 20, as the norvaline side chain is pointing into the solvent, in contrast to the native conformation in the proline side chain of foldon wild-type.

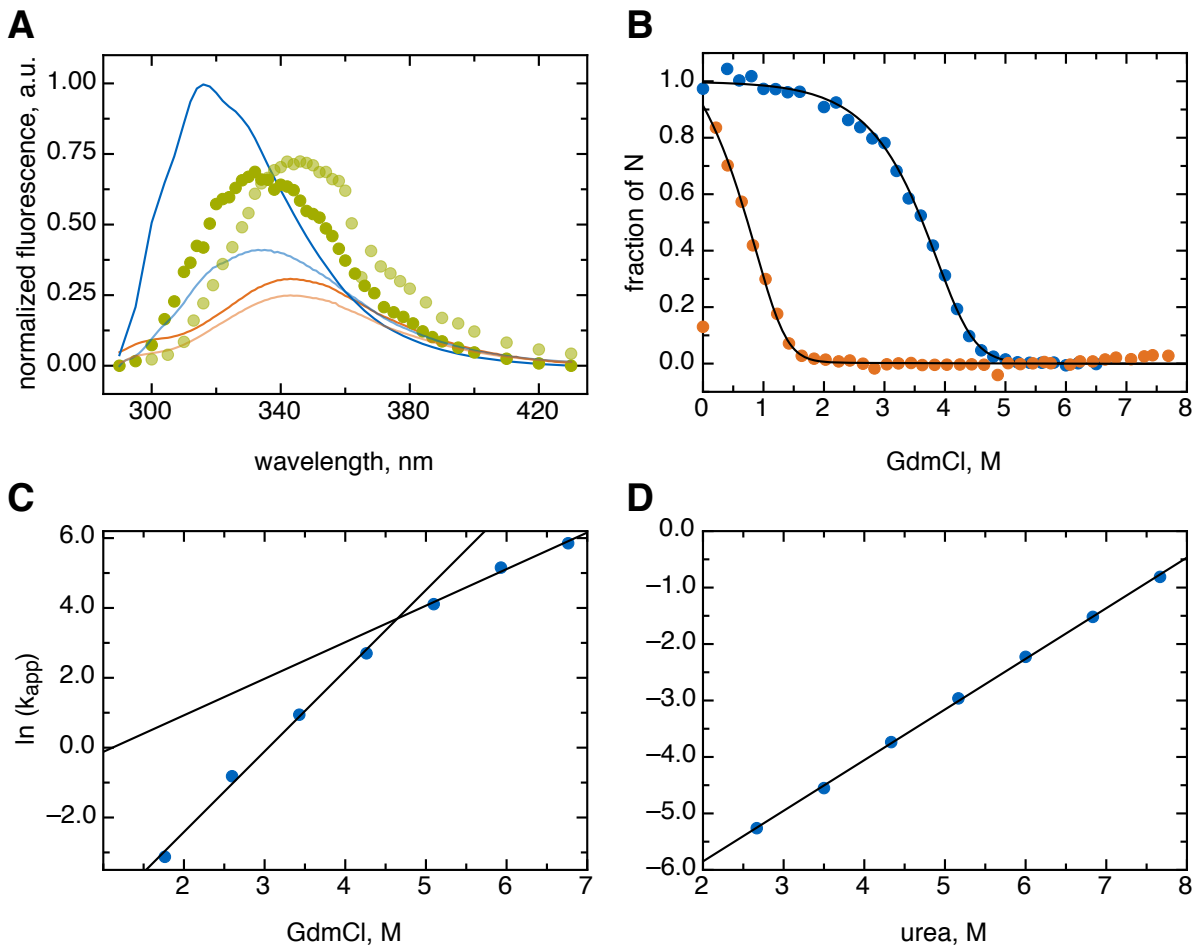


Figure 4.22.: A) Fluorescence spectra of foldon wild-type and foldon P7Nva. The native state spectra are colored in blue, the spectra of the unfolded state in orange and the burst-phase intermediate spectra in green. The fluorescence spectra of foldon P7Nva are lighter colored than the corresponding foldon wild-type spectra. B) Comparison of the foldon wild-type (blue) and foldon P7Nva (orange) GdmCl transitions at 30 μ M monomer concentration. C) Linear plot of the common logarithms of the apparent rate constants of the foldon P7Nva GdmCl unfolding experiments against the corresponding GdmCl concentrations. D) Linear plot of the common logarithms of the apparent rate constants of the foldon P7Nva GdmCl unfolding experiments against the corresponding urea concentrations.

Plotting the common logarithm of the apparent rate constants of GdmCl unfolding experiments of foldon P7Nva against the varying final GdmCl concentrations shows a kink in the unfolding limb. The kink is shifted by 0.5 M GdmCl concentration to approximately 4.5 M GdmCl compared to the variants at position 6, indicating that the high-energy intermediate has a more native-like structure. The linear fit results $k_{-T} = (3.83 \pm 0.43) \cdot 10^{-3} \text{ s}^{-1}$ and $m_U = 4.39 \pm 0.33 \text{ kJ} \cdot \text{mol}^{-1} \cdot \text{M}^{-1}$. The apparent rate constants are determined by single-exponential fitting of the unfolding traces. Fitting of the urea unfolding traces of foldon P7Nva at varying final

4. Results

urea concentrations is also achieved by a single-exponential equation. The common logarithm of the apparent rate constant is plotted against the final urea concentrations, yielding a linear slope. Thus foldon P7Nva is too unstable to allow detection of the refolding limb of the chevron plot in urea. The plot is fitted with a linear equation, yielding an unfolding rate constant $k_u = (4.79 \pm 0.03) \cdot 10^{-4} \text{ s}^{-1}$ and a $m_u = (2.19 \pm 0.02) \cdot 10^3 \text{ kJ} \cdot \text{mol}^{-1} \cdot \text{M}^{-1}$. The m_u -value is significantly larger than the ones obtained for the foldon variants at position 6, but it is not clear whether this is an effect of the mutation or of the fitting procedure. The reaction orders determined from the half-life times in both GdmCl and urea are close to 2, indicating a strong influence of one of the association rate constants on the folding reaction. A reaction order of 2 is also found by determination of the initial slopes. The dimer association rate constant determined in this way for GdmCl and urea refolding are $(1.03 \pm 0.12) \cdot 10^4 \text{ M}^{-1} \cdot \text{s}^{-1}$ and $(5.89 \pm 0.09) \cdot 10^2 \text{ M}^{-1} \cdot \text{s}^{-1}$, respectively.

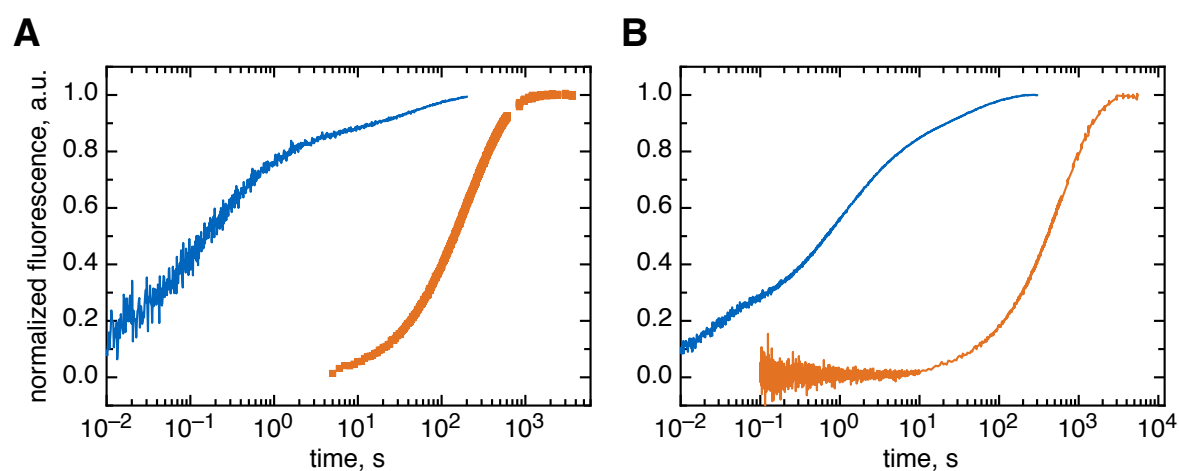


Figure 4.23.: Comparison of foldon wild-type (blue) and foldon P7Nva (orange) refolding traces A) 10 μM foldon monomer refolding traces in 0.58 M GdmCl. B) 10 μM foldon monomer refolding traces in 1.0 M urea.

The model used for the global fit of the foldon P7Nva refolding traces does not show any *cis-trans* isomerization reaction of proline peptide bonds. This is feasible due to the finding that only the proline at position 7 is responsible for the *cis-trans* isomerization, as known from the mutations at position 4 and 6. The fit model is thus reduced to a linear model consisting of dimerization, trimerization and the rearrangement step from the trimeric intermediate to the native state for GdmCl fitting. The refolding traces are adjusted using a free folding enthalpy of $58.4 \text{ kJ} \cdot \text{mol}^{-1}$. The fit traces show a systematic shift in the lower foldon monomer con-

centrations. Only minor changes are observed upon a global fit with all rate constants run freely. Both association rate constants are lowered by two orders of magnitude, while the dissociation rate constants of the intermediates are only slightly changed compared to foldon wild-type. The dissociation rate constant of the native state is 10fold larger than in the foldon wild-type, as expected due to the decreased stability. This value is similar to those determined for mutations at position 6, thus the free enthalpy of folding is not the sole cause for an increase or decrease of k_{-T} .

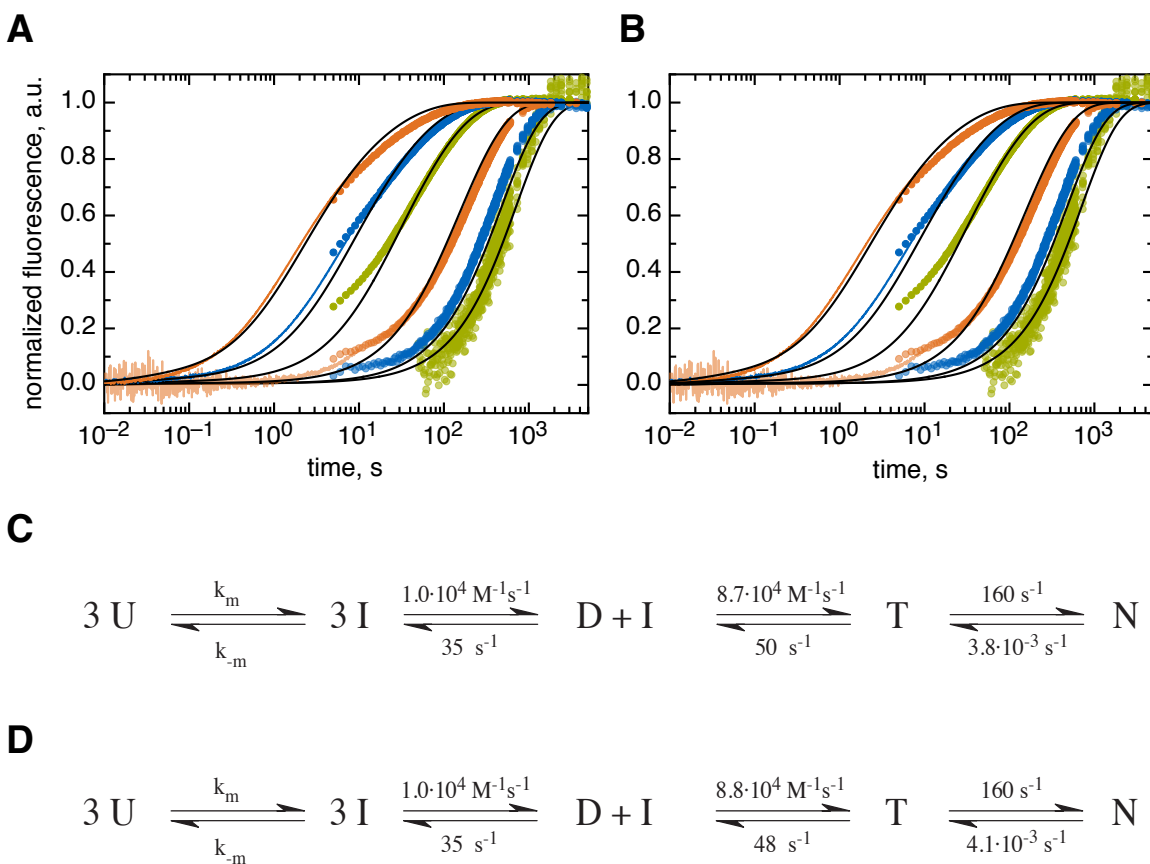


Figure 4.24.: Normalized global fit traces of foldon P7Nva GdmCl refolding kinetics. A) Normalized GdmCl refolding traces after fitting with fixed dimerisation rate constant k_D . B) Normalized GdmCl refolding traces after fitting with all parameters run freely. C) Fitting model with all determined rate constants with fixed dimerisation rate constant k_D . D) Fitting model with all determined rate constants with all parameters run freely.

4. Results

The fit mechanism for refolding of foldon P7Nva in urea is even more simple than in GdmCl as the rearrangement step from the intermediate trimer T_{trans} to the native state N is negligible in urea. The mechanism is shown in Figure 4.25C and D, together with the global fits with both fixed and free dimerization rate constant k_D .

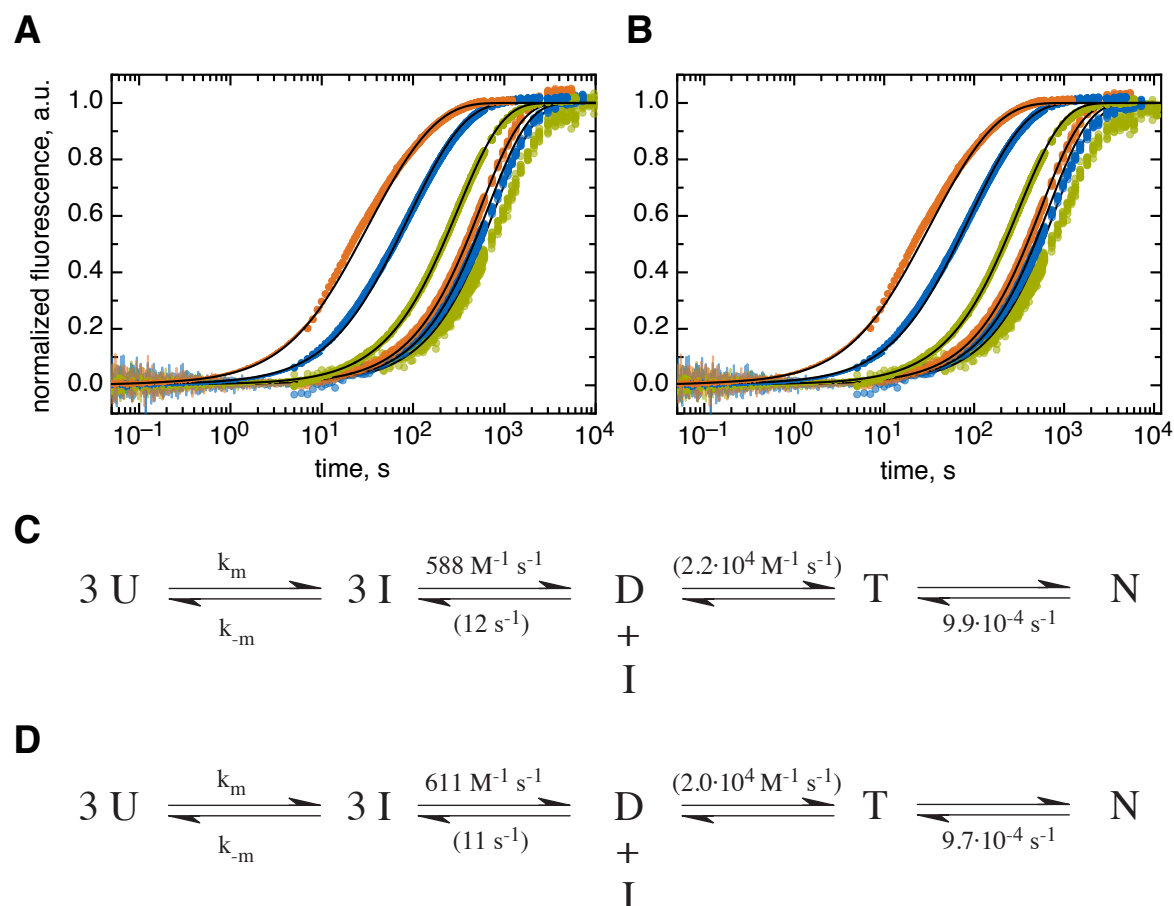


Figure 4.25.: Normalized global fit traces of foldon P7Nva urea refolding kinetics. A) Normalized urea refolding traces after fitting with fixed dimerisation rate constant k_D . B) Normalized urea refolding traces after fitting with all parameters run freely. C) Fitting model with all determined rate constants with fixed dimerisation rate constant k_D . D) Fitting model with all determined rate constants with all parameters run freely.

Foldon P7Nva shows the strongest effect on the dimerization reaction of all studied foldon variants in urea, with a decrease of the dimerization rate constant by three orders of magnitude, compared to foldon wild-type. Although only four rate constants are used in this folding model and two are already pre-determined by other experiments and evaluations, this is not

sufficient to yield robustly defined trimerization and dimer dissociation rate constants k_T and k_{-D} . These two rate constants show a strong coupling over several orders of magnitude without significantly changing the characteristics of the fit traces. The overall effect of the mutation at position 7 is that a loss of the critical *cis-trans* isomerization reaction leads to a tremendous deceleration of the dimerization reaction. The gain in backbone entropy seems to cause a large shift towards the unfolded state, as the backbone is no longer fixed in the correct conformation. This disfavors efficient dimerization of foldon monomers.

4.9. Foldon P4T P7Nva

The double-mutant foldon P4T P7Nva does not show significant differences in the X-ray structure of the native state at the mutated locations and especially not in the overall conformation of the N-terminal region when compared to both the wild-type and the single-point mutations at position 4 and 7. These comparisons are done both by an overlay of the structures and comparison of the backbone angles. In the foldon P4T variant the angles for the C_α at position 4 are $\phi = -85.68^\circ$ and $\psi = 167.87^\circ$, compared to $\phi = -83.71^\circ$ and $\psi = 163.82^\circ$ at the same position in the double-mutant. For foldon P7Nva, the angles are $\phi = -55.72^\circ$ and $\psi = 147.62^\circ$, while they are $\phi = -54.83^\circ$ and $\psi = 145.14^\circ$ in the double-mutant. Even in the absence of both proline residues, the poly-proline II structure of the N-terminal region is preserved in the native state, as seen by an alignment of both wild-type and mutant crystal structures. The side chains of the two mutated residues in the double-mutant have a different conformation compared to the single-mutant crystal structures, indicating that the mutant side chains gain conformational freedom in the native state.

The burst-phase intermediate spectrum of foldon P4T P7Nva shows the largest deviation from all other variants and foldon wild-type. A sharp peak with a maximum at 342 nm is observed that even exceeds the native state fluorescence of foldon wild-type. This spectrum indicates that the intermediate state has a tremendously changed hydrophobic core conformation in the intermediate state that is not seen in neither the native nor the unfolded state of foldon P4T P7Nva. The native and unfolded state spectra closely resemble the foldon wild-type spectra, with decreased fluorescence intensities. The change in the intermediate spectrum could indicate a changed packing of the N-terminal region of the foldon domain against the β -turn, which could lead to a strong enhancement in fluorescence.

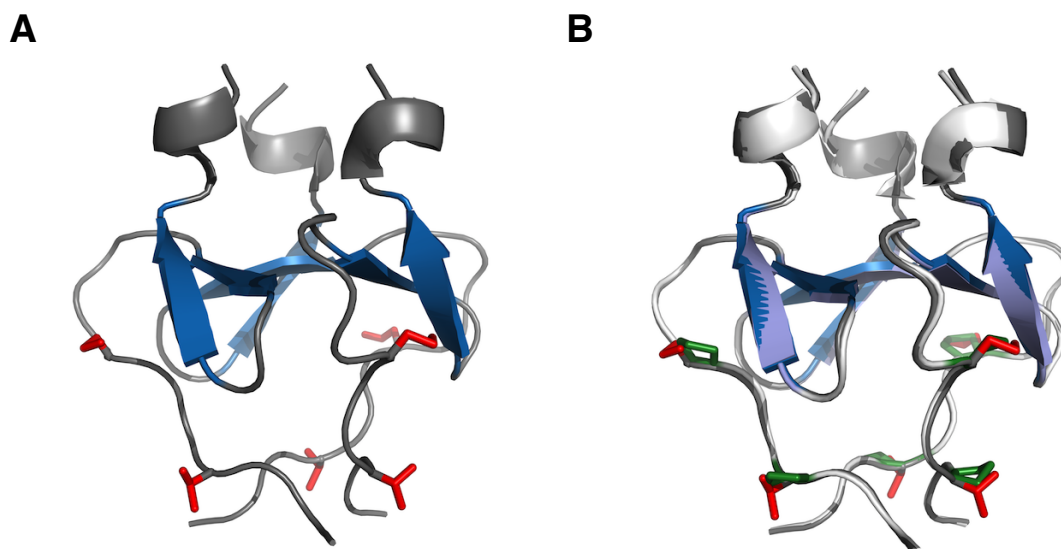


Figure 4.26.: Crystal structure of foldon P4T P7Nva. A) Crystal structure of foldon P4T P7Nva in ribbon representation, with the β -strands highlighted in blue, while the 3_{10} -helix and the other regions of the structure are colored in grey. The side chains of the mutated residues 4 and 7 are shown in stick representation and the side chain colored in red. B) Backbone alignment of foldon P4T P7Nva and foldon wild-type. Foldon wild-type, also shown in ribbon representation, is colored in white, with the exception of the β -strands, which are colored in light blue. The proline 4 and 7 residues are shown in stick representation and colored in green. The foldon P4T P7Nva crystal structure is shown and colored as in A).

As substitutions of the proline residues at position 4 and 7 lead to opposing effects regarding the stability of the native trimer, the question arises whether the effects are additive or not. The stability determined for the foldon P4T P7Nva double substitution in GdmCl is $\Delta G^0(\text{H}_2\text{O}) = (-70.2 \pm 6.0) \text{ kJ} \cdot \text{mol}^{-1}$, with an m_{eq} -value of $m_{eq} = (14.2 \pm 3.2) \text{ kJ} \cdot \text{mol}^{-1} \cdot \text{M}^{-1}$. From the comparison of the $\Delta\Delta G^0(\text{H}_2\text{O})$ values of the individual mutations ($-12.3 \text{ kJ} \cdot \text{mol}^{-1}$ for foldon P4T, $30.5 \text{ kJ} \cdot \text{mol}^{-1}$ for foldon P7Nva) with the double mutation ($19.0 \text{ kJ} \cdot \text{mol}^{-1}$), it is likely that the effects exerted by the mutations are additive. The m_{eq} -value is relatively large, as seen before for the single-point mutations at position 7.

Foldon P4T P7Nva shows the most complex unfolding kinetics of all studied variants in GdmCl. Despite its comparable stability with other foldon variants, fitting of the unfolding traces shows a significantly different unfolding behavior, both in the shape of the unfolding traces and the length of the unfolding reaction. Fitting of these traces has to be done using a three-exponential equation, and all three phases contribute significantly to the overall fluores-

cence signal change. Plotting of the common logarithms of the three apparent rate constants against the final GdmCl concentrations yields three linear unfolding traces which are grouped due to their amplitude. The lowest of these three rate constants has the largest contribution to the overall fluorescence change and is thus considered to correspond to the dissociation rate constant with $k_{-T} = (7.48 \pm 0.49) \cdot 10^{-5} \text{ s}^{-1}$. This rate constant is even 5fold lower than the dissociation rate constant of foldon wild-type, despite the overall destabilization of foldon P4T P7Nva compared to foldon wild-type.

The double mutant foldon P4T P7Nva is an exception with regard to the global GdmCl refolding fit (Figure 4.28). It is the only variant with a lower stability than foldon wild-type where the exact stability for each monomer concentration at the GdmCl refolding conditions can be used and global fitting remains possible. Additionally, no initial slope measurement is needed to gain robust fit results. It is a common observation throughout this study that the fidelity of the association rate constants is increased with lower rate constants. The other exception of foldon P4T P7Nva refolding kinetics is that the concentration-independent folding step from the trimeric intermediate to the native state N is not needed for a suitable fit any more. The same model as for foldon P7Nva in urea is used for global fitting. The reaction order determined from the half-life times of the concentration-dependent reaction steps is $t_{1/2} = 2.00 \pm 0.11$, as observed before for other foldon variants with a strong deceleration of the association steps. The dimerization rate constant is lowered by a factor of 1,000, while the dissociation rate constant of the dimer is the largest found throughout this study. The effects on the trimerization rate constant are in the same range as the other foldon variants with its 10fold decrease. The dissociation rate constant of the native state is almost identical to the dissociation rate of foldon P4T, despite the lowered stability of foldon P4T P7Nva. As seen for foldon mutations at position 6, the dissociation rate constant is not only ruled by the overall stability. Instead, an exchange of the proline at position 4 seems to raise the free activation enthalpy of unfolding, regardless of the overall stability of the foldon variant.

4. Results

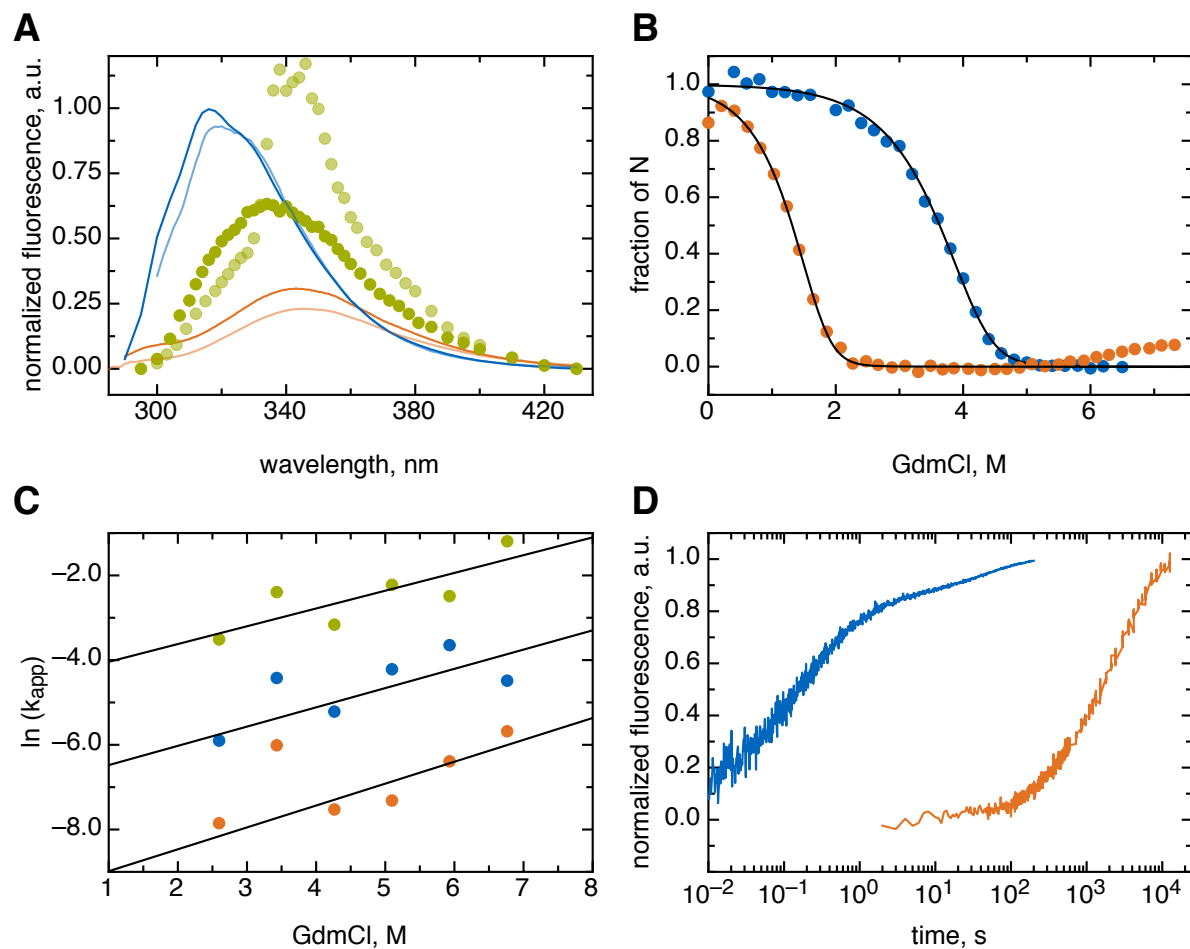


Figure 4.27.: A) Fluorescence spectra of foldon wild-type and foldon P4T P7Nva. The native state spectra are colored in blue, the spectra of the unfolded state in orange and the burst-phase intermediate spectra in green. The fluorescence spectra of foldon P4T P7Nva are lighter colored than the corresponding foldon wild-type spectra. B) Comparison of the foldon wild-type (blue) and foldon P4T P7Nva (orange) GdmCl transitions at 30 μM monomer concentration. C) Linear plot of the common logarithms of the three unfolding rate constants of the foldon P4T P7Nva GdmCl unfolding kinetics against the corresponding GdmCl concentrations. D) Comparison of the foldon wild-type (blue) and foldon P4T P7Nva refolding traces in 0.58 M GdmCl 20 mM sodium cacodylic acid pH 7.0. The monomer concentration is 10 μM .

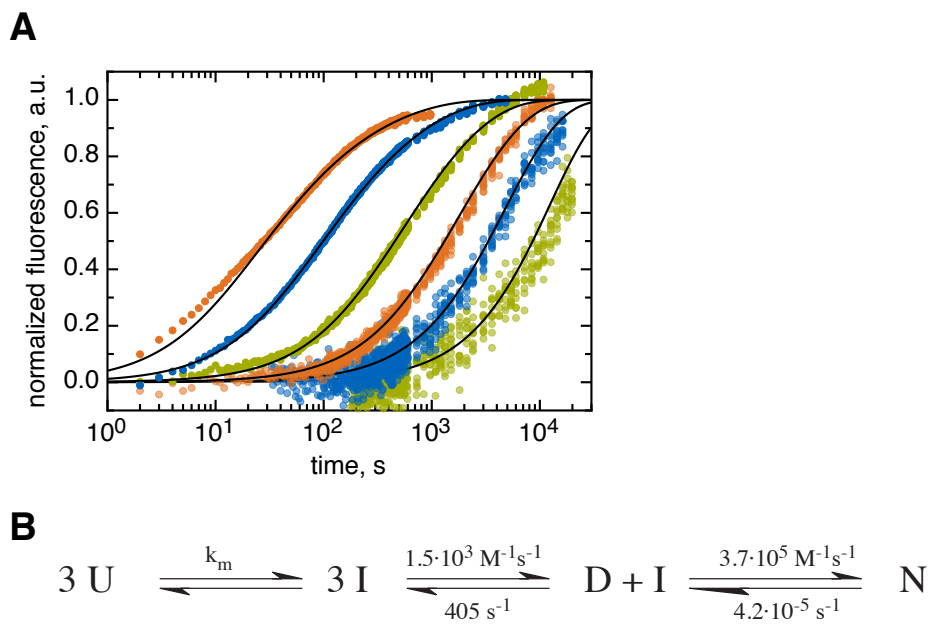


Figure 4.28.: Normalized global fit traces of foldon P4T P7Nva GdmCl refolding kinetics. A) Normalized GdmCl refolding traces after fitting. B) Fitting model with all determined rate constants with all parameters run freely.

4.10. Foldon Y13F

Tyrosine 13 is the only amino acid within the hydrophobic cluster that has a polar side chain. The hydroxyl group of the side chain is part of a hydrogen bond network in the first type-II-turn structure formed by the residues 9 to 11. The hydrogen bonds are formed between the side chain oxygens of glutamate 9, the side chain nitrogen of glutamine 11 and the hydroxyl group of the tyrosine 13 side chain, while another hydrogen bond in this network is formed between one oxygen of the glutamate 9 side chain and the backbone nitrogen of glutamine 11. Thus tyrosine acts as a link between the hydrophobic cluster and the first type-II-turn. Removal of the hydroxyl group by mutating the tyrosine to a phenylalanine at position 13 disrupts the connection to the turn, which is still able to form hydrogen bonds within the turn structure as seen from the X-ray structure, but lacks connectivity to other regions of the foldon monomer. Next to formation of the discussed intra-molecular hydrogen bond, the hydroxyl group of tyrosine 13 makes up a minor region of the foldon surface, interacting with the surrounding solvent. Removal of the hydroxyl group results in direct contact of the hydrophobic side chain of the phenylalanine to the solvent, which is entropically unfavorable. An almost perfect alignment of the phenyl rings of residue 13 in foldon wild-type and foldon Y13F is observed.

Foldon Y13F shows a strong resemblance to foldon A6V in its fluorescence spectra. The native state fluorescence is decrease to around 40 % of the foldon wild-type intensity while still showing the characteristic shape of the native state. The decrease can be due to a lower stability of this variant, but tryptophan fluorescence can also be decreased due to the missing FRET from the tyrosine in close proximity within the hydrophobic core. In the unfolded state spectrum of foldon Y13F, where the intensity is decreased by approximately 25 %, the tyrosine shoulder with its maximum at 303 nm is not detectable any more, despite the fact that the other tyrosine at position 3 is still present. The intermediate spectrum has a red-shifted maximum at 346 nm and shows a slight peak, as observed for foldon A6S and P4T P7Nva. The intensity of the burst-phase intermediate is reduced by almost 50 %.

The substitution of a tyrosine with a phenylalanine is probably the most conservative substitution in the whole study. Yet this exchange leads to a destabilization by $24.8 \text{ kJ} \cdot \text{mol}^{-1}$ compared to foldon wild-type, resulting in a $\Delta G^0(\text{H}_2\text{O}) = (-64.4 \pm 0.1) \text{ kJ} \cdot \text{mol}^{-1}$ with an $m_{eq} = (13.8 \pm 0.1) \text{ kJ} \cdot \text{mol}^{-1} \cdot \text{M}^{-1}$ in GdmCl. This result shows that not only the insertion, but also the removal of a hydroxyl group within the hydrophobic core can lead to an increased

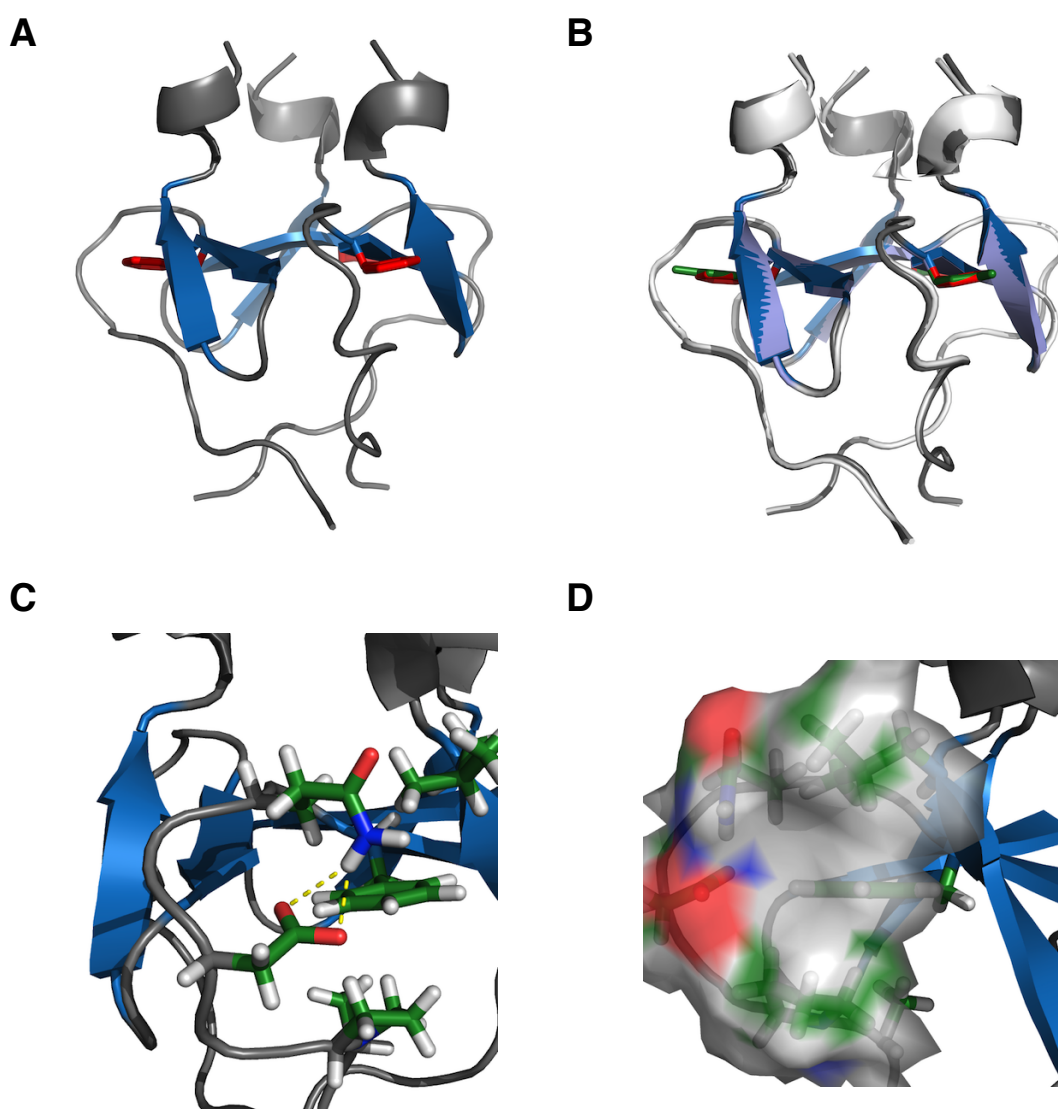


Figure 4.29.: Crystal structure of foldon Y13F. A) Crystal structure of foldon Y13F in ribbon representation, with the β -strands highlighted in blue, while the 3_{10} -helix and the other regions of the structure are colored in grey. The side chain of the mutated residue 13 is shown in stick representation and the side chain colored in red. B) Backbone alignment of foldon Y13F and foldon wild-type. Foldon wild-type, also shown in ribbon representation, is colored in white, with the exception of the β -strands, which are colored in light blue. The tyrosine 13 residue is shown in stick representation and colored in green. The foldon Y13F crystal structure is shown and colored as in A). C) The hydrogen bond network in the first β -turn of the foldon Y13F mutant. The hydrogen bond from glutamate 9 to tyrosine 13 present in foldon wild-type is replaced by a bond to the glutamine 12 nitrogen atom in the side chain. D) Transparent surface representation of the first β -turn region, showing that the former hydrophilic region of the Tyr13 hydroxyl group is now a hydrophobic patch where the solvent has direct access to the hydrophobic core of the native trimer.

4. Results

susceptibility of the native trimer to the denaturant. The effect on stability is within the range of the other mutations. The Y13F foldon variant shows the highest susceptibility to urea of all studied variants with a $m_{eq} = (6.9 \pm 0.4) \text{ kJ} \cdot \text{mol}^{-1} \cdot \text{M}^{-1}$, accompanied by the largest difference between the determined stabilities in GdmCl and urea of $6.2 \text{ kJ} \cdot \text{mol}^{-1}$, with $\Delta G^0(H_2O) = (-58.2 \pm 1.3) \text{ kJ} \cdot \text{mol}^{-1}$. This low stability could lead to ill-defined native baselines and very high errors of the determined results if compared to the other variants and foldon wild-type, as shown in Figure A.9. A global fit using all transition curves of both denaturants results in a free folding enthalpy of $\Delta G^0(H_2O) = (-59.1 \pm 0.4) \text{ kJ} \cdot \text{mol}^{-1}$, the m_{eq} -values for the respective denaturants are $m_{eq}(\text{GdmCl}) = (10.9 \pm 0.2) \text{ kJ} \cdot \text{mol}^{-1} \cdot \text{M}^{-1}$ and $m_{eq}(\text{urea}) = (7.1 \pm 0.2) \text{ kJ} \cdot \text{mol}^{-1} \cdot \text{M}^{-1}$.

Foldon Y13F does neither show complex unfolding kinetics nor a kink or roll-over in the plot of the common logarithms of the apparent rate constants against the GdmCl concentrations. Fitting is achieved by a single linear equation, resulting in a k_u of $5.4 \cdot 10^{-4} \text{ s}^{-1}$ and a m_u -value of $4.6 \text{ kJ} \cdot \text{mol}^{-1} \cdot \text{M}^{-1}$. The unfolding traces are fitted with a single-exponential equation. The same holds true for urea unfolding, resulting $k_u = (1.39 \pm 0.01) \cdot 10^{-4} \text{ s}^{-1}$ and a $m_u = (1.66 \pm 0.01) \cdot 10^3 \text{ kJ} \cdot \text{mol}^{-1} \cdot \text{M}^{-1}$. The apparent rate constants are obtained by single-exponential fitting of the unfolding traces. The m_u -value is in the same range as the one of foldon P7Nva, but significantly higher than for the foldon variants at position 6. It is thus not clear if the obtained kinetic m_u -values can be directly compared or if they represent local unfolding events in one case and in the other case a global unfolding reaction. A comparison with GdmCl unfolding reactions to determine which of the two cases is true is however not possible, as unfolding with GdmCl always results in global unfolding, regardless of the used GdmCl concentration.

The reaction order determined from the half-life times in both GdmCl in urea refolding experiments is close to two, as is the reaction order determined from the initial slopes for both denaturants. The dimer association rate constant is $k_D = (1.79 \pm 0.12) \cdot 10^4 \text{ M}^{-1} \cdot \text{s}^{-1}$ in GdmCl and $k_D = (2.57 \pm 0.22) \cdot 10^3 \text{ M}^{-1} \cdot \text{s}^{-1}$ in urea.

The modified model for foldon association which allows productive dimer and trimer formation of monomers with a *cis* conformation of the Ala6-Pro7 peptide bond is used for global fitting of the foldon Y13F refolding traces in GdmCl. The refolding traces are adjusted using a free enthalpy of folding of $59.1 \text{ kJ} \cdot \text{mol}^{-1}$. The fit traces are in good agreement with the

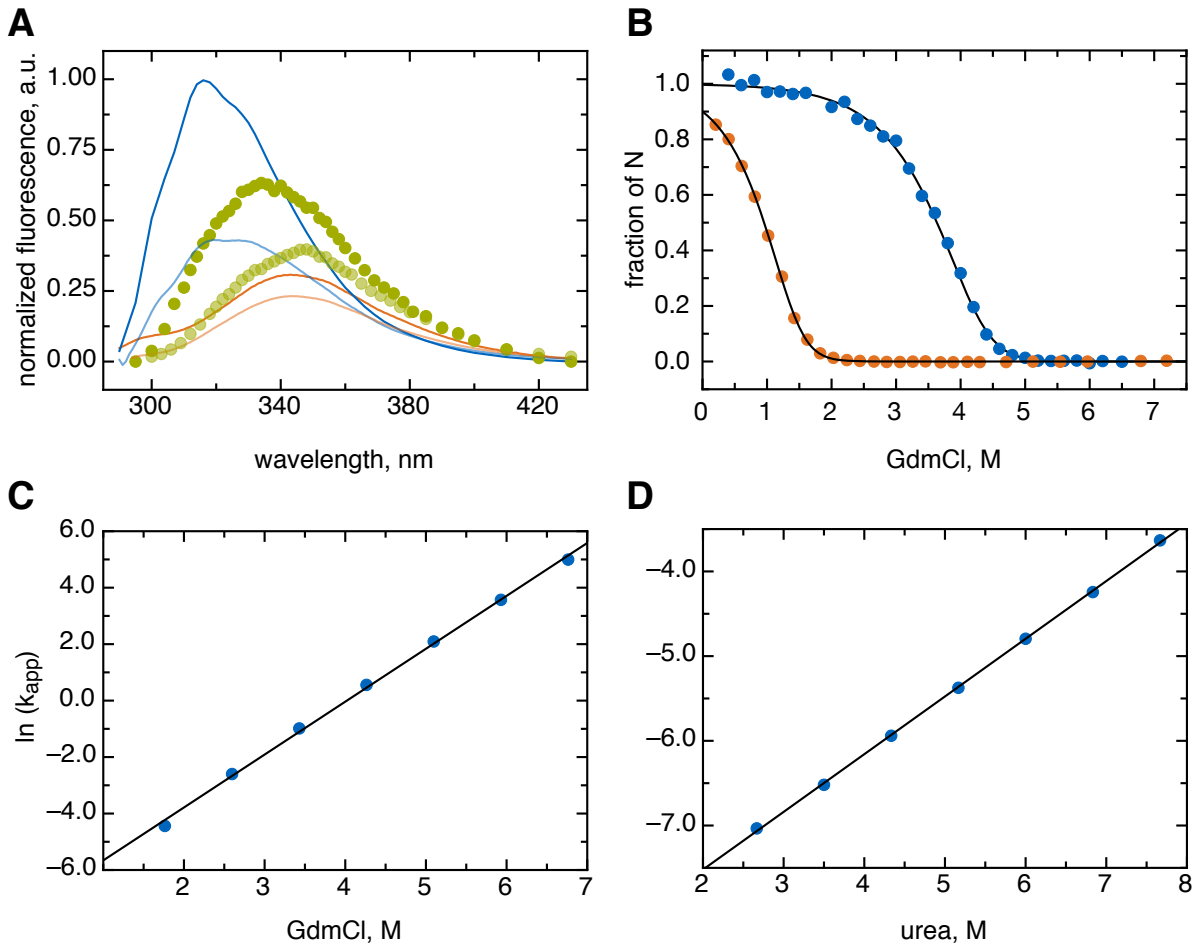


Figure 4.30.: A) Fluorescence spectra of foldon wild-type and foldon Y13F. The native state spectra are colored in blue, the spectra of the unfolded state in orange and the burst-phase intermediate spectra in green. The fluorescence spectra of foldon Y13F are lighter colored than the corresponding foldon wild-type spectra. B) Comparison of the foldon wild-type (blue) and foldon Y13F (orange) GdmCl transitions at 30 μ M monomer concentration. C) Linear plot of the common logarithms of the three unfolding rate constants of the foldon Y13F GdmCl unfolding kinetics against the corresponding GdmCl concentrations. D) Chevron plots of the common logarithms of the major phase unfolding rate constants of the foldon Y13F urea unfolding kinetics against varying urea concentrations.

4. Results

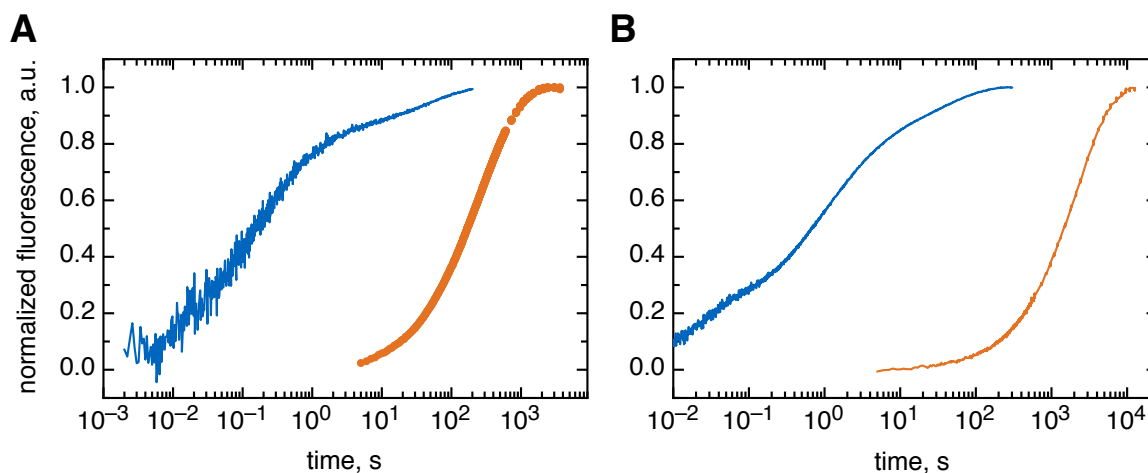


Figure 4.31.: Comparison of foldon wild-type (blue) and foldon Y13F (orange) refolding traces A) 10 μM foldon monomer refolding traces in 0.58 M GdmCl. B) 10 μM foldon monomer refolding traces in 1.0 M urea.

refolding traces, with only a slight deviation for the 2 and 100 μM foldon monomer concentration traces, as seen in Figure 4.32. The most tremendous effect is observed for the dissociation rate constant of the dimeric intermediate, k_{-D} . The rate constant of 731 s^{-1} is however lowered to 257 s^{-1} after free fitting of all rate constants. The caused changes in the dimerization rate constant are too small to be detectable after rounding of the results. However, they have a tremendous effect on both k_{-D} and k_T . The association rate constants are decreased by one order of magnitude for the trimerization and two orders of magnitude for the dimerization rate constant. The highest impact of the amino acid exchange is found on the dimerization, as the monomer-to-dimer equilibrium is shifted by a factor of 500 towards the monomeric state compared to foldon wild-type.

Deviations from the global fit traces in urea are seen for 20 and 50 μM foldon monomer concentration. For the 50 μM foldon monomer refolding experiment, this is rather an artifact of the measurement for the longer time-points than poor fitting. The systematic deviations for lower foldon monomer concentrations are more serious in this regard, although they could be explained by the error in the concentration determination. The dimerization rate is comparable to the ones determined for the foldon variants at position 6 and position 7. The effect on the trimerization reaction is even stronger, although the rate constant cannot be determined with high accuracy, for the reasons discussed before.

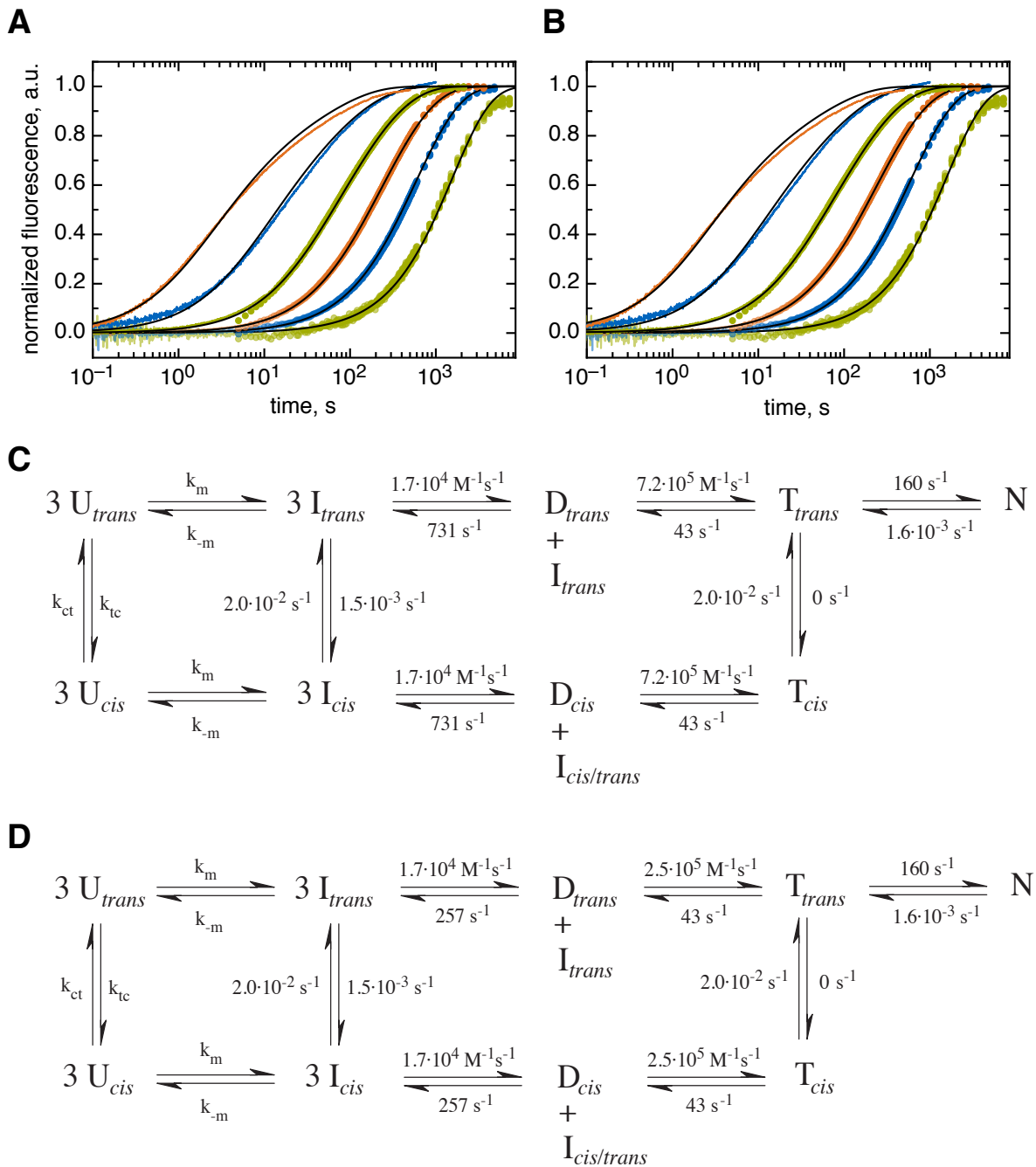


Figure 4.32.: Normalized global fit traces of foldon Y13F GdmCl refolding kinetics. A) Normalized GdmCl refolding traces after fitting with fixed dimerisation rate constant k_D . B) Normalized GdmCl refolding traces after fitting with all parameters run freely. C) Fitting model with all determined rate constants with fixed dimerisation rate constant k_D . D) Fitting model with all determined rate constants with all parameters run freely.

4. Results

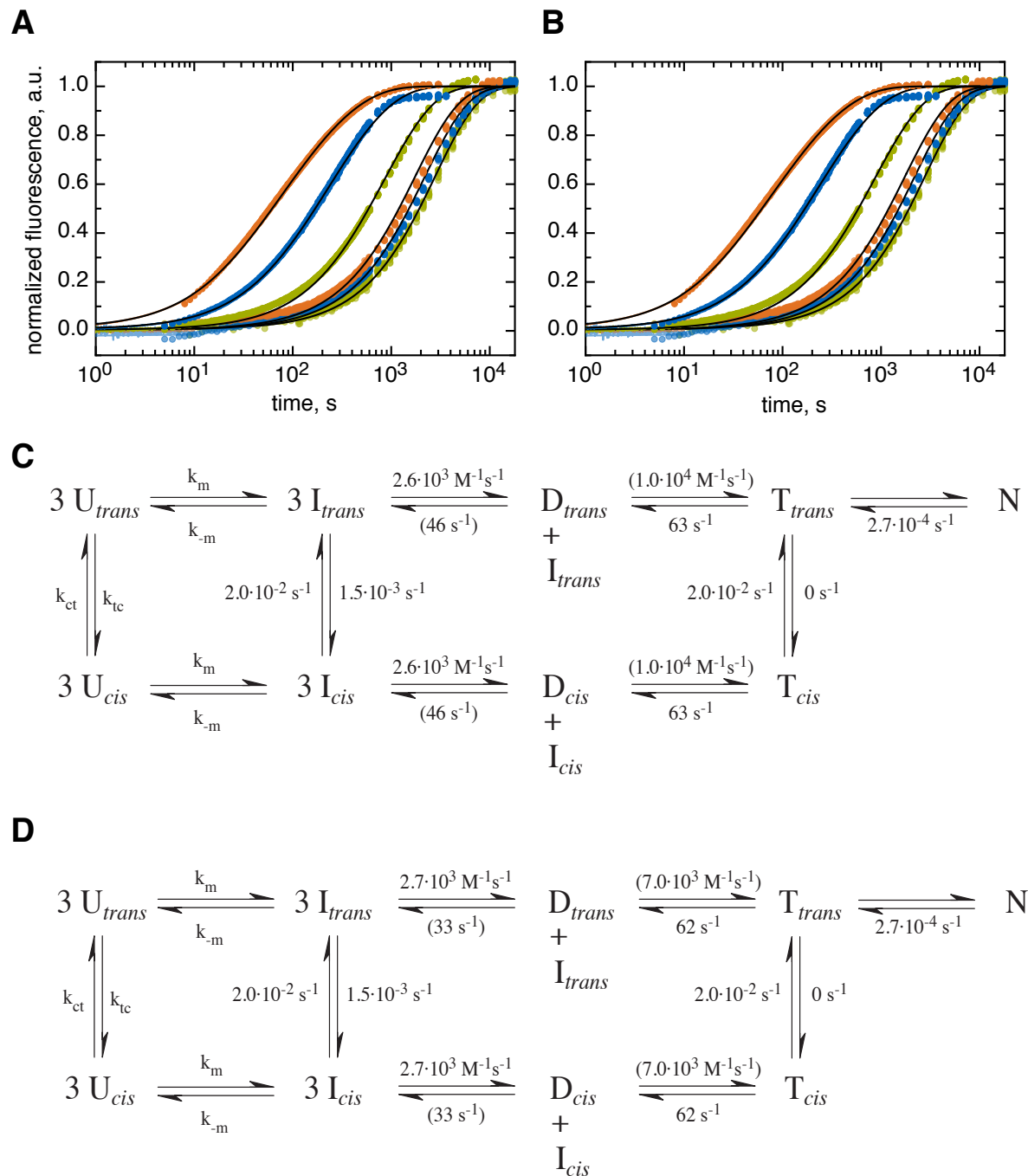


Figure 4.33.: Normalized global fit traces of foldon Y13F urea refolding kinetics. A) Normalized urea refolding traces after fitting with fixed dimerisation rate constant k_D . B) Normalized urea refolding traces after fitting with all parameters run freely. C) Fitting model with all determined rate constants with fixed dimerisation rate constant k_D . D) Fitting model with all determined rate constants with all parameters run freely.

Due to the large dissociation rate constant k_{-D} , foldon Y13F has the lowest k_T/k_{-D} ratio of all studied foldon variants in both denaturants. However, as the rate constant for trimerization k_T is similar in size to other foldon variants, this effect is solely caused by the dissociation of the dimer. The low stability of the dimer compared to foldon wild-type is very likely caused by the loss of the direct interaction of the hydrophobic cluster and the hydrogen bonding network of the first turn-II-structure.

4.11. Foldon W20Nal

Foldon W20Nal is considered the least destabilizing of all mutants inserted into foldon. An overlay of the foldon W20Nal crystal structure with the foldon wild-type structure shows that the naphthyl- and indole-ring align almost perfectly with only slight deviations for naphthylalanine due to the changed binding angle, as the indole is bound via a five-atom heterocyclic ring to the backbone. But otherwise, no steric hindrance or clashes occur, and all hydrophobic interactions are still present. The only deviation seen for the naphthylalanine variant is the expected loss of the hydrogen bond from the indole nitrogen atom N' to the oxygen of the residue 4 backbone. All other interactions of the N-terminus with the rest of the molecule are prevailed, so the missing bond does not cause any observable changes of the structure.

In contrast to all previously discussed foldon variants, the mutations at position 20 lead to a tremendous change in the fluorescence spectra, as the tryptophan is removed from foldon. The spectra of the foldon variants W20Nal and W20H are therefore not directly comparable to the respective foldon wild-type spectra. Additionally, the measurement of the burst-phase intermediate is not possible, as refolding experiments of the tryptophan 20 variants cannot be carried out due to their extremely low stability and slow refolding reaction, as discussed below. Thus only the native and unfolded state spectra are compared. The spectra of foldon W20Nal are dominated by the fluorescence of the naphthylalanine. The fluorescence of naphthylalanine has a wide excitation range, which was probed to determine a suitable excitation wavelength. In the end, 273 nm was chosen as excitation wavelength as tyrosine fluorescence is also strongly excited at this wavelength. The fluorescence intensity is increased upon unfolding. Additionally, upon unfolding, a second minor peak is observed with a maximum at 303 nm. The existence of this peak indicates that FRET from the tyrosine residues to the naphthylalanine occurs in the folded state. The FRET efficiency is obviously strongly reduced in

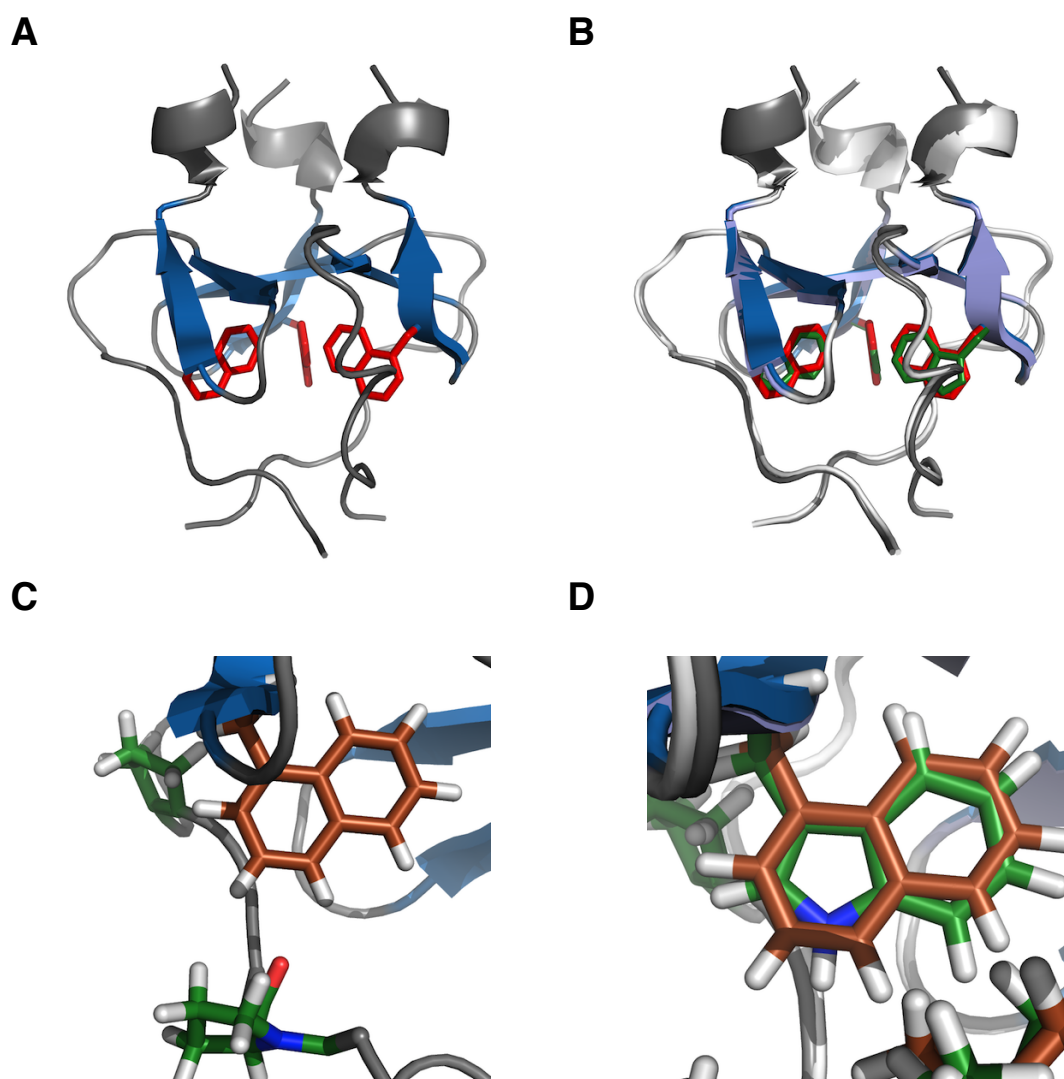


Figure 4.34.: Crystal structure of foldon W20Nal. A) Crystal structure of foldon W20Nal in ribbon representation, with the β -strands highlighted in blue, while the 3_{10} -helix and the other regions of the structure are colored in grey. The side chain of the mutated residue 20 is shown in stick representation and the side chain colored in red. B) Backbone alignment of foldon W20Nal and foldon wild-type. Foldon wild-type, also shown in ribbon representation, is colored in white, with the exception of the β -strands, which are colored in light blue. The tryptophan 20 residue is shown in stick representation and colored in green. The foldon W20Nal crystal structure is shown and colored as in A). C) Stick representation of the residues Nal20 and Pro4. The carbon atoms are colored in green (with the exception of Nal20, where the carbon atoms are shown in brown), the oxygen atoms in red, nitrogens in blue and the hydrogen atoms in white. D) Overlay of the residue 20 side chains of both foldon wild-type and foldon W20Nal, together with the neighboring residue Pro4. The proline at position 4 in shown in an all-stick representation. The color scheme is the same as in C).

the unfolded state, making the tyrosine fluorescence directly observable. That way the emission change at 303 nm after excitation at 273 nm can be used for further experiments. The unfolded and native state spectra of foldon W20Nal variants are shown in Figure 4.35A.

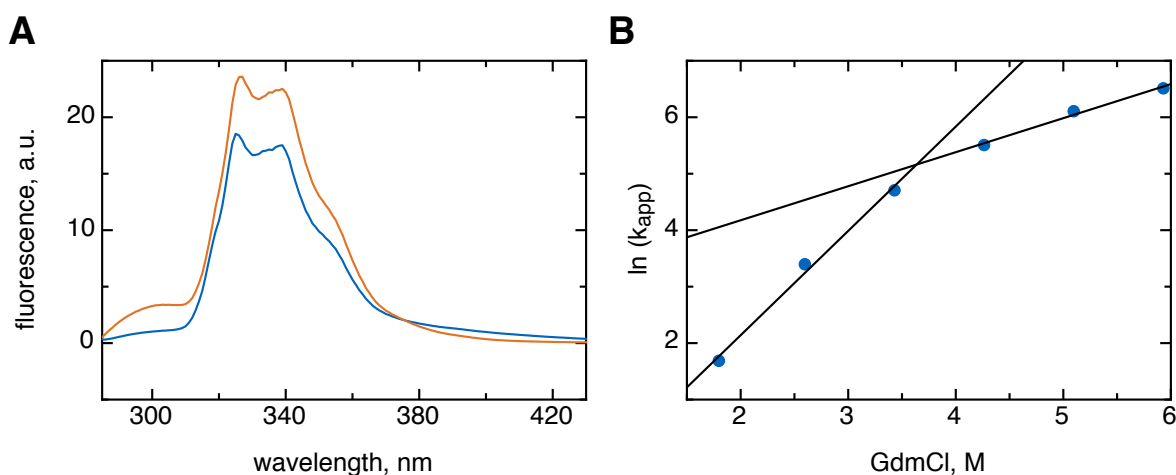


Figure 4.35.: A) Fluorescence spectra foldon W20Nal. The native state spectrum is colored in blue, the unfolded state spectrum in orange. B) Linear plot of the common logarithm of the foldon W20Nal GdmCl unfolding apparent rate constants against the corresponding GdmCl concentrations.

It was not possible to determine the stability of foldon W20Nal by either a GdmCl or urea transition due to its low free folding enthalpy. The denaturant transition of this variant was thus carried out in the presence of 500 mM sodium sulfate, as was a transition of foldon wild-type to allow a comparison between the mutant and foldon wild-type. The free folding enthalpy of foldon wild-type is $\Delta G^0(\text{H}_2\text{O}) = (-108.3 \pm 1.5) \text{ kJ} \cdot \text{mol}^{-1}$, the m_{eq} -value is $m_{eq} = (13.3 \pm 0.4) \text{ kJ} \cdot \text{mol}^{-1} \cdot \text{M}^{-1}$. The normalized transition curves are shown in Figure 4.38A. The difference of the free folding enthalpy of foldon wild-type to the transition carried out without sodium sulfate is $19.1 \text{ kJ} \cdot \text{mol}^{-1}$, corresponding to a shift towards the native state by a factor of 40,000. The increase in stability is accompanied by an increase of the m_{eq} -value by $3.1 \text{ kJ} \cdot \text{mol}^{-1} \cdot \text{M}^{-1}$. A transition of foldon wild-type in urea is not possible as urea is not able to unfold the highly stabilized foldon wild-type within the solubility range of urea under these conditions. The free folding enthalpy of foldon W20Nal cannot be determined precisely due to the low stability of foldon W20Nal, despite the presence of 500 mM sodium sulfate. Several global fits together with foldon W20H and foldon wild-type are performed, which are discussed in detail in Chapter 4.12. In general, one can determine the stability of foldon W20Nal to be increased by about $40 \text{ kJ} \cdot \text{mol}^{-1}$ compared

4. Results

to foldon wild-type in both denaturants. Unfolding experiments of foldon W20Nal against varying GdmCl concentrations reveals a kink in the plot of the common logarithm of the apparent rate constant determined by single-exponential fitting against the effective GdmCl concentration. While the slow phase has an already remarkable fast dissociation constant with $k_{u,slow} = (0.21 \pm 0.06) \text{ s}^{-1}$ and an $m_{u,slow} = (4.41 \pm 0.40) \text{ kJ} \cdot \text{mol}^{-1} \cdot \text{M}^{-1}$, the fast phase has a rate constant of $k_{u,fast} = (19.49 \pm 2.26) \text{ s}^{-1}$ with $m_{u,fast} = (1.47 \pm 0.16) \text{ kJ} \cdot \text{mol}^{-1} \cdot \text{M}^{-1}$. These are the fastest determined dissociation rate constants throughout this study and a good explanation for the low stability of foldon W20Nal.

Despite the fact that equilibrium transitions in both GdmCl and urea and unfolding experiments in GdmCl can be performed and evaluated using foldon W20Nal, it was not possible to determine refolding experiments using this foldon variant. All refolding traces determined were strongly governed by mixing artifacts and photo-bleaching, if any fluorescence changes were determinable at all. This indicates that refolding is decreased to an extent that determination of the refolding reaction is not possible any more. Yet this foldon variant is still able to fold into the native structure, given enough time. The extremely low stability of foldon W20Nal is thus a consequence of both a fast unfolding and an extremely slow refolding reaction, strongly shifting the equilibrium towards the unfolded state.

4.12. Foldon W20H

The exchange of tryptophan to histidine serves the purpose to control if formation of the hydrogen bond disrupted by the exchange of tryptophan to naphthylalanine in the W20Nal mutant is restored. And although the hydrogen bond is formed, as can be seen clearly from the crystal structure, the loss of hydrophobic interactions in the hydrophobic core and the likely trapping of water within the native trimer overcompensate for the gained free folding enthalpy by the formed hydrogen bonds. The water molecules within the hydrophobic core are not defined in the crystal structure. This indicates that any solvent molecules trapped within the hydrophobic core do not show specific binding to any groups forming the shell of the cavern.

In foldon W20H, the fluorescence emission spectra of both states is dominated by the tyrosine fluorescence with its characteristic maximum at 303 nm after excitation at 273 nm. An increase of the fluorescence intensity is observed upon unfolding in high GdmCl concentrations

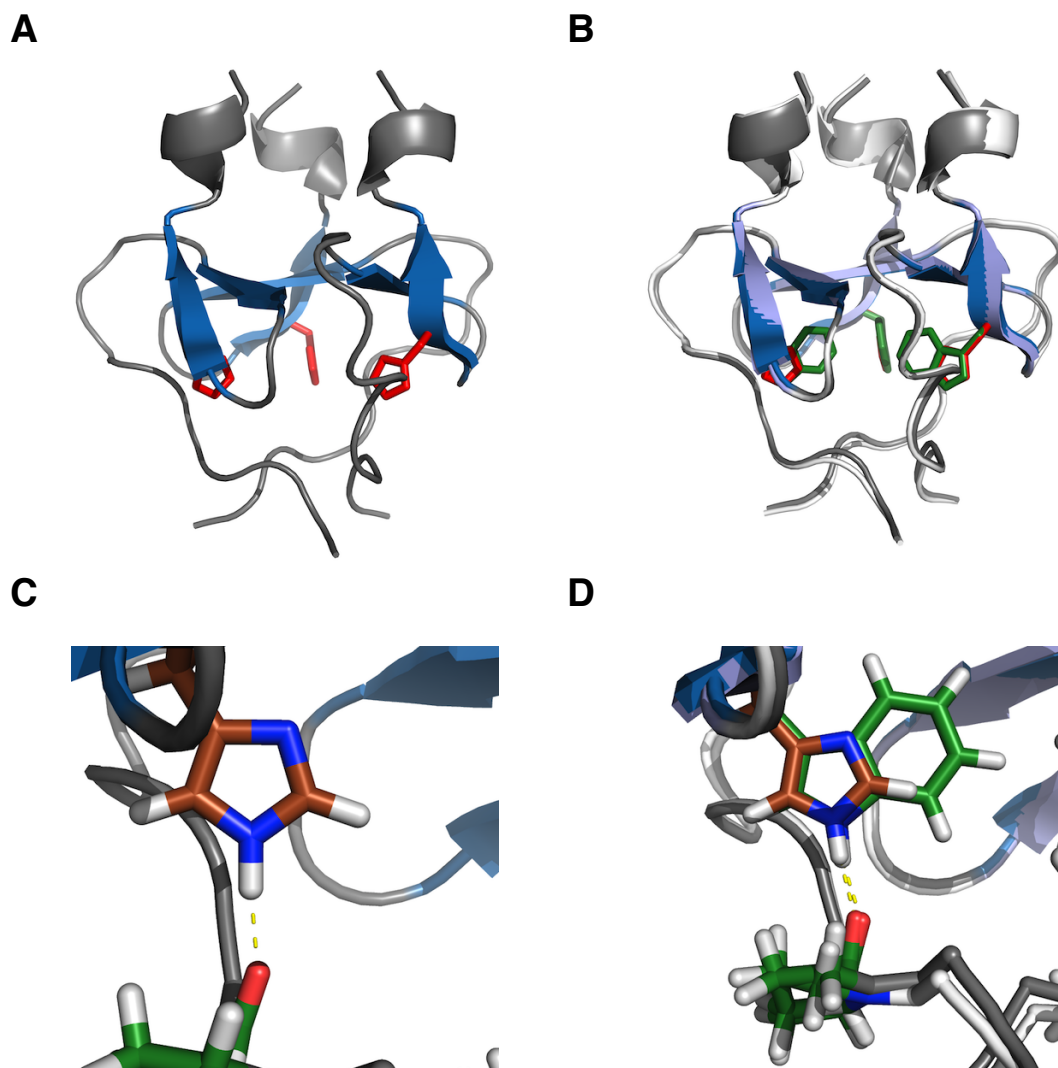


Figure 4.36.: Crystal structure of foldon W20H. A) Crystal structure of foldon W20H in ribbon representation, with the β -strands highlighted in blue, while the 3_{10} -helix and the other regions of the structure are colored in grey. The side chain of the mutated residue 20 is shown in stick representation and the side chain colored in red. B) Backbone alignment of foldon W20H and foldon wild-type. Foldon wild-type, also shown in ribbon representation, is colored in white, with the exception of the β -strands, which are colored in light blue. The tryptophan 20 residue is shown in stick representation and colored in green. The foldon W20H crystal structure is shown and colored as in A). C) Stick representation of the residues His20 and Pro4 in an all-stick representation. The carbon atoms are colored in green (with the exception of His20, where the carbon atoms are shown in brown), the oxygen atoms in red, nitrogens in blue and the hydrogen atoms in white. The hydrogen bond between the backbone oxygen of residue 4 and the N_ϵ of the histidine 20 side chain is depicted as a dashed yellow line. D) Overlay of the residue 20 side chains of both foldon wild-type and foldon W20H, together with the neighboring residue Pro4 in an all-stick representation. The color scheme is the same as in C). The overlay shows almost perfect alignment of the side chains.

4. Results

without any significant changes in the shape of the spectrum. This fluorescence increase upon unfolding can be used to probe the transition between the unfolded and native state in other experiments (Figure 4.37A).

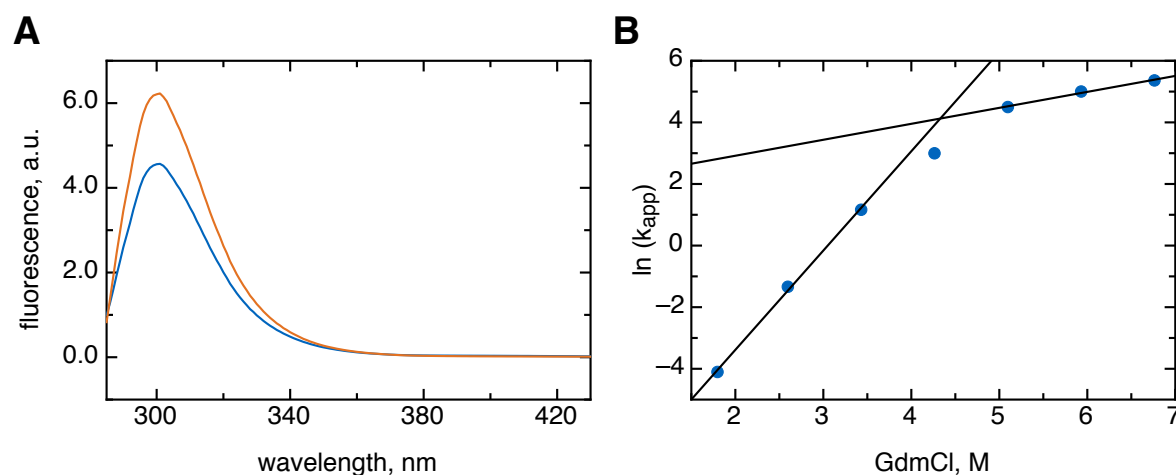


Figure 4.37.: A) Fluorescence spectra of foldon W20H. The native state spectrum is colored in blue, the unfolded state spectrum in orange. B) Linear plot of the common logarithms of the apparent rate constants of foldon W20H GdmCl unfolding kinetics against the corresponding GdmCl concentrations.

Despite the stabilization by sulfate and high foldon monomer concentrations, fitting of the foldon W20H transitions leads to artifacts: Although the transition mid-points are clearly at lower GdmCl concentrations compared to foldon W20Nal, the fits yield a lower free folding enthalpy, accompanied with an increased m_{eq} -value of almost $20 \text{ kJ} \cdot \text{mol}^{-1} \cdot \text{M}^{-1}$. Several strategies are applied to overcome this problem: Using a fixed m_{eq} -value determined from fitting of the higher foldon monomer concentration transition curve is one possibility. In addition, a global fit of both the W20Nal and W20H transition curves is performed, using a global m_{eq} -value for both foldon variants, as the effect of both mutations on the susceptibility to GdmCl should be rather similar. For this global fit, the traces of foldon wild-type are used as well, as the clear native and unfolded baselines of the foldon wild-type transition result in a well-defined overall m_{eq} -value. The normalized transition curves are shown in Figure 4.38, the stabilities finally determined are $\Delta G^0(\text{H}_2\text{O}) = (-65.5 \pm 0.4) \text{ kJ} \cdot \text{mol}^{-1}$ and $\Delta G^0(\text{H}_2\text{O}) = (-55.0 \pm 0.7) \text{ kJ} \cdot \text{mol}^{-1}$ for foldon W20Nal and W20H, respectively, with a m_{eq} -value of $m_{eq} = (12.9 \pm 0.2) \text{ kJ} \cdot \text{mol}^{-1} \cdot \text{M}^{-1}$ for both variants. The free folding enthalpy determined for foldon wild-type by this method is $\Delta G^0(\text{H}_2\text{O}) = (-107.3 \pm 0.9) \text{ kJ} \cdot \text{mol}^{-1}$, which is within the error range of the value determined by fitting of the foldon wild-type data alone.

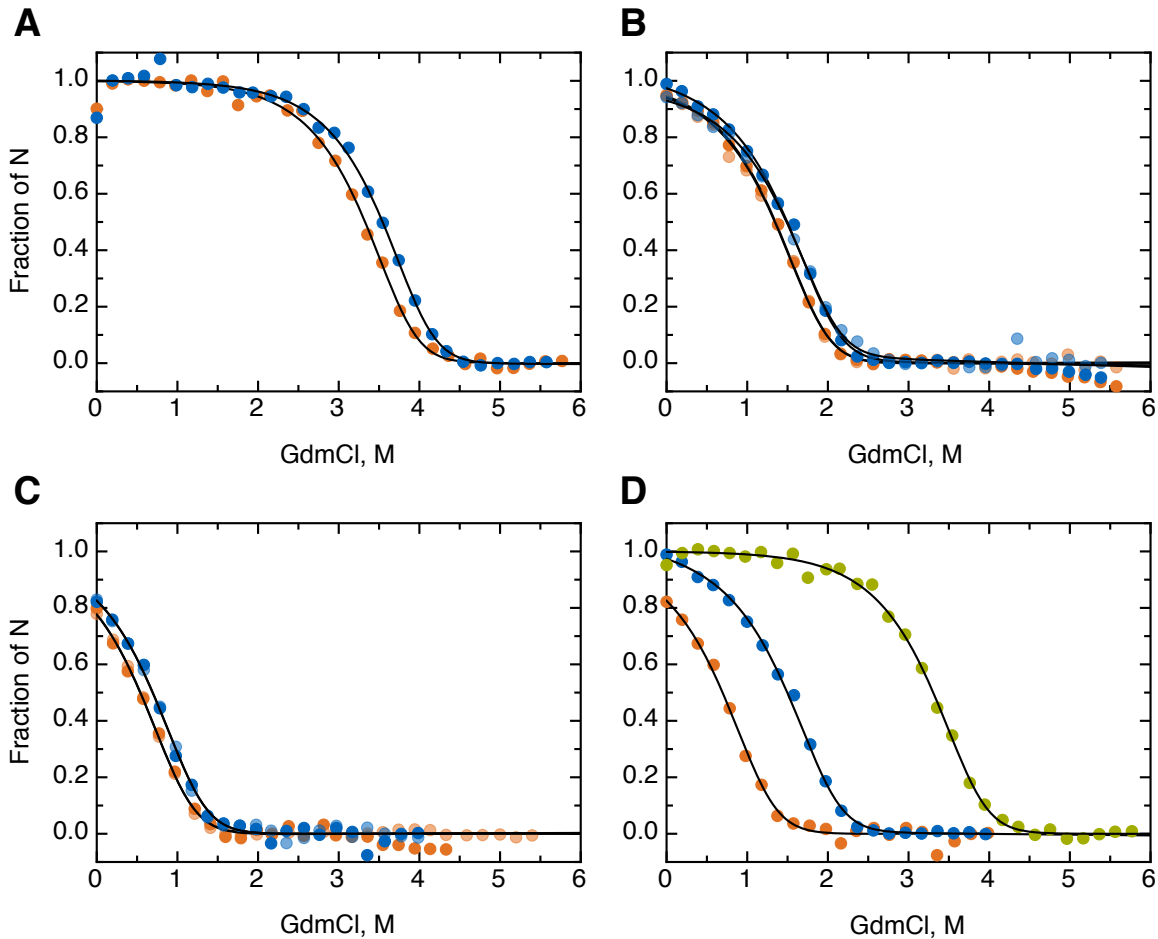


Figure 4.38.: Normalized GdmCl transition curves of foldon wild-type and both tryptophan 20 substitutions in the presence of 500 mM sodium sulfate. A) Normalized GdmCl transition curves in the presence of 500 mM sodium sulfate for 2 μ M (orange) and 4 μ M (blue) monomer concentrations of foldon wild-type, measured by intrinsic tryptophan fluorescence change. B) Normalized GdmCl transition curves in the presence of 500 mM sodium sulfate for 60 μ M (orange) and 90 μ M (blue) monomer concentrations of foldon W20Nal, measured by intrinsic tyrosine fluorescence change (opaque points) or change in circular dichroism absorption at 228 nm (translucent points). C) Normalized GdmCl transition curves in the presence of 500 mM sodium sulfate for 60 μ M (orange) and 90 μ M (blue) monomer concentrations of foldon W20H, measured by intrinsic tyrosine fluorescence change (opaque points) or change in circular dichroism absorption at 228 nm (translucent points). D) Comparison of a 2 μ M foldon wild-type monomer concentration (green) with 90 μ M foldon W20H (orange) and 90 μ M foldon W20Nal (blue) transition curves, measured by change in intrinsic tryptophan or tyrosine fluorescence, respectively.

The same problems encountered during the evaluation of the GdmCl data are also present during the evaluation of the urea transition data: Fitting of the foldon W20H transition curves

4. Results

alone results in a higher stability than found for foldon W20Nal, again accompanied by a high m_{eq} -value. Thus, like for the GdmCl transitions, a global fit with a global m_{eq} -value for the urea data is carried out, resulting $\Delta G^0(\text{H}_2\text{O}) = (-62.7 \pm 0.8) \text{ kJ} \cdot \text{mol}^{-1}$ and $\Delta G^0(\text{H}_2\text{O}) = (-49.8 \pm 0.8) \text{ kJ} \cdot \text{mol}^{-1}$ for foldon W20Nal and W20H, respectively. The m_{eq} -value is $m_{eq} = (5.8 \pm 0.2) \text{ kJ} \cdot \text{mol}^{-1} \cdot \text{M}^{-1}$. The transition curves in urea and a comparison of both variants are given in Figure 4.39.

Comparing the transitions in both GdmCl and urea reveals that the free enthalpy of folding determined in GdmCl is lower than in urea. Thus, an additional global fit is carried out. The m_{eq} -values for foldon W20H for this fit are fixed to the values determined by the previously performed global fits. All different free folding enthalpies $\Delta G^0(\text{H}_2\text{O})$ are given in Table 4.2. For both variants at position 20, the determined stabilities have to be seen rather as an accumulation point than an exact result, although all determined values are within error range of each other.

W20Nal	GdmCl		urea	
	$\Delta G^0(\text{H}_2\text{O})$ kJ · mol ⁻¹	m_{eq} kJ · mol ⁻¹ · M ⁻¹	$\Delta G^0(\text{H}_2\text{O})$ kJ · mol ⁻¹	m_{eq} kJ · mol ⁻¹ · M ⁻¹
normal fitting	-69.9 ± 1.5	15.1 ± 0.7	-61.5 ± 0.4	5.4 ± 0.1
global fit with W20H using a global m_{eq}	-65.5 ± 0.4	12.9 ± 0.2	-62.7 ± 0.8	5.8 ± 0.2
global fit of GdmCl & urea	-64.1 ± 0.5	12.2 ± 0.3		6.1 ± 0.1
W20H	GdmCl		urea	
	$\Delta G^0(\text{H}_2\text{O})$ kJ · mol ⁻¹	m_{eq} kJ · mol ⁻¹ · M ⁻¹	$\Delta G^0(\text{H}_2\text{O})$ kJ · mol ⁻¹	m_{eq} kJ · mol ⁻¹ · M ⁻¹
normal fitting	-64.6 ± 5.2	19.5 ± 3.7	-69.1 ± 7.2	13.8 ± 3.2
global fit with W20Nal using a global m_{eq}	-55.0 ± 0.7	12.9 ± 0.2	-49.8 ± 0.8	5.8 ± 0.2
global fit of GdmCl & urea	-53.4 ± 0.5	12.9		5.8

Table 4.2.: All free folding enthalpies and m_{eq} -values of both foldon W20Nal and W20H determined by different global fitting attempts described in detail in the text. Results with the same colour have been obtained from the same global fit, except for results colored in black, which are the results of individual fits of the corresponding denaturant transitions.

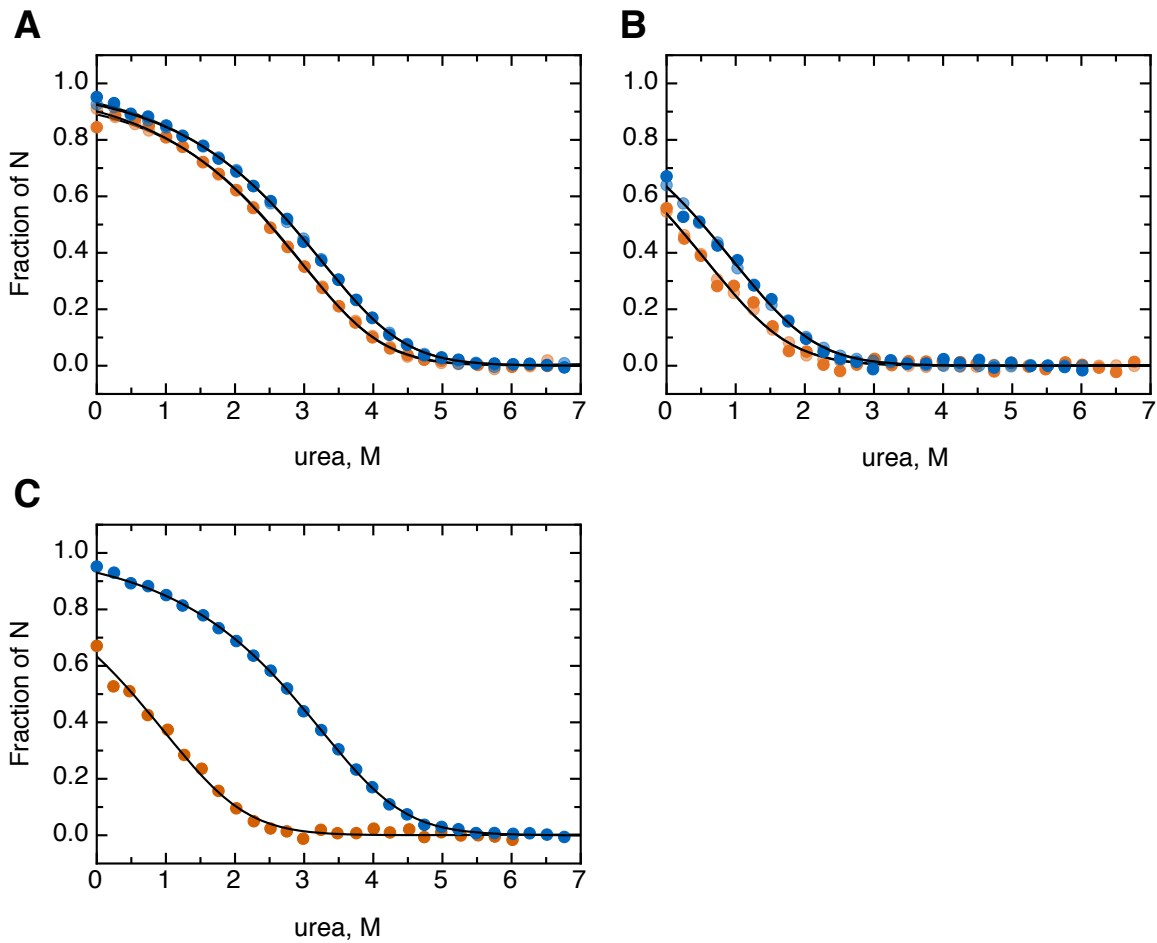


Figure 4.39.: Normalized urea transition curves of both tryptophan 20 substitutions in the presence of 500 mM sodium sulfate. A) Normalized urea transition curves in the presence of 500 mM sodium sulfate for 60 μ M (orange) and 90 μ M (blue) monomer concentrations of foldon W20Nal, measured by intrinsic tyrosine fluorescence change (opaque points) or change in circular dichroism absorption at 228 nm (translucent points). B) Normalized urea transition curves in the presence of 500 mM sodium sulfate for 60 μ M (orange) and 90 μ M (blue) monomer concentrations of foldon W20H, measured by intrinsic tyrosine fluorescence change (opaque points) or change in circular dichroism absorption at 228 nm (translucent points). C) Direct comparison both foldon tryptophan mutant transitions at 90 μ M foldon monomer concentration. Foldon W20H (orange) and foldon W20Nal (blue) transition curves were measured by change in intrinsic tyrosine fluorescence.

The variant W20H of foldon was chosen for the particular reason that a hydrogen bond is formed between the nitrogen's hydrogen atom in the side chain of the tryptophan at position 20 and the backbone carbonyl group of proline 4. This hydrogen bond is missing in the W20Nal variant and could be restored by the mutation to a histidine. The inserted histidine side chain is however highly sensible to the surrounding pH, which can be used to determine

4. Results

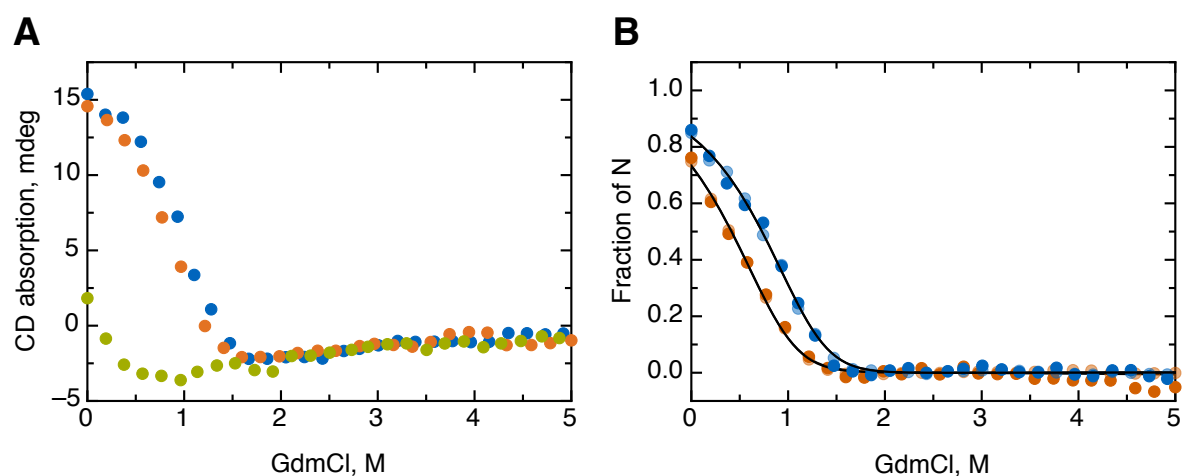


Figure 4.40.: GdmCl transition curves of foldon W20H in the presence of 500 mM sodium sulfate at varying pH-values. A) Original circular dichroism absorption data of foldon W20H GdmCl transitions at 60 μ M monomer concentration in the presence of 500 mM sodium sulfate at pH 5.5 (green), pH 7.0 (orange), and pH 8.0 (blue). B) Normalized GdmCl transition curves in the presence of 500 mM sodium sulfate of foldon W20H at pH 7.0 (orange) and pH 8.0 (blue), measured by intrinsic tyrosine fluorescence change (opaque points) or change in circular dichroism absorption at 228 nm (translucent points).

the effect of this hydrogen bond on the folding mechanism and stability. Two additional transitions were carried out at pH 5.5 and pH 8.0 to cover the complete area around the pK_a of the imidazole side chain. A protonation of the histidine side chain leads to a positive charge of the side chain within the hydrophobic core of foldon. This leads to a strong destabilization, making the detection of a native base-line impossible. This is shown in Figure 4.40A. The determination of a free folding enthalpy for foldon W20H is therefore impossible at pH 5.5. Fitting of the transition data at pH 8.0 and also comparison of the transition curves at pH 7.0 and pH 8.0 however shows a stabilization by $2.4 \text{ kJ} \cdot \text{mol}^{-1}$ to $\Delta G^0(\text{H}_2\text{O}) = (-57.4 \pm 0.6) \text{ kJ} \cdot \text{mol}^{-1}$. The transition was only determined at one foldon monomer concentration (60 μ M), hence fitting was carried out using the same fixed m_{eq} -value as for the previous transitions of foldon W20H. The increase in free folding enthalpy is caused by the fact that at pH 8.0, all side chains of histidine are uncharged, while there is still a small fraction (9%) of protonated and thus charged side chains at pH 7.0. This positive charge within the hydrophobic core does have a strong destabilizing effect, as seen from the almost complete unfolding of W20H at pH 5.5, where 76 % of the histidine side chains are protonated, according to the Henderson-Hasselbalch equation²⁸ and assuming a pK_a of 6.0 for the histidine side chain¹. This reveals that the pK_a of the histidine side chain within the hydrophobic cluster is only

slightly perturbed compared to histidine free in solution. The transition curves at all pH-values and a comparison of the pH 7.0 and pH 8.0 normalized fluorescence transition curve are given in Figure 4.40. As seen for foldon W20Nal, unfolding of foldon W20H with GdmCl can be measured, while refolding experiments fail due to an undeterminable fluorescence signal change or covering of the signal by mixing artifacts or photo-bleaching. Plotting of the common logarithm of the apparent unfolding rate constants against the effective GdmCl concentration reveals a kink in the plot. The slow unfolding rate constant is determined with $k_{u,slow} = (5.35 \pm 0.19) \cdot 10^{-5} \text{ s}^{-1}$ and an $m_{u,slow} = (7.86 \pm 0.32) \text{ kJ} \cdot \text{mol}^{-1} \cdot \text{M}^{-1}$, the fast phase has a rate constant of $k_{u,fast} = (6.55 \pm 1.03) \text{ s}^{-1}$ with $m_{u,fast} = (1.26 \pm 0.12) \text{ kJ} \cdot \text{mol}^{-1} \cdot \text{M}^{-1}$. It is striking that unfolding of W20Nal is the fastest unfolding rate constant observed of all foldon variants, while unfolding of W20H is among the slowest unfolding variants in this study. Refolding of foldon W20H thus must be even slower than refolding of foldon W20Nal, explaining the failure of the refolding experiments.

4.13. Foldon L22F

In the X-ray structure of foldon L22F, the distance between the carbon atoms of the side chains at position 13 and 22 is still around 3.5 Å, but the phenyl ring of the phenylalanine is not able to interact favorably with the tyrosine side chain. For favorable interactions between two aromatic side chains, the aromatic side chain rings have to be ordered in a parallel fashion to allow formation of π - π bonds between the delocalized π -electrons. The aromatic groups in the L22F variant however are not ordered in parallel fashion, but the phenyl ring of phenylalanine is tilted with respect to the tyrosine 13 side chain. This minimizes the possible van der Waals interactions and is contrary to the desired effect, as the exchange of the leucine to a phenylalanine should have led to a higher contact interface between the residues 13 and 22. The interaction of the leucine with its neighboring tyrosine occurs only by the C₅ carbon atom. The bonds of this carbon atom to the C₄ of the leucine side chain is oriented in an almost parallel plane with respect to the tyrosine's phenol group. One of the hydrogens bound to the C₅ is thus very likely to point directly towards the delocalized π -electron system, which will lead to a favorable interaction due to the partial positive charge of the hydrogen atom. The hydrogens bound to the phenyl ring of phenylalanine are not in a favorable orientation to interact in the described fashion. Instead, the orientation of the hydrogen atoms results in steric hindrance. This steric hindrance between the side chains of tyrosine 13 and phenylalanine 22

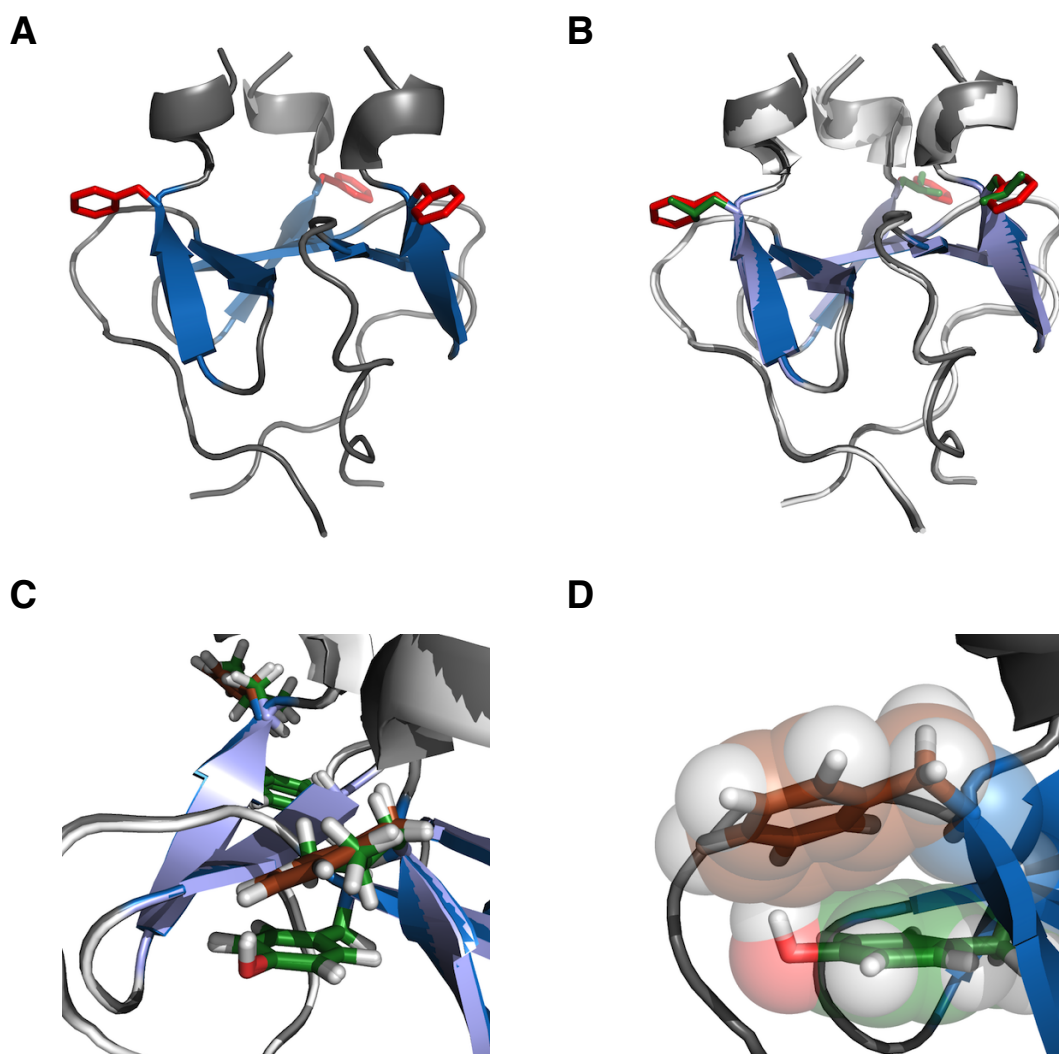


Figure 4.41.: Crystal structure of foldon L22F. A) Crystal structure of foldon L22F in ribbon representation, with the β -strands highlighted in blue, while the 3_{10} -helix and the other regions of the structure are colored in grey. The side chain of the mutated residue 22 is shown in stick representation and the side chain colored in red. B) Backbone alignment of foldon L22F and foldon wild-type. Foldon wild-type, also shown in ribbon representation, is colored in white, with the exception of the β -strands, which are colored in light blue. The leucine 22 residue is shown in stick representation and colored in green. The foldon L22F crystal structure is shown and colored as in A). C) Overlay of the residue 22 side chains of both foldon wild-type and foldon L22F, together with the neighboring residue Tyr13. The all-atom representation shows all carbon atoms in green (with the exception of Phe22, where the carbon atoms are shown in brown), the oxygen atoms in red and the hydrogen atoms in white. D) Stick representation of the residues Tyr13 and Phe22 colored as described in Figure 4.2. The van der Waals-radii for all atoms are presented to show the unfavorable overlap of the π orbitals of the side chains, due to the tilted orientation of the phenylalanine side chain at position 22 with respect to the tyrosine side chain at position 13.

rather results in a repulsion, probably hindering efficient hydrophobic cluster formation and thereby the orientation of the C-terminus with respect to the rest of the molecule.

The fluorescence spectra of all states of foldon L22F largely resembles the foldon wild-type spectra, with the exception that the intensity is reduced in all cases by up to 50 %. The amino acid exchange results in a destabilization of the native state, with $\Delta G^0(\text{H}_2\text{O}) = (-72.3 \pm 0.7) \text{ kJ} \cdot \text{mol}^{-1}$ and $m_{eq} = (12.4 \pm 0.3) \text{ kJ} \cdot \text{mol}^{-1} \cdot \text{M}^{-1}$ for the GdmCl transition. In urea the determined stability is $\Delta G^0(\text{H}_2\text{O}) = (-68.4 \pm 0.7) \text{ kJ} \cdot \text{mol}^{-1}$ and $m_{eq} = (6.2 \pm 0.2) \text{ kJ} \cdot \text{mol}^{-1} \cdot \text{M}^{-1}$, making L22F more stable in urea than A6S, which has a lower free folding enthalpy in guanidinium chloride. GdmCl unfolding experiments of foldon L22F show a high similarity to foldon Y13F: The plot of the common logarithms of the unfolding apparent rate constant is linear over the whole range of used GdmCl concentrations, unfolding is fast and all traces can be sufficiently fitted to a single-exponential equation.

In urea unfolding experiments of foldon L22F, two phases contribute significantly to the overall fluorescence signal change. Fitting of the unfolding traces is achieved by a two-exponential equation. Plotting of the common logarithm of the apparent rate constants of the two phases against the final urea concentrations reveals that both curves yield a chevron-plot. Fitting of these plots results in a folding rate constant $k_f = (4.68 \pm 5.81) \cdot 10^3 \text{ s}^{-1}$ with a $m_f = (10.02 \pm 1.01) \text{ kJ} \cdot \text{mol}^{-1} \cdot \text{M}^{-1}$ and $k_u = (1.42 \pm 0.73) \cdot 10^{-4} \text{ s}^{-1}$ with $m_u = (1.14 \pm 0.20) \text{ kJ} \cdot \text{mol}^{-1} \cdot \text{M}^{-1}$ for the fast phase. The plot of the slower apparent rate constant results a folding rate constant $k_f = (8.33 \pm 3.26) \cdot 10^{-4} \text{ s}^{-1}$ with a $m_f = (1.67 \pm 0.38) \text{ kJ} \cdot \text{mol}^{-1} \cdot \text{M}^{-1}$ and $k_u = (1.00 \pm 0.28) \cdot 10^{-5} \text{ s}^{-1}$ with $m_u = 1.30 \pm 0.95 \text{ kJ} \cdot \text{mol}^{-1} \cdot \text{M}^{-1}$. As can be clearly seen from the plot in Figure 4.42D, the unfolding limbs are almost parallel and hence yield nearly identical m_u -values, while the resulting k_u deviate by one order of magnitude. For the refolding limbs a strong difference is seen, as refolding of the faster apparent rate constant is 7 orders of magnitude faster than the slower one, accompanied by a five-fold higher m_f -value for the faster phase. The slower of the two apparent rate constants however contributes always at least 80 % to the complete amplitude change while the fast apparent rate contributes between 5 and 20 % of the total amplitude, depending on the final GdmCl concentration. The largest contribution of the fast apparent rate constant is observed for the highest final urea concentration, and this decreases in an almost linear fashion towards lower urea concentrations, as shown in Figure A.22.

4. Results

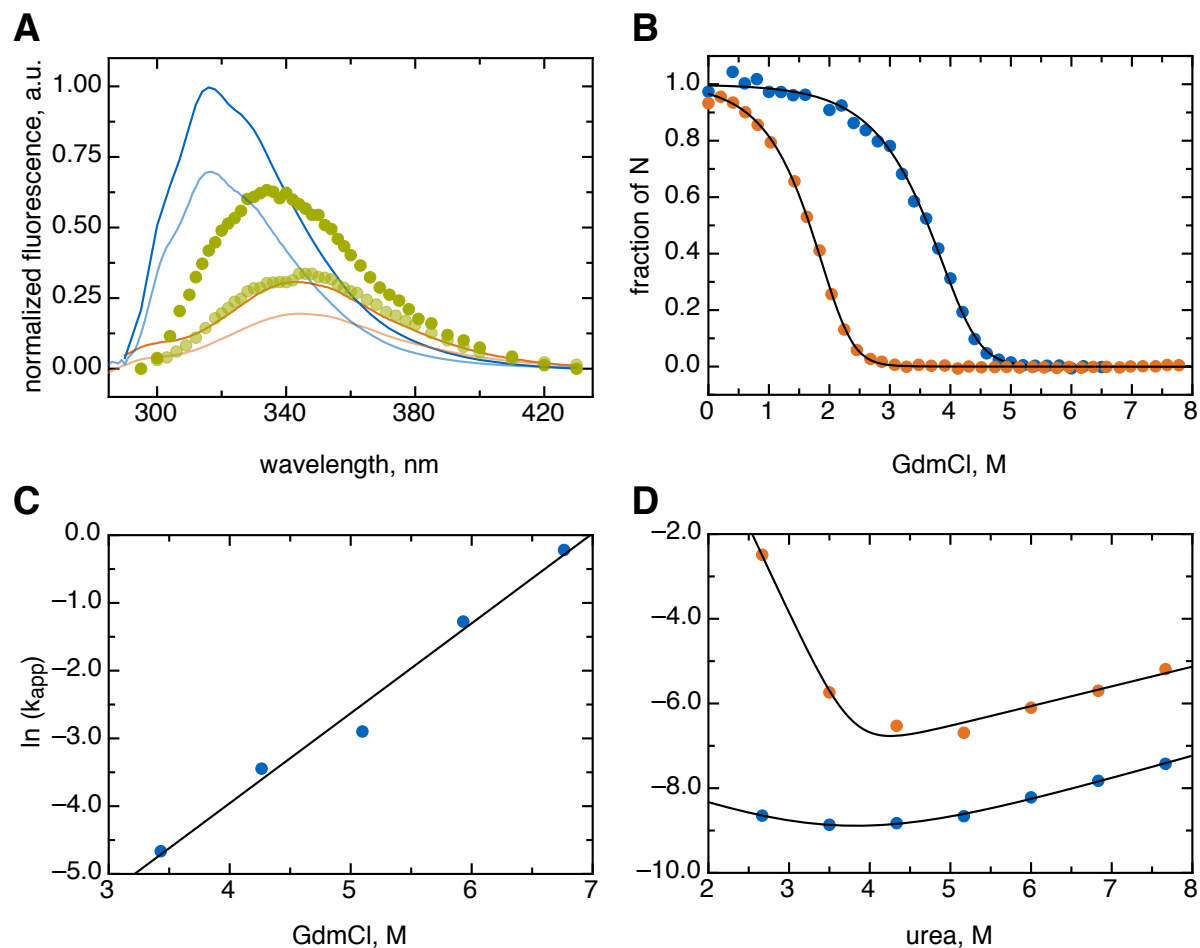


Figure 4.42.: A) Fluorescence spectra of foldon wild-type and foldon L22F. The native state spectra are colored in blue, the spectra of the unfolded state in orange and the burst-phase intermediate spectra in green. The fluorescence spectra of foldon L22F are lighter colored than the corresponding foldon wild-type spectra. B) Comparison of the foldon wild-type (blue) and foldon L22F (orange) GdmCl transitions at 30 μ M monomer concentration. C) Linear unfolding plot of the common logarithms of the apparent rate constants of foldon L22F GdmCl unfolding kinetics against the corresponding GdmCl concentrations. D) Chevron plots of the two phases of the foldon L22F urea unfolding kinetics. The common logarithms of the apparent rate constants are plotted against the corresponding urea concentrations.

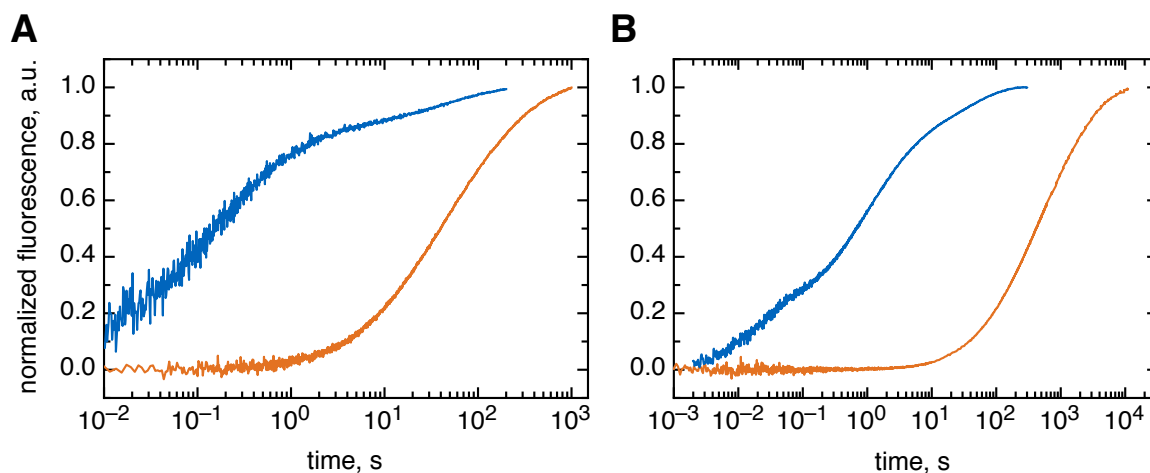


Figure 4.43.: Comparison of foldon wild-type (blue) and foldon L22F (orange) refolding traces A) 10 μ M foldon monomer refolding traces in 0.58 M GdmCl. B) 10 μ M foldon monomer refolding traces in 1.0 M urea.

The determination of the reaction order for refolding of foldon L22F from the half-life times of the concentration-independent phases reveals broken reaction orders for both denaturants, with 2.82 ± 0.03 in GdmCl and 2.89 ± 0.11 in urea. The reaction order determined from the initial slope is in both cases close to 2, and the dimerization rate constant determined from the y-intercept is $k_D = (5.76 \pm 0.09) \cdot 10^4 \text{ s}^{-1} \cdot \text{M}^{-1}$ in GdmCl and $k_D = (1.35 \pm 0.32) \cdot 10^3 \text{ s}^{-1} \cdot \text{M}^{-1}$ in urea. The refolding traces of foldon L22F are adjusted using a free folding enthalpy of $-72.2 \text{ kJ} \cdot \text{mol}^{-1}$. The association rate constants are decelerated by two orders of magnitude for k_D and one order of magnitude for k_T in GdmCl compared to foldon wild-type. The dissociation rate constant of the native state shows a significant shift by a factor of two after the free global fit. This is accompanied by a decrease of the dimer dissociation rate from 85 s^{-1} to 29 s^{-1} , which are both in the range of the wild-type dissociation rate constant. The dissociation rate constant of the trimeric intermediate T is decreased by a factor of 500 compared to foldon wild-type. This indicates a strong stabilizing effect on the trimeric intermediates of the amino acid exchange from the leucine to a phenylalanine at position 22. The rate constant of the last rearrangement step to the native state is not perturbed by this stabilization.

The rate constants determined for foldon L22F in urea are overall highly comparable to the other foldon variants that make use of the modified model for rate constant determination. However, the rate constant for the dissociation of the trimeric intermediate states, k_{-T} , is only $1.1 \cdot 10^{-2}$ or $3.3 \cdot 10^{-2}$ for the fixed and free fit, respectively. This deceleration by more than

4. Results

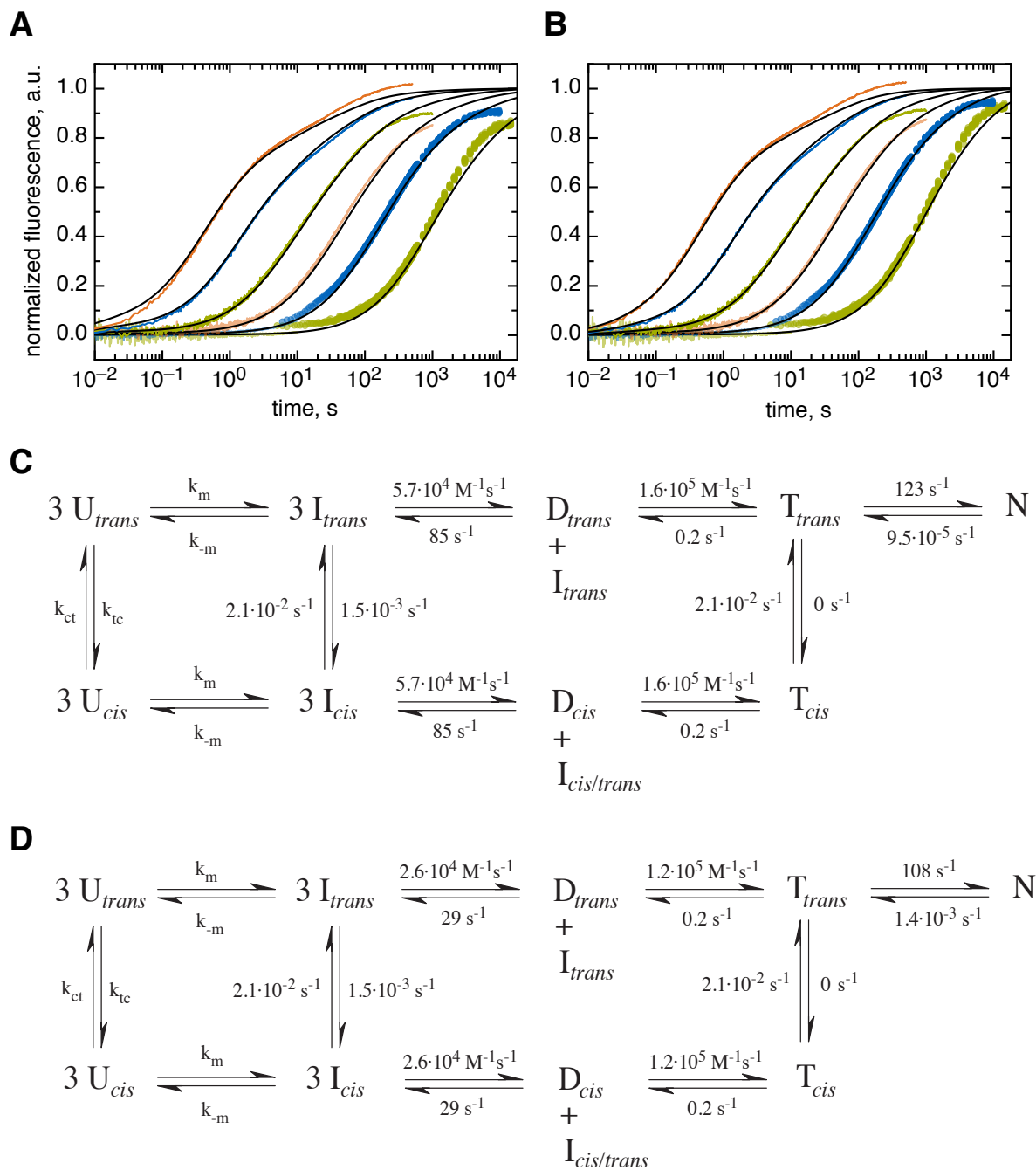


Figure 4.44.: Normalized global fit traces of foldon L22F GdmCl refolding kinetics. A) Normalized GdmCl refolding traces after fitting with fixed dimerization rate constant k_D . B) Normalized GdmCl refolding traces after fitting with all parameters run freely. C) Fitting model with all determined rate constants with fixed dimerization rate constant k_D . D) Fitting model with all determined rate constants with all parameters run freely.

three orders of magnitude compared to foldon P4T and foldon Y13F hints to a strong stabilization of the trimeric intermediate states, once successful association is achieved. This finding is supported by the extremely slow unfolding of foldon L22F in urea, despite the comparable stability with the other variants. The slow unfolding rates result in two chevron plots observed upon unfolding in urea. The rate constants for dimerization and trimerization are of the same magnitude as in foldon Y13F. Unfortunately, the global fit is not as well defined as for the other foldon variants, especially regarding the higher foldon monomer concentrations.

For foldon wild-type, an N -test had to be performed to allow exact determination of all rate constants. Although this is not the case while using the KinTek Software, undefined rate constants still occur. Especially k_T and k_{-D} are highly coupled in the urea refolding fits. To gain a more robust and better defined fit, an N -test was carried out for foldon L22F at 100 μM foldon monomer concentration after initiating refolding. The normalized amplitudes are given in Figure 4.46, together with the fit of the traces after global fitting to both the refolding and the N -test plots.

It is obvious from the plots that N alone cannot account for the complete amplitude detected from the N -test, especially for the early time-points. However, if including T_{cis} into the fit, the trace represents the data points sufficiently. Observable unfolding of the trimeric intermediate states of foldon is an explanation for the fast phases with low amplitudes observed for GdmCl and urea unfolding traces, which are predominantly observed for low denaturant concentrations, as the trimeric intermediate states T_{cis} and T_{trans} should only be a minor species during the folding reaction and should unfold considerably faster than the native state N . Although the N -test can be fitted with the used model, it does not result in a better definition of the rate constants k_T and k_{-D} , which remain coupled and cannot be completely defined. Thus only one N -test was carried out, and fitting for all foldon variants was achieved without additional information from interrupted refolding experiments.

4. Results

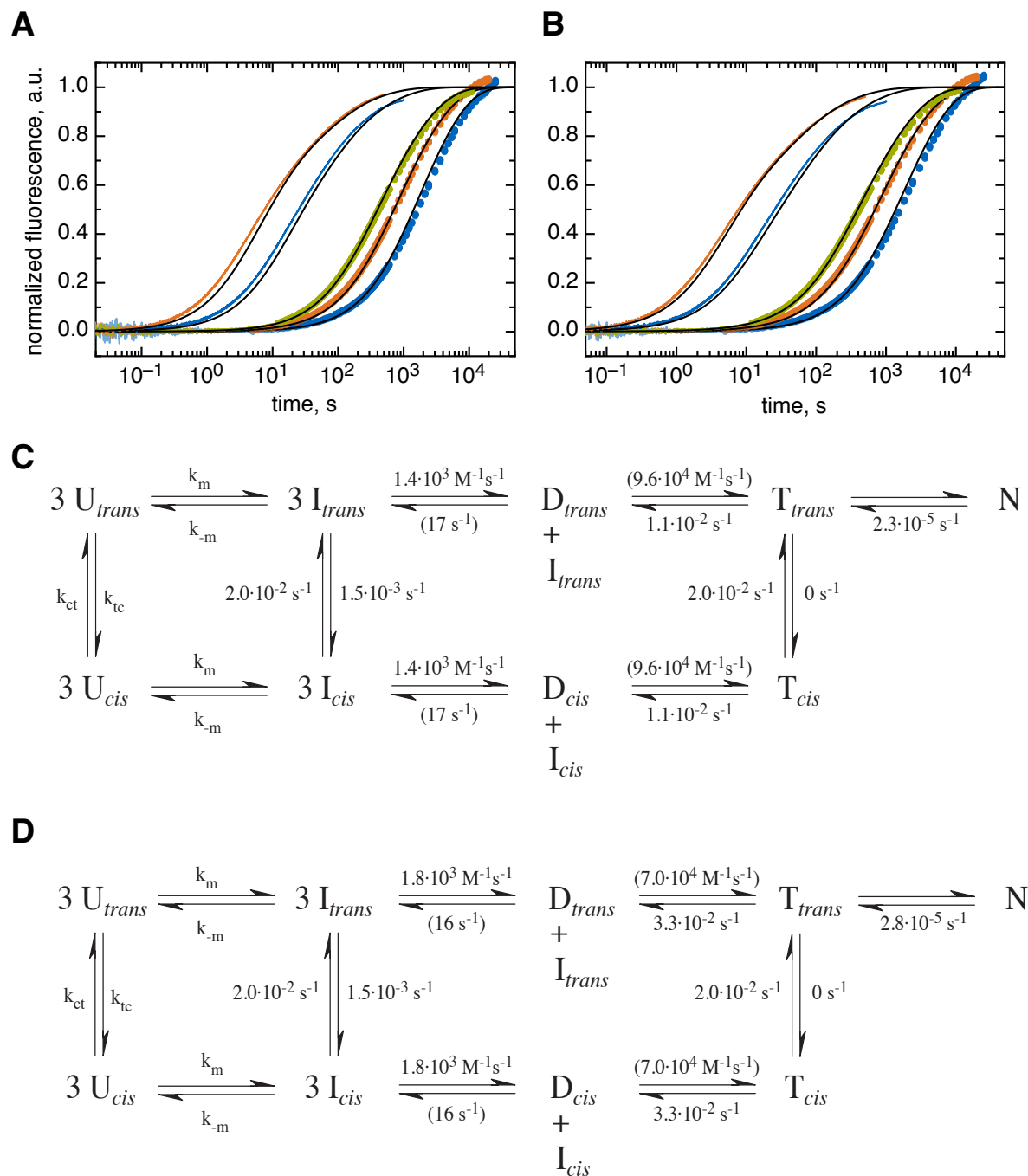


Figure 4.45.: Normalized global fit traces of foldon L22F urea refolding kinetics. A) Normalized urea refolding traces after fitting with fixed dimerisation rate constant k_D . B) Normalized urea refolding traces after fitting with all parameters run freely. C) Fitting model with all determined rate constants with fixed dimerisation rate constant k_D . D) Fitting model with all determined rate constants with all parameters run freely.

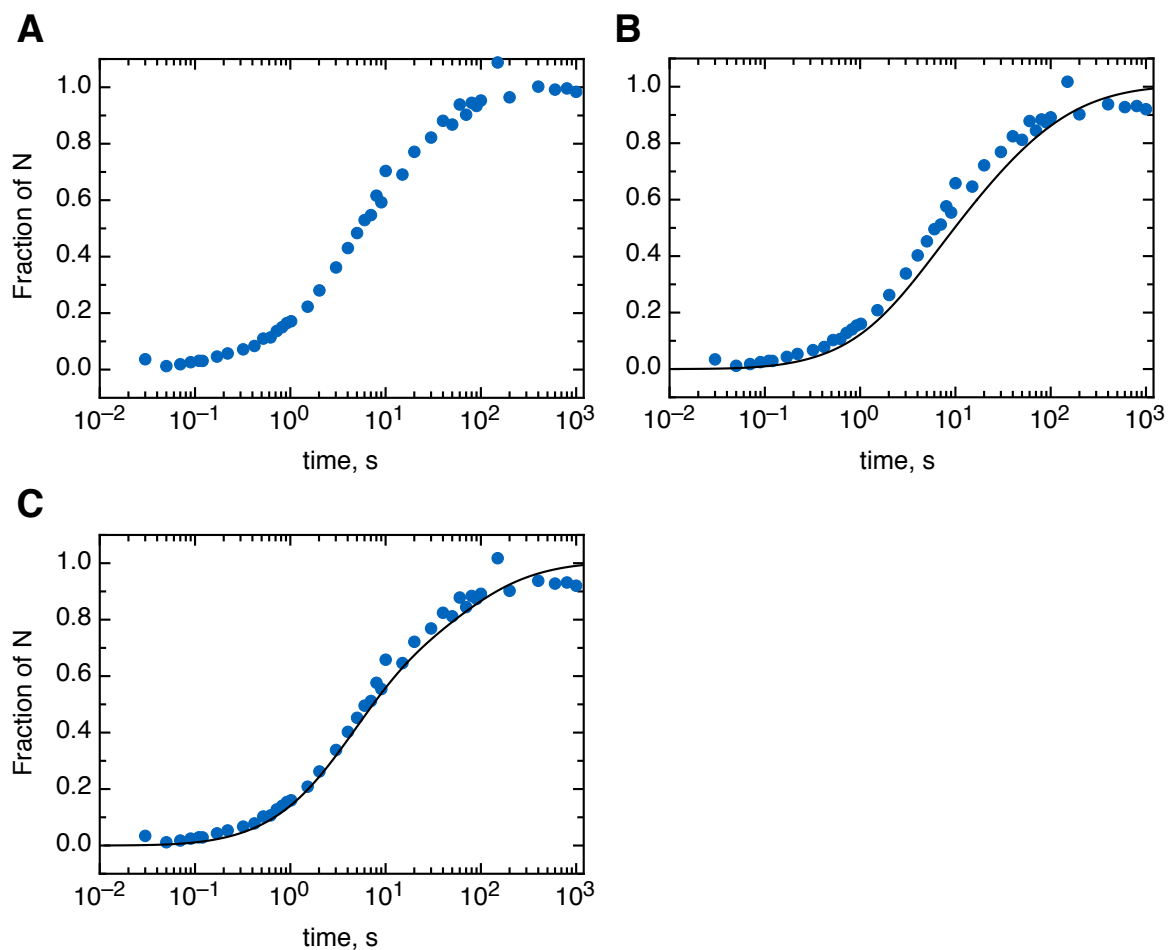


Figure 4.46.: Interrupted refolding experiment of foldon L22F in urea. A) Normalized amplitudes of urea unfolding after refolding for varying time-spans. B) Global fit of the amplitude, using the fraction of the native state N only. The final fraction of N is adjusted to the stability under the used experimental conditions. C) Global fit of the amplitude, using the fraction of the native state N and the trimer in *cis* conformation T_{cis} . The trimer in all-*trans* conformation is not included, as it is spectroscopically silent under the chosen experimental conditions. The final fraction of N is adjusted to the stability under the used experimental conditions.

4.14. Denaturant transitions

Denaturant transitions of proteins are an efficient way to determine the free folding enthalpy $\Delta G^0(H_2O)$. In this study, determination of the stability of the different foldon variants allows a first glimpse of the impact of each mutation. Furthermore, the stability allows the calculation of the natively folded fraction of the total protein present in solution. This allows to determine if the fraction of N is high enough at the refolding conditions to measure and evaluate refolding traces. The fraction of N is also necessary for the exact evaluation of the refolding experiments by global fitting. The stability of every obtained foldon variant is determined by GdmCl transitions at at least two different foldon monomer concentrations. An overview of all determined GdmCl transition results is given in Table 4.3, which also allows a comparison of the $\Delta\Delta G^0(H_2O)$ values between the different foldon variants. The transition plots are shown in Appendix A.2. An interesting finding is the lacking difference in stability of foldon A6Abu and A6S. The incorporation of an additional methyl group or an additional hydroxyl group in the side chain has a rather small impact on the folding stability, while the impact on the m_{eq} -value is almost 30 %. Thus, stability in this case could be directly related to the steric hindrance in the tightly packed hydrophobic core, while the m_{eq} -value is proportional to the fact that the hydrophobic packing is disturbed by the insertion of a hydrophilic hydroxyl group.

	$\Delta G^0(H_2O)$ kJ · mol ⁻¹	m_{eq} kJ · mol ⁻¹ · M ⁻¹	$\Delta\Delta G^0(H_2O)$ kJ · mol ⁻¹
P4T	-101.5 ± 3.0	11.4 ± 0.7	-12.3
A6V	-62.6 ± 0.3	11.9 ± 0.1	26.6
A6S	-70.2 ± 0.3	14.2 ± 0.1	19.0
A6Abu	-73.7 ± 1.4	11.0 ± 0.5	15.5
P7F	-46.5 ± 11.3	13.2 ± 4.2	42.7
P7V	-53.0 ± 6.0	10.7 ± 2.4	36.2
P7Nva	-58.7 ± 1.3	12.4 ± 0.7	30.5
P4T P7Nva	-70.2 ± 6.0	14.2 ± 3.2	19.0
Y13F	-64.4 ± 0.1	13.8 ± 0.1	24.8
L22F	-72.3 ± 0.7	12.4 ± 0.3	16.9
wild-type	-89.2 ± 0.6	10.4 ± 0.2	0.0

Table 4.3.: Folding stabilities and m_{eq} -values of all foldon variants, determined by guanidinium chloride transitions in 20 mM sodium cacodylic acid buffer pH 7.0, 20.0 °C.

Urea transitions of 5 chosen foldon variants were carried out and their stabilities determined. These variants are all destabilized compared to foldon wild-type, which allows complete unfolding in urea. A comparison of all free folding enthalpies determined by urea transitions is given in Table 4.4, the transition plots shown in Appendix A.3. A comparison of the determined stabilities is given in Table 4.5. The m_{eq} -values of all foldon variants are larger than for foldon wild-type in both denaturants. In average, the m_{eq} -value is about a factor 2 smaller in urea than in GdmCl, ranging from 2.4 for foldon A6S to 1.9 in foldon wild-type. This corresponds to the fact that GdmCl is twice as strong as a denaturant than urea⁵⁷.

	$\Delta G^0(H_2O)$ kJ · mol ⁻¹	m_{eq} kJ · mol ⁻¹ · M ⁻¹	$\Delta\Delta G^0(H_2O)$ kJ · mol ⁻¹
A6S	-65.3 ± 0.7	5.8 ± 0.2	23.9
A6Abu	-69.3 ± 0.7	5.3 ± 0.1	25.3
P7Nva	-54.6 ± 0.6	6.2 ± 0.2	34.6
Y13F	-58.2 ± 1.3	6.9 ± 0.4	31.0
L22F	-68.4 ± 0.7	6.2 ± 0.2	20.8
wild-type	-89.2 ± 0.3	5.4 ± 0.1	0.0

Table 4.4.: Folding stabilities and m_{eq} -values of chosen foldon variants, determined by urea transitions in 10 mM sodium cacodylic acid buffer pH 7.0, 20.0 °C.

Comparison of the urea and GdmCl transitions of a given foldon variant shows that the determined stabilities differ for the two denaturants, as shown in Table 4.5. The difference is in average about 4.7, with the GdmCl transitions resulting in an apparently higher stability for all foldon variants for which the stability has been determined using both denaturants. The reason could be the poorly defined native baseline for most GdmCl transitions, compared to the transitions carried out using urea. This is caused by the stronger denaturing effect of GdmCl, possibly leading to an incorrect extrapolation to 0.0 M GdmCl, resulting in a shifted folding stability. This is supported by the observation that for foldon wild-type, no difference in the $\Delta G^0(H_2O)$ -values determined by GdmCl and urea transitions is observed. As the stability of foldon wild-type is higher than for all studied mutants, the native baselines for both denaturants and all chosen foldon monomer concentrations are well-defined.

To overcome the problem of differing $\Delta G^0(H_2O)$ -values, a global fit of all denaturant transition curves in both GdmCl and urea is carried out as described in Chapter 3.5. The folding stability $\Delta G^0(H_2O)$ and the m_{eq} -values for both GdmCl and urea are given in Table 4.6. The determined stabilities are located between the values for GdmCl and urea. The globally de-

4. Results

	$\Delta G^0(H_2O)$ (GdmCl) kJ · mol ⁻¹	$\Delta G^0(H_2O)$ (urea) kJ · mol ⁻¹	$\Delta\Delta G^0(H_2O)$ kJ · mol ⁻¹
A6S	-70.2 ± 0.3	-65.3 ± 0.7	4.9
A6Abu	-73.7 ± 1.4	-69.3 ± 0.7	4.4
P7Nva	-58.7 ± 1.3	-54.6 ± 0.6	4.1
Y13F	-64.4 ± 0.1	-58.2 ± 1.3	6.2
L22F	-72.3 ± 0.7	-68.4 ± 0.7	3.9
wild-type	-89.2 ± 0.6	-89.2 ± 0.3	0.0

Table 4.5.: Comparison of the $\Delta G^0(H_2O)$ -values of foldon variants determined by both GdmCl and urea transitions in their corresponding buffers at 20 °C.

terminated stability of foldon L22F is close to the value for GdmCl, while it is near the value determined in urea for foldon P7Nva. All other globally determined stabilities are approximately in the middle between the two values determined from the single GdmCl and urea transitions. The m_{eq} -value for GdmCl of the global fit is almost always lower than the fit for GdmCl alone, again with the exception of L22F. For urea, the opposing effect is seen, where all m_{eq} -value are increased by approximately 1 kJ · mol⁻¹ · M⁻¹ compared to the original urea transition. This leads to a shift in the quotient of these values from approximately 2 found for the individual fits to an average value of about 1.7.

	$\Delta G^0(H_2O)$ kJ · mol ⁻¹	m_{eq} (GdmCl) kJ · mol ⁻¹ · M ⁻¹	m_{eq} (urea) kJ · mol ⁻¹ · M ⁻¹
A6S	-68.1 ± 0.5	13.1 ± 0.3	6.4 ± 0.1
A6Abu	-70.7 ± 0.8	9.7 ± 0.3	5.5 ± 0.2
P7Nva	-57.8 ± 0.4	11.8 ± 0.2	7.2 ± 0.2
Y13F	-59.1 ± 0.4	10.9 ± 0.2	7.1 ± 0.2
L22F	-72.2 ± 0.6	12.4 ± 0.3	7.4 ± 0.1
wild-type	-89.2	10.2	5.4

Table 4.6.: Folding stabilities and m_{eq} -values of chosen foldon variants, determined by global fitting of all transition curves using a global $\Delta G^0(H_2O)$ and different m_{eq} -values for both GdmCl and urea traces in their corresponding buffers at 20.0 °C. The values for foldon wild-type are taken from their individual transition fits and are shown for direct comparison; they are hence given without error range.

The systematic deviations between the stabilities determined by GdmCl and urea raised the question if this is just a fitting artifact due to the ill- or non-defined native baselines found for GdmCl. The finding that the difference between the determined stabilities is around 4 kJ ·

mol^{-1} and the fact that the ratio of the m_{eq} -values is 2 if evaluating the denaturant transitions independently speak against this explanation. Furthermore are the problems encountered during the global fitting of the GdmCl refolding traces discussed for the individual foldon variants another evidence for the different influence the denaturants exert. The fraction of native foldon at varying denaturant concentrations can be determined by carrying out unfolding experiments of foldon solution in buffer with varying GdmCl concentration and comparing the amplitudes of the respective unfolding experiments. For foldon Y13F, unfolding experiments of $10 \mu\text{M}$ foldon stock solutions are performed and the fraction of N is plotted against the exact GdmCl concentration. Comparison with the equilibrium transition of foldon Y13F at the same monomer concentration shows that the determined fractions superimpose very well with the equilibrium transition data points and the determined transition curve for this concentration. The most intriguing finding is the increase of natively folded trimer upon the increase of denaturant from 0.0 M to 0.2 M GdmCl. This indicates that for destabilized foldon monomers, low concentrations of GdmCl can have a stabilizing effect on the native state. This contradicts the Hofmeister effect of guanidinium chloride, which is considered strongly destabilizing⁴⁶. An inversion of the Hofmeister series may be observed for pH-values below the pI of a protein⁴⁹, but the isoelectric point of foldon is 4.78, as determined by ProtParam¹⁵⁹, which is well below pH 7.0. The increase in stability is also observed by an fluorescence increase in equilibrium transition experiments, but was always assigned to a fluorescence artifact due to the increased ionic strength without any impact on the stability.

The same experiment is also carried out using a $60 \mu\text{M}$ foldon L22F stock solution in buffer with varying GdmCl concentrations. As no equilibrium GdmCl transition was carried out for this monomer concentration, a direct comparison is not possible. However, equation 3.8 allows the calculation of a transition curve for any given monomer concentration and stability. Two curves are plotted together with the fraction of native state for the corresponding GdmCl concentration in Figure 4.47. The upper one corresponds to the stability determined by the GdmCl equilibrium transition with the corresponding m_{eq} -value, while the lower curve represents a hypothetical curve with the stability determined from the urea equilibrium transition and the m_{eq} -value of the GdmCl equilibrium transition. It is clear that the obtained data points correspond to the GdmCl transition rather than the hypothetical transition curve of the stability determined in urea. The N -tests carried out with both foldon Y13F and foldon L22F support the chosen stabilities used for the global GdmCl refolding experiments, although it is not clear how GdmCl can exert a stabilizing effect on foldon association and folding.

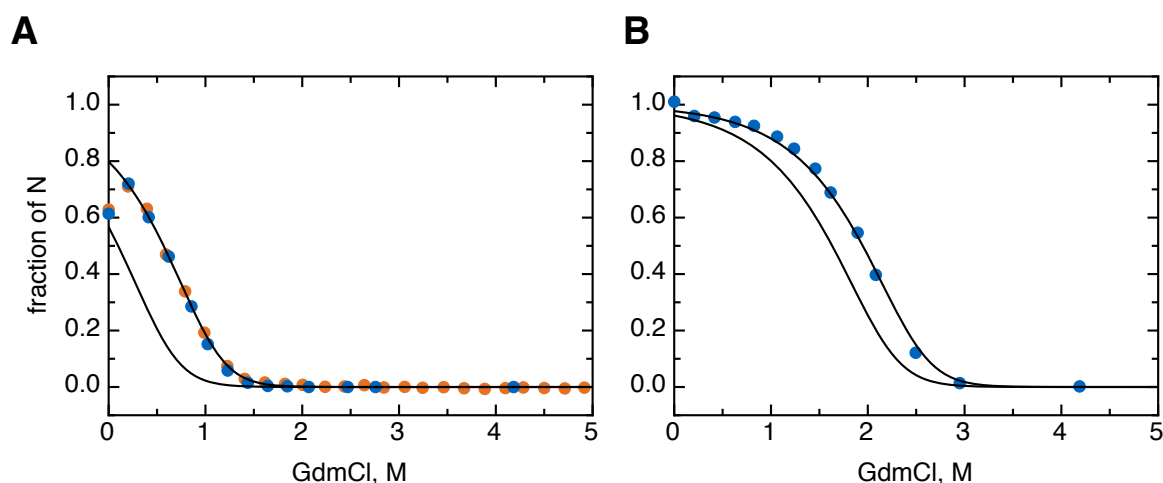


Figure 4.47.: Equilibrium N -test of foldon variants at varying foldon monomer concentrations. A) Foldon Y13F equilibrium N -test at $10 \mu\text{M}$ foldon monomer concentration. The determined fraction by the equilibrium N -test are shown in blue, while the orange data points are points determined from the equilibrium GdmCl transition. The upper solid line is the transition curve determined from the equilibrium GdmCl transition, while the lower one represents a hypothetical curve with the stability determined from the urea equilibrium transition and the m_{eq} -value of the GdmCl equilibrium transition. B) The same experiment as in A) is performed with $60 \mu\text{M}$ foldon L22F. The solid lines correspond to the lines in A).

4.15. Unfolding kinetics

Denaturant-dependent unfolding experiments of the various foldon variants allow the determination of the unfolding rate constant of the native state k_{-T} . Additionally, it is possible to determine the kinetic unfolding m_u -value. This allows the analysis of the transition state location by a rate-equilibrium free energy relationship (REFER) of the denaturation, resulting α_D (see Chapter 4.16).

Unfolding experiments were carried out in both GdmCl and urea. All unfolding traces can be fit with either a single- or a double-exponential equation, with the exception of foldon P4T P7Nva, which shows a more complicated unfolding behavior. The only exceptions are foldon wild-type and foldon P4T, which are too stable to determine denaturant-dependent unfolding kinetics at neutral pH. Unfolding by GdmCl allows only the determination of the unfolding limb of the chevron plot. For all foldon variants in the N-terminal poly-proline II helix, a kink

of the unfolding limb is observed, which can be interpreted as a shift of the transition state or the presence of a high-energy intermediate. When more than one apparent rate constant is determined, these are grouped according to their contribution to the total amplitude of the fluorescence change. The unfolding traces of foldon Y13F and foldon L22F can be readily fit with a single-exponential equation and yield a linear unfolding limb. The unfolding rate constant of the native trimer can directly be determined from the linear fits, together with the m_U -value. A comparison of all variants is given in Table 4.7. The unfolding traces and determination of k_{-N} are shown in Appendix A.4. When several rate constants can be determined, the slowest one is always considered to be k_{-T} . Any other observed rate constants are considered to be unfolding events of folding intermediates, especially T_{trans} and T_{cis} . These rates cannot be attributed directly and may be a mixture of several unfolding events. It is obvious by a comparison of the unfolding rate constants of foldon P4T P7Nva and foldon L22F with foldon wild-type that a decreased stability is not necessarily accompanied by an increase of the unfolding rate constant. The mutations at position 4 and 22 seem to increase the activation enthalpy of unfolding, resulting in a decrease of the unfolding rate constant k_{-T} compared to foldon wild-type. This is discussed in detail in Chapter 5.

	$k_{u,1}$ s^{-1}	$m_{u,1}$ $\text{kJ} \cdot \text{mol}^{-1} \cdot \text{M}^{-1}$	$k_{u,2}$ s^{-1}	$m_{u,2}$ $\text{kJ} \cdot \text{mol}^{-1} \cdot \text{M}^{-1}$
A6Abu	$(6.52 \pm 0.81) \cdot 10^{-4}$ $(4.09 \pm 0.38) \cdot 10^{-2}$	4.55 ± 0.64 1.88 ± 0.16	$(4.43 \pm 0.26) \cdot 10^{-2}$	2.22 ± 0.08
A6S	$(3.52 \pm 0.01) \cdot 10^{-3}$	4.57 ± 0.01	$(7.90 \pm 0.62) \cdot 10^{-2}$	2.54 ± 0.09
A6V	$(3.98 \pm 0.11) \cdot 10^{-3}$	4.45 ± 0.12	0.12 ± 0.06	2.52 ± 0.05
P7Nva	$(3.83 \pm 0.43) \cdot 10^{-3}$	4.39 ± 0.33		
P4T	$(1.15 \pm 0.19) \cdot 10^{-2}$	1.02 ± 0.36		
P7Nva	$(9.74 \pm 1.28) \cdot 10^{-4}$	1.08 ± 0.43		
	$(7.48 \pm 0.49) \cdot 10^{-5}$	1.26 ± 0.30		
Y13F	$(5.35 \pm 0.10) \cdot 10^{-4}$	4.57 ± 0.08		
L22F	$(9.46 \pm 0.57) \cdot 10^{-5}$	3.24 ± 0.24		
wild-type	$4.2 \cdot 10^{-4}$			

Table 4.7.: GdmCl unfolding experiments. The rate constants and kinetic m -values are determined by a linear fit to the common logarithm of the apparent unfolding rate constants, plotted against the final GdmCl concentration. The rate constant for foldon wild-type is determined from a global fit⁹⁵, thus no m_u -value is known.

4. Results

Urea unfolding kinetics are in general slower than unfolding in GdmCl. The first reason for this behavior is that urea is a weaker denaturant, making unfolding less strong and therefore slower. Additionally, refolding of all foldon variants, including wild-type, is slower in urea than in GdmCl. As the stability and therefore the overall equilibrium constant is similar at both refolding conditions, unfolding must be slower. Furthermore, due to the less strong unfolding strength of urea, unfolding does not occur over the whole range of used urea concentrations for three of the five used foldon variants. Instead, refolding at low urea concentrations is observed, giving the plots of the common logarithms of the apparent rate constants against the final urea concentrations the characteristic V-shape of chevron plots. An overview of all determined rate constants and kinetic m -values determined by urea unfolding experiments is given in Table 4.8. The unfolding traces and determination of k_{-N} are shown in Appendix A.5. Urea unfolding traces of foldon A6Abu, foldon A6S and foldon Y13F are fitted with a two-exponential equation, and foldon L22F unfolding traces are fitted with a three-exponential equation. The major phase of the fits contributes at least 80 % to the total amplitude change in all cases. A typical V-shape of a chevron plot is observed when the common logarithm of the apparent rate constant of the dominant phase is plotted against the final urea concentration. For foldon L22F, the two major phases contribute significantly to the overall fluorescence signal change. The third phase neglected, as it does not contribute significantly to the overall fluorescence signal change. Plotting of the apparent rate constants of the two major phases against the final urea concentration reveals that both curves yield a chevron-plot.

4.16. Determination of α_D -values

The properties of the transition barrier can be probed by linear rate-equilibrium free energy relationships (REFERs). They were first proposed by Leffler, stating that the free enthalpy of folding and the free activation enthalpy change linearly with a change of external factors as temperature, pressure or change in the composition of the buffer¹⁶⁷. Evaluation of linear REFERs allows the characterization of the transition state position and its properties, defining the α_x -values. The α_D -value represents the α -value determined by changes of the denaturant concentration. It can be readily determined by a plot of the common logarithm of the folding rate constant k_f determined from a chevron plot against the common logarithm of the equilibrium constant K_{eq} at the same conditions. The slope of this plot allows the determination of the α_D value. This value is directly proportional to the buried accessible surface area in the

	k_f s^{-1}	m_f $\text{kJ} \cdot \text{mol}^{-1} \cdot \text{M}^{-1}$	k_u s^{-1}	m_u $\text{kJ} \cdot \text{mol}^{-1} \cdot \text{M}^{-1}$
A6Abu	$(3.41 \pm 1.15) \cdot 10^{-2}$	0.98 ± 0.33	$(3.20 \pm 2.80) \cdot 10^{-4}$	1.34 ± 0.27
A6S	$(8.59 \pm 4.31) \cdot 10^{-2}$	1.99 ± 0.50	$(1.91 \pm 0.29) \cdot 10^{-3}$	1.24 ± 0.05
P7Nva			$(4.79 \pm 0.03) \cdot 10^{-4}$	2.19 ± 0.02
Y13F			$(1.39 \pm 0.01) \cdot 10^{-4}$	1.66 ± 0.01
L22F	k_{fast} $(4.68 \pm 5.81) \cdot 10^3$ k_{slow} $(8.33 \pm 3.26) \cdot 10^{-4}$	10.02 ± 1.01 1.67 ± 0.38	$(1.42 \pm 0.73) \cdot 10^{-4}$ $(1.00 \pm 0.28) \cdot 10^{-5}$	1.14 ± 0.20 1.30 ± 0.95
wild-type	$2.0 \cdot 10^{-5}$			

Table 4.8.: Urea unfolding experiments. The rate constants and kinetic m -values are either determined by a linear fit to the common logarithm of the apparent unfolding rate constants, plotted against the final urea concentration, or fitting of equation 3.17 to the plot. The rate constant for foldon wild-type is determined from a global fit¹⁵⁰, thus no m_u -value is known.

4. Results

transition state⁷¹, which follows from:

$$\alpha_D = \frac{\delta\Delta G_f^{0\dagger}/\delta[\text{Denaturant}]}{\delta\Delta G^0/\delta[\text{Denaturant}]} = \frac{m_f}{m_{eq}} \quad (4.1)$$

Most α_D -values reach from 0.6 to 1.0, indicating that the transition state has a higher similarity to the native than to the unfolded state⁵⁷. Curvatures in chevron plots can be analyzed by non-linear REFERs, and these are a powerful tool to characterize the transition state region. If a high-energy intermediate exists on the folding pathway, changes in denaturant concentration can shift the height of the energy barrier and thus the rate-limiting step during the refolding reaction. The high-energy intermediate is normally not detectable with equilibrium or kinetic measurements due to its low stability. Tendamistat was the first protein where the analysis of non-linear REFERs has led to the detection of an on-pathway high-energy intermediate¹⁶⁸. This led to an extensive analysis of apparent two-state folding systems with curvatures in either the unfolding and refolding limb of the chevron plot, where high-energy intermediates were found as well¹⁶⁹. The detection of a high-energy intermediate on the folding pathway is also a direct evidence when it is not clear whether or not an intermediate is located on or off the folding pathway. With both the equilibrium m_{eq} -value and the kinetic unfolding m_u -value known from equilibrium denaturant transition and denaturant-dependent unfolding experiments, respectively, it is possible to determine the α_D -value for each variant and denaturant. The m_{eq} -value is defined by:

$$m_{eq} = |m_f| + |m_u| \quad (4.2)$$

It is thus possible to calculate $|m_f|$ from determined m_{eq} and m_u values, assuming a two-state transition. As the folding kinetics of foldon are more complex, the evaluation of the α_D must be performed with caution. The α_D is calculated using equation 4.1. For several foldon variants, however, more than one m_u -value is determined from the GdmCl unfolding experiments. In this case, the m_u -value of the slowest unfolding rate constant is used for α_D determination. All results are given in Table 4.9.

The α_D -values are all in a range between 0.59 and 0.74, which is clearly within the range normally found for small proteins⁵⁷. The location of the transition state is thus more or less identical for all foldon variants. The only exception is foldon P4T P7Nva. As for all other foldon variants, only the m_u -value of the slowest unfolding step is used to determine α_D . This

	m_{eq} kJ · mol ⁻¹ · M ⁻¹	m_u kJ · mol ⁻¹ · M ⁻¹	m_f kJ · mol ⁻¹ · M ⁻¹	$\alpha_{D(f)}$
A6Abu	10.97 ± 0.54	4.55 ± 0.64	6.42	0.59
A6S	14.22 ± 0.07	4.57 ± 0.10	9.65	0.68
A6V	11.94 ± 0.07	4.45 ± 0.12	7.49	0.63
P7Nva	12.36 ± 0.71	4.39 ± 0.33	7.97	0.64
Y13F	13.80 ± 0.05	4.57 ± 0.08	9.23	0.67
L22F	12.45 ± 0.30	3.24 ± 0.24	9.21	0.74
P4T P7Nva	14.22 ± 3.19	1.26 ± 0.30	12.96	0.91

Table 4.9.: Determination of the α_D -values for all GdmCl unfolding experiments. The m_f -values are calculated using equation 4.2.

value is only $\frac{1}{3}$ of the values determined for all other foldon variants, hence the high α_D of 0.91. This indicates a strongly shifted transition barrier towards the folded state. For urea, the α_D is determined as described for GdmCl. All results are given in Table 4.10.

	m_{eq} kJ · mol ⁻¹ · M ⁻¹	m_u kJ · mol ⁻¹ · M ⁻¹	m_f kJ · mol ⁻¹ · M ⁻¹	α_D
A6Abu	5.26 ± 0.13	1.34 ± 0.27	3.92	0.74
A6S	5.77 ± 0.16	1.24 ± 0.05	4.53	0.78
P7Nva	6.23 ± 0.18	2.19 ± 0.02	4.04	0.65
Y13F	6.87 ± 0.39	1.66 ± 0.01	5.21	0.76
L22F	6.21 ± 0.15	1.30 ± 0.95	4.91	0.79

Table 4.10.: Determination of the α_D -values for all urea unfolding experiments. The m_f -values are calculated using equation 4.2.

The transition barrier for the transition in urea is shifted towards the native state compared to the same variant in GdmCl. However, while it is clear that global unfolding is achieved by using GdmCl as a denaturant, it is not clear if complete unfolding of the native state is within reach of all urea concentrations. This is obvious from the chevron plots of the foldon variants A6Abu, A6S and L22F: The sum of the determined m_u and m_f -values does not result the m_{eq} -value determined by the equilibrium transition, but falls short by approximately 3 kJ · mol⁻¹ M⁻¹, which is almost half of the overall m_{eq} -value. The determined m_u -values and thus the α_D 's have to be evaluated with caution.

4.17. Folding kinetics

A comparison of the different foldon variants can only be complete if the individual rate constants for each detectable folding and association step are known. Refolding experiments of several foldon variants are performed in both GdmCl and urea at varying foldon monomer concentrations. As several foldon variants refold too slowly to determine the refolding kinetics within the time-range of the stopped-flow, mixing and consequent fluorescence measurements in a fluorimeter were carried out for these variants. Global fitting of the refolding traces in both GdmCl and urea is carried out by the KinTek program, as described in detail in Chapter 3.10. The fidelity of the fit can be enhanced as the dimerization rate constant k_D and the unfolding rate constant k_{-N} are known from previous experiments. The initial slope of all foldon refolding experiments is used to both determine the reaction order at the start of the reaction and the dimerization rate constant k_D . The reaction order is 2 in all cases, as determined from the slope of the common logarithm of the initial slope plotted as a function of the common logarithm of the foldon monomer concentration. The ordinate intercept yields the dimerization rate constant k_D . The determined results for GdmCl and urea are summarized in Table 4.11 and 4.12, respectively. Each mutation results in a decrease of the dimerization rate constant by at least a factor of five, compared to foldon wild-type in GdmCl. The deceleration is largest for foldon P7Nva, which shows a deceleration by two orders of magnitude.

	reaction order	ordinate intercept	dimerization rate constant k_D $M^{-1} \cdot s^{-1}$
P4T	2.00 ± 0.06	5.86 ± 0.32	$(7.24 \pm 0.05) \cdot 10^5$
A6Abu	2.04 ± 0.05	5.52 ± 0.23	$(3.32 \pm 0.04) \cdot 10^5$
A6S	2.06 ± 0.14	5.31 ± 0.71	$(2.06 \pm 0.13) \cdot 10^5$
A6V	2.15 ± 0.15	4.16 ± 0.77	$(1.45 \pm 0.19) \cdot 10^4$
P7Nva	2.00 ± 0.10	4.01 ± 0.47	$(1.03 \pm 0.12) \cdot 10^4$
Y13F	2.00 ± 0.10	4.25 ± 0.49	$(1.79 \pm 0.12) \cdot 10^4$
L22F	1.94 ± 0.09	4.76 ± 0.43	$(5.76 \pm 0.09) \cdot 10^4$
wild-type	2.0 ± 0.1	6.0 ± 3.0	$(1.1 \pm 0.5) \cdot 10^6$

Table 4.11.: Initial slope values all foldon variants from the GdmCl refolding experiments.

The results given in Table 4.12 show that each mutation causes an at least 10fold deceleration of the dimerization reaction compared to foldon wild-type. Amino acid exchanges at position 6 have the least impact on dimerization, while an exchange at position 7 slows down the dimer-

	reaction order	ordinate intercept	dimerization rate constant k_D $M^{-1} \cdot s^{-1}$
P4T	2.09 ± 0.06	5.68 ± 0.32	$(4.80 \pm 0.06) \cdot 10^5$
A6Abu	2.09 ± 0.13	4.76 ± 0.62	$(5.75 \pm 0.13) \cdot 10^4$
A6S	1.93 ± 0.08	4.72 ± 0.38	$(5.25 \pm 0.08) \cdot 10^4$
P7Nva	2.01 ± 0.06	2.77 ± 0.26	$(5.89 \pm 0.09) \cdot 10^2$
Y13F	2.11 ± 0.16	3.41 ± 0.74	$(2.57 \pm 0.22) \cdot 10^3$
L22F	1.98 ± 0.20	3.13 ± 1.01	$(1.35 \pm 0.32) \cdot 10^3$
wild-type	2.0 ± 0.1	6.1 ± 0.1	$(1.1 \pm 0.1) \cdot 10^6$

Table 4.12.: Initial slope values of all foldon variants from the urea refolding experiments.

ization reaction by a factor of 1,000. Foldon Y13F and L22F show an intermediate effect, as both mutations lead to a deceleration by a factor of 100. As the effect on the dimerization reaction of each mutation is similar to the effect of the mutation on the whole refolding kinetics, as seen from the comparison of the foldon wild-type and foldon variant refolding traces given for each foldon variant in the respective chapter, it seems obvious that the dimerization step is most sensitive to changes in the hydrophobic cluster, governing the rate of the complete association reaction. Determination of the reaction order allows a better understanding of the refolding reaction and especially gives a first hint on the used association and folding mechanism and the prominent steps within a reaction. As for foldon wild-type in GdmCl⁹⁵, the half-life time of the concentration-dependent steps is determined from the normalized refolding traces. The common logarithms of the half-life times are plotted against the foldon monomer concentration and fit with equation 3.12. For an overview of all results see Table 4.13.

Foldon P4T, A6Abu and L22F refolding reactions have an apparent reaction order of three, which deviate from linear behavior for higher foldon monomer concentrations in foldon P4T and A6Abu. This is the same result as in foldon wild-type, indicating that the reaction mechanism is similar and that a concentration-independent folding step has to be considered that becomes rate-limiting for higher foldon monomer concentrations. In foldon P7Nva guanidinium chloride refolding is a second-order reaction, indicating that either dimerization or trimerization is affecting the complete refolding process, making the overall reaction of second order. Foldon A6V, P4T P7Nva and Y13F show a dependence of the reaction order on the foldon monomer concentration. While they show second order for low foldon monomer

4. Results

	reaction order	ordinate intercept	reaction order	ordinate intercept
P4T	3.19 ± 0.19	-11.95 ± 1.01		
A6Abu	2.96 ± 0.26	-9.71 ± 1.48		
A6S	2.57 ± 0.02	-7.53 ± 0.11		
A6V	2.05 ± 0.09	-3.59 ± 0.47	$3.01 \pm n.d.$	$-7.93 \pm n.d.$
P7Nva	1.92 ± 0.11	-2.64 ± 0.58		
P4T P7Nva	2.00 ± 0.11	-1.90 ± 0.59	2.92 ± 0.03	-6.27 ± 0.13
Y13F	2.04 ± 0.15	-2.91 ± 0.77	$3.24 \pm n.d.$	$-8.39 \pm n.d.$
L22F	2.82 ± 0.03	-7.43 ± 0.13		
wild-type	3.15 ± 0.19	-11.90 ± 1.00		

Table 4.13.: Reaction order for all GdmCl refolding experiments determined from the half-life times. Two reaction orders and the ordinate intercepts are given if a clear kink in the plot is observed and both regions can be fitted using equation 3.12.

concentrations, refolding becomes apparent third-order for higher concentrations. The transition point for all three variants is between 10 and 20 μM foldon monomer. Foldon A6S finally shows a broken reaction order of 2.57 over the complete concentration range. All foldon variants show broken reaction orders for their urea refolding kinetics. Those can be either closer to 2 like in P7Nva and Y13F, which are of second order in GdmCl, or closer to third-order, as in A6Abu, A6S and L22F. While the result is not surprising for foldon A6S, which shows more or less the same reaction order as in GdmCl, the shift is not clear for A6Abu and L22F. Only foldon P4T has an apparent reaction order of three, as does foldon wild-type.

	reaction order	ordinate intercept		reaction order	ordinate intercept
P4T	3.03 ± 0.08	-10.04 ± 0.45	A6Abu	2.67 ± 0.03	-7.31 ± 0.17
A6S	2.65 ± 0.08	-7.09 ± 0.38	P7Nva	2.13 ± 0.06	-3.06 ± 0.26
Y13F	2.17 ± 0.07	-2.80 ± 0.35	L22F	2.89 ± 0.11	-6.80 ± 0.52
wild-type	2.9 ± 0.1	-9.7 ± 0.3			

Table 4.14.: Half-life values all foldon variants from the urea refolding experiments.

For three foldon variants, foldon A6V, foldon P7Nva, and foldon Y13F, a reaction order of two is determined by plotting the half-life time of the association reaction as a function of common logarithm of the foldon monomer concentration. Comparison of the apparent rate constant from the half-life times plot with the dimerization rate determined from the initial slope plot reveals that these rates are not identical, but that the apparent rate from the half-life times is two orders of magnitude lower. Any rates determined from the half-life times are

therefore not comparable with other rates determined, neither from the initial slopes fit nor from the global fits. Fitting of the GdmCl refolding data was accompanied by severe difficulties for all foldon variants except foldon P4T and foldon P4T P7Nva, discussed in detail in chapter 4.3. All global fits carried out for urea refolding experiments can be performed with adjusted traces to the free folding enthalpy of the respective refolding conditions. This is in contrast to the global fit carried out for the GdmCl refolding experiments, where fitting in such a manner is not possible. It is possible to omit the concentration-independent rearrangement step of the trimer T_{trans} to the native state N for all refolding experiments in urea, as known from half-life times reaction order determination. Additionally, k_{ct} can be determined directly for some foldon variants in GdmCl and urea from fitting of the original traces to a three- or four-exponential equation, resulting the rate constant of the slow concentration-independent step which is the rate constant of the *cis-trans* proline isomerization reaction⁹⁵. With the knowledge of the *cis-trans* ratio of the peptide bond between a proline and its N-terminal amino acid⁸¹, or its proposed ratio as for 2-(L)-aminobutyric acid, k_{tc} can be calculated. Isomerization of the trimeric state with all monomers in an all-*trans* conformation T_t back to the trimeric state with at least one monomer in the *cis* conformation of the Ala6-Pro7 peptide bond is considered to be such a rare event that this rate constant, k_{-6} , is fixed to 0 for all experiments. For all fitting procedures two fits are shown. In the first one the dimerization rate constant k_D is hold constant to the value determined from initial slope determination. This is also done for the dissociation rate constant k_{-N} , which is determined from the unfolding experiments in GdmCl and urea, shown in Chapter 4.15. Both rate constants are run freely in the second global fit to compensate for any errors of the previous rate constant determination that could lead to significant deviations of the global fit from the original traces.

Another puzzling observation from the global refolding fits of the GdmCl refolding experiments is the need to include the last concentration-independent rearrangement step from the trimer the all-*trans* conformation T_{trans} to the native state N , even if no deviation from linear behavior is observed in the plots of the half-life times. This last reaction step is well-defined, although it never becomes rate-limiting for most foldon variants. The need for this step probably arises from the concurrent insertion of the dissociation rate constant of the trimer k_{-T} , which can become rather large and is often well-separated from both the dissociation rate constants of the dimer and the native trimer, k_{-D} and k_{-N} . A direct observation of the rate constant k_{-T} is not possible, and the inclusion of this additional step is, like for the adjusted stabilities discussed before, a consequence of the fitting procedure rather than a direct physical

4. Results

observation. An overview of all determined rate constants for the GdmCl and urea refolding experiments is given in Table 4.15 and Table 4.16, respectively. Rate constants are shown for both the fit with a fixed dimerization rate constant k_D and with this rate constant allowed to run free. Differences between these two methods are only marginal and are rarely larger than a factor of two. As a general result, the dimerization is decreased by all amino acid exchanges in the hydrophobic cluster by at least a factor of two, and up to a decrease by four orders of magnitude. The effect of each mutation on the refolding kinetics is discussed in detail for each foldon variant in the respective chapter.

The global fits of all foldon variant studies by urea refolding experiments yield robust fits and rate constants for each individual folding step. The only exception are the second association rate constant k_T , yielding the native trimer N , and the dissociation rate constant of the dimer D , k_{-D} . These two rate constants are strongly coupled. The fit traces are insensitive to changes of these two rate constants over several orders of magnitude as long as the coupling is prevailed. Thus the numbers given in Table 4.16 are not exact results and have to be treated with caution. However, this finding provides further evidence that dimerization is the critical step in the foldon association, and that further association to the trimeric state has to be fast as otherwise dissociation of the dimer will occur. The impact of dimer destabilization can be seen by the k_T/k_{-D} ratios given for each studied foldon variant in urea in Table 4.18. The rate constants for the GdmCl refolding kinetics are all well-defined and little to no coupling between rate constants is observed. Nevertheless the ration between the second association rate constant k_T and the dissociation rate constant of the dimer k_{-D} is also determined for GdmCl refolding experiments. All rate constants and the resulting ratio are given in Table 4.17.

Ratios are given for all refolding fits both with and without fixed k_D . The larger this ratio of the two rate constants, the more favorable is the formation of the trimeric state in contrast to the dissociation of the formed dimer. From Table 4.17 and Table 4.18 it is clear that the ratio is lowest for foldon Y13F, and that the ratio decreases the closer the mutation is towards (or in) the β -hairpin. Comparison with foldon wild-type shows that this ratio is decreases for all foldon variants by at least a factor of 2, with the exception of foldon P4T. For this variant, the increase of the ratio can clearly be assigned to the decreased dissociation rate constant of the dimer, as the increase in stability is not accompanied by an increase of the refolding speed. For all other variants, it is not clear wether this effect is due to a decrease of the trimer formation

	k_{et} s ⁻¹	k_{tc} s ⁻¹	k_D M ⁻¹ · s ⁻¹	k_{-D} s ⁻¹	k_T M ⁻¹ · s ⁻¹	k_{-T} s ⁻¹	k_N s ⁻¹	k_{-N} s ⁻¹	$\Delta G^0(H_2O)$ kJ · mol ⁻¹	$\Delta \Delta G^0$ kJ · mol ⁻¹
P4T	fixed k_D	$3.5 \cdot 10^{-2}$	$2.6 \cdot 10^{-3}$	$7.2 \cdot 10^5$	32	$5.4 \cdot 10^6$	26	216	$1.5 \cdot 10^{-7}$	-12.3
	free k_D	$3.5 \cdot 10^{-2}$	$2.6 \cdot 10^{-3}$	$1.5 \cdot 10^6$	50	$4.2 \cdot 10^6$	66	216	$1.4 \cdot 10^{-5}$	
A6Abu	fixed k_D	$2.1 \cdot 10^{-2}$	$2.3 \cdot 10^{-2}$	$3.3 \cdot 10^5$	319	$5.6 \cdot 10^6$	6	124	$6.2 \cdot 10^{-3}$	15.5
	free k_D	$2.1 \cdot 10^{-2}$	$2.3 \cdot 10^{-3}$	$5.8 \cdot 10^5$	206	$2.0 \cdot 10^6$	5	143	$8.9 \cdot 10^{-3}$	
A6S	fixed k_D	$2.1 \cdot 10^{-2}$	$2.4 \cdot 10^{-3}$	$2.1 \cdot 10^5$	40	$7.6 \cdot 10^5$	11	160	$1.5 \cdot 10^{-2}$	19.0
	free k_D	$2.1 \cdot 10^{-2}$	$2.4 \cdot 10^{-3}$	$2.2 \cdot 10^5$	42	$8.0 \cdot 10^5$	17	160	$9.5 \cdot 10^{-3}$	
A6V	fixed k_D	$1.9 \cdot 10^{-2}$	$2.1 \cdot 10^{-3}$	$1.5 \cdot 10^4$	11	$8.6 \cdot 10^4$	77	160	$1.1 \cdot 10^{-3}$	26.6
	free k_D	$1.9 \cdot 10^{-2}$	$2.1 \cdot 10^{-3}$	$7.3 \cdot 10^3$	5	$8.4 \cdot 10^4$	87	160	$1.5 \cdot 10^{-3}$	
P7Nva	fixed k_D			$1.0 \cdot 10^4$	35	$8.7 \cdot 10^4$	50	160	$3.8 \cdot 10^{-3}$	30.5
	free k_D			$1.0 \cdot 10^4$	35	$8.8 \cdot 10^4$	48	160	$4.1 \cdot 10^{-3}$	
Y13F	fixed k_D	$2.0 \cdot 10^{-2}$	$1.5 \cdot 10^{-3}$	$1.7 \cdot 10^4$	731	$7.2 \cdot 10^5$	43	160	$1.6 \cdot 10^{-3}$	24.8
	free k_D	$2.0 \cdot 10^{-2}$	$1.5 \cdot 10^{-3}$	$1.7 \cdot 10^4$	257	$2.5 \cdot 10^5$	43	160	$1.6 \cdot 10^{-3}$	
L22F	fixed k_D	$2.1 \cdot 10^{-2}$	$1.5 \cdot 10^{-3}$	$5.7 \cdot 10^4$	85	$1.5 \cdot 10^5$	0.2	123	$9.5 \cdot 10^{-5}$	16.9
	free k_D	$2.1 \cdot 10^{-2}$	$1.5 \cdot 10^{-3}$	$2.6 \cdot 10^4$	29	$1.2 \cdot 10^5$	0.2	108	$1.4 \cdot 10^{-3}$	
P4T				$1.5 \cdot 10^3$	405	$3.7 \cdot 10^5$			$4.2 \cdot 10^{-5}$	19.0
P7Nva										
wild-type	$2.0 \cdot 10^{-2}$	$5.7 \cdot 10^{-3}$	$1.9 \cdot 10^6$	59	$5.4 \cdot 10^6$	110	210	210	$4.2 \cdot 10^{-4}$	0.0

Table 4.15.: Determined rate constants for guanidinium chloride refolding experiments of different foldon variants using a global fit by KinTek^{165,166}. The rate constants for foldon wild-type were determined using a custom-written Matlab program⁹⁵. The $\Delta \Delta G^0$ -values are calculated with respect to the foldon wild-type stability.

4. Results

	k_{ct} s^{-1}	k_{jc} s^{-1}	k_D $M^{-1} \cdot s^{-1}$	k_{-D} s^{-1}	k_T $M^{-1} \cdot s^{-1}$	k_{-T} s^{-1}	k_{-N} s^{-1}	$\Delta G^0(H_2O)$ $KJ \cdot mol^{-1}$	$\Delta \Delta G^0$ $KJ \cdot mol^{-1}$
P4T	fixed k_D	$4.4 \cdot 10^{-2}$	$3.3 \cdot 10^{-3}$	$4.8 \cdot 10^5$	45	$6.8 \cdot 10^5$	74	$5.3 \cdot 10^{-7}$	<i>n.d.</i>
	free k_D	$4.4 \cdot 10^{-2}$	$3.3 \cdot 10^{-3}$	$3.8 \cdot 10^5$	45	$6.7 \cdot 10^5$	74	$5.3 \cdot 10^{-7}$	<i>n.d.</i>
A6Abu	fixed k_D	$1.9 \cdot 10^{-2}$	$2.1 \cdot 10^{-3}$	$5.8 \cdot 10^4$	123	$1.1 \cdot 10^6$		$1.1 \cdot 10^{-3}$	-65.3
	free k_D	$1.9 \cdot 10^{-2}$	$2.1 \cdot 10^{-3}$	$5.7 \cdot 10^4$	125	$1.1 \cdot 10^6$		$1.2 \cdot 10^{-3}$	23.9
A6S	fixed k_D	$1.9 \cdot 10^{-2}$	$2.1 \cdot 10^{-3}$	$5.3 \cdot 10^4$	205	$1.1 \cdot 10^6$		$1.9 \cdot 10^{-3}$	-69.3
	free k_D	$2.0 \cdot 10^{-2}$	$2.3 \cdot 10^{-3}$	$5.5 \cdot 10^4$	218	$1.0 \cdot 10^6$		$2.0 \cdot 10^{-3}$	25.3
P7Nva	fixed k_D			588	12	$2.2 \cdot 10^4$		$9.9 \cdot 10^{-4}$	-54.6
	free k_D			611	11	$2.0 \cdot 10^4$		$9.7 \cdot 10^{-4}$	34.6
Y13F	fixed k_D	$2.0 \cdot 10^{-2}$	$1.5 \cdot 10^{-3}$	$2.6 \cdot 10^3$	46	$1.1 \cdot 10^4$	63	$2.7 \cdot 10^{-4}$	-58.2
	free k_D	$2.0 \cdot 10^{-2}$	$1.5 \cdot 10^{-3}$	$2.7 \cdot 10^3$	33	$7.0 \cdot 10^3$	62	$2.7 \cdot 10^{-4}$	31.0
L22F	fixed k_D	$2.0 \cdot 10^{-2}$	$1.5 \cdot 10^{-3}$	$1.4 \cdot 10^3$	17	$9.6 \cdot 10^4$	$1.1 \cdot 10^{-2}$	$2.3 \cdot 10^{-5}$	-68.4
	free k_D	$2.0 \cdot 10^{-2}$	$1.5 \cdot 10^{-3}$	$1.8 \cdot 10^3$	16	$7.0 \cdot 10^4$	$3.3 \cdot 10^{-2}$	$2.8 \cdot 10^{-5}$	20.8
wild-type	$2.2 \cdot 10^{-2}$	$1.1 \cdot 10^{-3}$	$1.1 \cdot 10^6$	35	$4.5 \cdot 10^5$		$2.0 \cdot 10^{-5}$	-89.2	0.0

Table 4.16.: Determined rate constants for urea refolding experiments of different foldon variants using a global fit by KinTek^{165,166}. The foldon wild-type rate constants were taken from¹⁵⁰. The fit was performed using the "classic" foldon folding model lacking the last rearrangement step from T to N , given in Figure 3.2B. No deviations between a run with fixed and free dimerization rate constant k_D were observed, thus only one set of rate constants is shown (K. Stecher, personal communication). The $\Delta \Delta G^0$ -values are calculated with respect to the foldon wild-type stability.

	fixed k_D			free k_D		
	k_T	k_{-D}	k_T/k_{-D}	k_T	k_{-D}	k_T/k_{-D}
	$M^{-1} \cdot s^{-1}$	s^{-1}	M^{-1}	$M^{-1} \cdot s^{-1}$	s^{-1}	M^{-1}
P4T	$5.4 \cdot 10^6$	32	$1.7 \cdot 10^5$	$4.2 \cdot 10^6$	50	$8.4 \cdot 10^5$
A6Abu	$5.6 \cdot 10^6$	319	$1.8 \cdot 10^4$	$2.0 \cdot 10^6$	206	$9.7 \cdot 10^3$
A6S	$7.6 \cdot 10^5$	40	$1.9 \cdot 10^4$	$8.0 \cdot 10^5$	42	$1.9 \cdot 10^4$
A6V	$8.6 \cdot 10^4$	11	$7.8 \cdot 10^3$	$8.4 \cdot 10^4$	5	$1.7 \cdot 10^4$
P7Nva	$8.7 \cdot 10^4$	35	$2.5 \cdot 10^3$	$8.8 \cdot 10^4$	35	$2.5 \cdot 10^3$
Y13F	$7.2 \cdot 10^5$	731	$9.8 \cdot 10^2$	$2.5 \cdot 10^5$	257	$9.7 \cdot 10^2$
L22F	$1.6 \cdot 10^5$	85	$1.9 \cdot 10^3$	$1.2 \cdot 10^5$	29	$4.1 \cdot 10^3$
P4T P7Nva				$3.7 \cdot 10^5$	405	$9.1 \cdot 10^2$
wild-type				$5.4 \cdot 10^6$	59	$9.2 \cdot 10^4$

Table 4.17.: The k_T/k_{-D} ratio of GdmCl refolding global fits using both fixed and free k_D -values.

	fixed k_D			free k_D		
	k_T	k_{-D}	k_T/k_{-D}	k_T	k_{-D}	k_T/k_{-D}
	$M^{-1} \cdot s^{-1}$	s^{-1}	M^{-1}	$M^{-1} \cdot s^{-1}$	s^{-1}	M^{-1}
P4T	$6.8 \cdot 10^5$	45	$1.5 \cdot 10^4$	$6.7 \cdot 10^5$	32	$2.0 \cdot 10^4$
A6Abu	$1.0 \cdot 10^6$	123	$8.1 \cdot 10^3$	$1.0 \cdot 10^6$	125	$8.2 \cdot 10^3$
A6S	$1.0 \cdot 10^6$	205	$5.0 \cdot 10^3$	$1.0 \cdot 10^6$	218	$4.8 \cdot 10^3$
P7Nva	$2.2 \cdot 10^4$	11	$1.9 \cdot 10^3$	$2.0 \cdot 10^4$	11	$1.8 \cdot 10^3$
Y13F	$1.0 \cdot 10^4$	46	$2.2 \cdot 10^2$	$7.0 \cdot 10^3$	33	$2.1 \cdot 10^2$
L22F	$9.6 \cdot 10^4$	17	$5.6 \cdot 10^3$	$7.0 \cdot 10^4$	16	$4.3 \cdot 10^3$
wild-type				$4.5 \cdot 10^5$	36	$1.3 \cdot 10^4$

Table 4.18.: The k_T/k_{-D} ratio of urea refolding global fits using both fixed and free k_D -values.

or dissociation of the dimer due to the strong coupling of these two rate constants. The ratios determined in urea and GdmCl show a high resemblance, indicating that both the unfolding of the dimer and the formation of the trimer are equally effected by the change of the ionic strength. The only exception is foldon P4T.

5. Discussion

Foldon shows a high structural robustness upon mutations

A large amount of structural data had been known from NMR measurements for the fibrin foldon domain, starting from truncated fibrin variants of varying length¹⁴, NMR data for the native state, the acid-unfolded A-state and the monomeric E5R variant^{95,149,151}. In this study crystal structures of 10 different foldon variants were obtained. Despite some conformational change at the N-terminal end due to the missing coiled-coil segments of fibrin⁹⁵, all structures except the acid-denatured A-state show a remarkable high structural similarity, with backbone root mean square deviations of approximately 0.3 Å, regardless of the used structure resolution method and variant. The large structural resemblance of the different foldon variants is however accompanied by a large change in stability of the native trimer by more than 50 kJ · mol⁻¹ for the most destabilizing mutant. Although this finding is astounding at first sight, a high elasticity of protein structures to mutations is a common observation. The most prominent example is T4 lysozyme, with its currently almost 600 crystal structures deposited in the RCSB Protein Data Bank¹⁷⁰: Although all structures, regardless of inserted mutations, are almost similar in their structure, the stability determined by thermal transitions can both be significantly higher or lower¹⁷¹. Thus, high robustness and compensation of unfavorable interactions is a common feature found in many proteins. Furthermore, when discussing the strong effects of single-point mutations in foldon, one always has to keep in mind that each mutation is present three times in the native trimer, thus the effect on stability is triplefold. The stabilizing effect of one exchange of proline at position 4 to a threonine is only 4.1 kJ · mol⁻¹, which corresponds to the stability gained by a newly formed hydrogen bond¹⁷². The strongest destabilizing mutation of the proline at position 7 to a phenylalanine shows a $\Delta\Delta G^0 = 42.7$ kJ · mol⁻¹ for the trimer. However, the destabilization per foldon monomer is only 14.2 kJ · mol⁻¹, which corresponds to commonly observed destabilizations for single-point mutations^{173,174}.

5. Discussion

No crystal structures could be resolved for the foldon variants A6V and P7V, as no suitable crystals could be grown. It is striking that both mutations that insert a valine into the hydrophobic cluster hinder crystallization of foldon completely. Although both valine mutants are among the most destabilized mutants, stability alone cannot be the sole criterion, as the variants at position 20, which destabilize the native state even more, are able to crystallize. Both mutations incorporate a valine at a critical location in foldon. Mutations at position 6 lead to a change of the overall reaction mechanism, making dimerization of monomers with a *cis* conformation of the Xaa6-Pro7 peptide bond unproductive for the assembly pathway. The valine side chain probably hinders the N-terminal region to adopt a backbone conformation that allows near-native packing of this region against the β -hairpin. Although foldon A6V is still able to fold in GdmCl refolding experiments, the unfavorable conformation of the N-terminal region seems to have a stronger inhibiting effect on crystallization than low stability, as a low stability can be overcome by higher monomer concentrations or adjustments of the crystallization conditions. The reasons why it is impossible to obtain crystals of the foldon P7V variant are likely very similar to the ones for foldon A6V. It was further shown by Eckhardt et al. that the incapability to obtain crystals and solve the crystal structure is not necessarily accompanied by the inability to fold into the native trimer¹⁷⁵.

Mutations inserted into foldon can lead to local perturbations of the side chain conformations, but do not have an impact on the overall backbone conformation. This is seen from the high resemblance of the trimer structure. Despite the fact that mutations within the hydrophobic cluster do not have any severe effect on the native structure, the effects on the fraction of natively formed trimers are strong. All foldon variants except foldon P4T show a strong destabilization upon insertion of a mutation into the hydrophobic cluster by at least $17 \text{ kJ} \cdot \text{mol}^{-1}$. This destabilization corresponds to a shift in the $N \rightleftharpoons U$ equilibrium towards U by a factor of 1,000 at 1 M foldon monomer concentration. At a physiological foldon monomer concentration of $5 \mu\text{M}$, the fraction of N at 293.15 K is shifted from 0.988 for foldon wild-type to 0.885 for foldon A6Abu, the least destabilized foldon variant. A decrease to 0.014 at $5 \mu\text{M}$ is observed for the least stable foldon variant P7F. The m_{eq} -value of all variants is higher than the value for foldon wild-type, representing a higher susceptibility of all foldon variants to denaturing agents, and additionally representing a higher change in accessible surface area upon the transition. For some variants, the start of a second transition for high GdmCl concentrations is seen. This could be interpreted that the first transition is only able to induce the dissociation of the native foldon into monomers, and that there is still some residual structure present in the

monomers. This residual structure could be the β -turn motif, because a strong resistance of the β -turn against denaturation is seen in the acid-denatured A-state NMR structures¹⁴⁹. This residual structure may still be present in the denaturant-unfolded state of foldon wild-type, leading to the second GdmCl transition. The insertion of the different mutants could however induce an unfolding of this structure at high denaturant concentrations, making the unfolded state of the variants more expanded and thus more accessible for denaturing agents, finally yielding a higher m_{eq} -value. This ground-state effect is commonly observed, and is very often caused by a change of the m_f -value while the m_u -value is almost unchanged⁷⁴. This effect will be discussed in detail below.

Mutations can in- or decrease the unfolding rate constant of the native trimer

A shift of the equilibrium towards the unfolded state of a protein unfolding-refolding equilibrium is often accompanied by an increase of the unfolding rate constant. This increase is found for most foldon variants, with dissociation rate constants of the native trimer k_{-T} in the range of 10^{-3} to 10^{-2} in GdmCl and 10^{-4} to 10^{-2} in urea. The increase of k_{-T} can however be as large as five orders of magnitude compared to foldon wild-type. A k_{-T} of $(0.21 \pm 0.06) \text{ s}^{-1}$ is found for foldon W20NaI in the presence of 500 mM sodium sulfate. Due to the sodium sulfate, this unfolding rate constant is not directly comparable to the other foldon variants. However, in three mutants the shift of the equilibrium towards the unfolded state is not caused by an increased unfolding rate constant k_{-T} . In these cases, the unfolding rate constant is lower than the one of foldon wild-type. These are the foldon variants P4T P7Nva, foldon L22F and foldon W20H. The rate constant of foldon W20H, as the one of foldon W20NaI, is determined in the presence of 500 mM sodium sulfate, thus comparability is constricted. It is striking that the other two mutations found to cause decelerated unfolding are located at the most N- or C-terminal positions of the hydrophobic cluster.

According to the Arrhenius equation, rate constants are generally dependent the temperature and the free activation enthalpy²⁸. As all experiments are performed at $T = 293.15 \text{ K}$, only the free activation enthalpy and pre-exponential factor can cause changes of the rate constant. A decrease of the pre-exponential factor would lead to overall lower rate constants. However, as the change introduced into the peptide by a single-point mutation is small, this effect can most likely be neglected. The inserted mutations could, on the other hand, cause an increase of the free activation enthalpy. The increase of the free activation enthalpy might be due to

decreased or weakened side chain interactions in the transition state. A possible explanation for this finding is that foldon wild-type is not optimized for a high unfolding free activation enthalpy. The moment the native trimeric structure is formed in fibrin, coiled-coil formation is initiated. Thus foldon has served its purpose, and it is irrelevant for the overall stability of fibrin if foldon can undergo local unfolding of its C-terminal domain. Secondly, a decrease of the free activation enthalpy caused by increased side chain interactions could also result in either an increased propensity of non-native secondary structure formation during the folding and assembly process, or a stabilization of the intermediate compared to both the unfolded and native state. Both cases would strongly perturb folding, as the formed structures would have to be broken prior to folding. A similar process is observed for a single-chain foldon variant, a protein where three foldon sequences are connected by flexible (GS)₇-loops. Designed to allow a concentration-independent determination of the association rate constants, it was shown that folding can only occur rapidly at a low GdmCl concentration, as a partial unfolding of non-native structure motifs seems necessary to allow correct formation of the native trimeric structure¹⁵⁰. The foldon domain is not optimized for slowest possible unfolding or highest possible stability, as seen from the foldon P4T variant, but for fast association and rapid contact formation of its native interactions.

Intra-molecular hydrogen-bonds in foldon

The hydrophobic cluster of foldon is not only stabilized by van der Waals interactions, but also by two intramolecular hydrogen bonds. One bond is formed between the backbone oxygen of the proline 4 and the N₁ of the indole of tryptophan 20. This interaction could assure correct packing of the N-terminal region against the β -turn during the folding reaction. The second hydrogen bond in the hydrophobic cluster is formed between the tyrosine 13 side chain hydroxyl group and the aspartate 9 and glutamine 11 side chains in the first type-II-turn. These amino acids constitute a hydrogen bond network between the hydrophobic cluster and this type-II-turn leading to the first β -strand, as shown in Figure 5.1B. This hydrogen bond network is crucial for the folding and assembly of the native trimer. Eckhardt et al. showed that a disruption of this network by a single-point mutation of the aspartate 9 to a glutamine disrupts cooperative folding into the native trimer¹⁷⁵. The effect of hydrogen bonds within the hydrophobic cluster on the stability and folding of the native structure is thereby of interest.

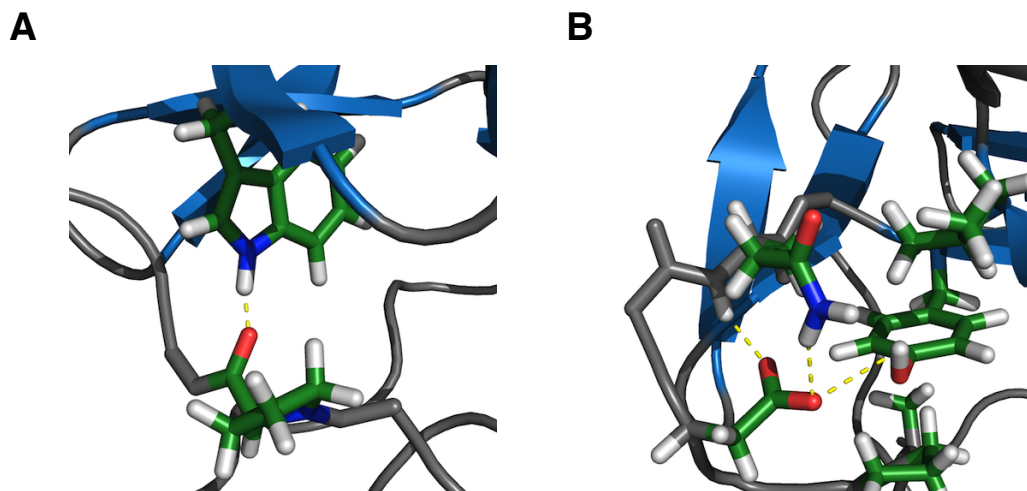


Figure 5.1.: Hydrogen bonds in the hydrophobic cluster. A) The hydrogen bond between the proline 4 backbone oxygen and the hydrogen of the tryptophan N_1 of the indole side group. The side chains are shown in stick representation and the atoms colored in green for carbon, red for oxygen, blue for nitrogen and white for hydrogen, while the backbone of the loops is colored in gray and the β -strands are in blue. The hydrogen bond is depicted as a dashed yellow line. B) The hydrogen bonding network between the tyrosine 13 and the aspartate 9 and glutamine 11. The same color scheme is used as in A).

In general, the formation of hydrogen bonds is one of the driving forces of the folding reaction. The enthalpy of the formation of hydrogen bonds between a protonated nitrogen and an oxygen, the most common hydrogen bond formed in proteins, is between -12 to $-40 \text{ kJ} \cdot \text{mol}^{-1}$ per hydrogen bond¹⁷⁶. This type of hydrogen bond is predominately found in secondary structure motifs like α -helices and β -pleated sheets. The vast majority of hydrogen bonds in proteins is formed within single elements of secondary structure²⁵. Removal of a hydrogen bond by mutation of the respective amino acid however only results in a commonly observed decrease of stability of 2 to 7 $\text{kJ} \cdot \text{mol}^{-1}$ per hydrogen bond¹⁷². One reason for this deviation is that the protein can also form hydrogen bonds with the solvent, though these are not as stable as hydrogen bonds within a hydrophobic environment¹⁷⁶. Additionally, binding of the solvent water to the protein via hydrogen bonds causes a loss of solvent entropy, further raising the free enthalpy²⁷.

The energetic contribution and the influence on the association reaction of two hydrogen bonds in the hydrophobic cluster are determined in this study. Removal of the hydrogen bond between the backbone carbonyl group of proline 4 and the N_1 of the tryptophan 20 indole group by an exchange of the tryptophan to a naphthylalanine results in a change of the free folding

5. Discussion

enthalpy of $\Delta\Delta G^0 = 41.8 \text{ kJ} \cdot \text{mol}^{-1}$ compared to foldon wild-type. The only difference observed for the crystal structure of foldon W20Nal variant with regard to structure of foldon wild-type is the loss of this hydrogen bond. The orientation of the naphthyl group and the indole group in the hydrophobic core are alike, as can be seen in Figure 4.34. Neither are there any other interactions of the naphthyl group within the hydrophobic cluster, nor is the rest of the foldon domain perturbed. Therefore this mutation gives a grand estimate for the contribution of the hydrogen bond of $13.9 \text{ kJ} \cdot \text{mol}^{-1}$ to the free folding enthalpy of foldon. This value is higher than the stability normally found for hydrogen bonds. Nevertheless, contributions in a similar range have been observed before¹⁷⁷. Is there an explanation for this large effect? The large contribution of the hydrogen bond is assigned to the hydrophobic environment of the hydrogen bond. The test whether a restoration of the hydrogen bonds by insertion of a histidine at position 20 would again lead to a stabilization of the native structure partially failed, as the loss of hydrophobic interactions within the hydrophobic core and the probable inclusion of water in the formed cavern overcompensate for the gained energy of the formed hydrogen bonds between the histidine's N' and the residue 4 backbone carbonyl group. The effect of the hydrogen bond in this W20H mutant is however not negligible, as shown by the pH-titrations of the histidine. This titration shows an increase of stability upon changing the pH from 7.0 to 8.0, while folding into the native trimer is completely omitted at pH 5.5. This effect was also shown before for villin head-piece^{178,179}.

The second hydrogen bond disrupted by an amino acid exchange in the hydrophobic cluster is the one between the tyrosine hydroxyl group and the side chain of the aspartate at position 9. Although the amino acid exchange of the tyrosine to a phenylalanine leads to a decrease of the free folding enthalpy of $\Delta\Delta G^0 = 24.8 \text{ kJ} \cdot \text{mol}^{-1}$, foldon is still able to fold into the native structure, in contrast to the results of Eckhardt et al. for different mutations¹⁷⁵. The change in free folding enthalpy cannot be attributed completely to the loss of the hydrogen bond. First, the hydrogen bonding network in the first β -turn compensates by forming a new hydrogen bond between the nitrogen of the glutamine at position 11 and the aspartate at position 9. Second, the removal of the hydroxyl group leads to a change of the foldon surface, as the former hydrophilic patch formed by the tyrosine's hydroxyl group is now replaced by a hydrophobic surface from the phenyl group of the phenylalanine. A clear attribution of the energetic contribution of this hydrogen bond is hence not possible. The role of this hydrogen bond can rather be seen as a link of the hydrophobic cluster to the first type-II-turn. Despite the fact that a removal of the two hydrogen bonds located within the hydrophobic cluster does not prevent

the folding and association into the native state, the loss of energy leads to a strong shift of the equilibrium towards the unfolded state. Both hydrogen bonds have therefore a vital contribution to the overall stability and probably help finding the correct native-like interactions during the folding reaction.

The function of the N-terminal poly-Pro II structure

The N-terminal region of foldon is a poly-proline (poly-Pro) type II helix. These proline-rich motifs are commonly found in many organisms and make up a large fraction of all sequence motifs¹⁸⁰. The poly-Pro type II helix is left-handed and has three residues per turn with an overall shape which is similar to a triangular prism^{181,182}. Poly-Pro II helices are a favorable binding partner for recognition domains¹⁸³. All backbone carbonyl groups in the poly-Pro II structure are exposed to the solvent due to the extended conformation, and the amino acids are ordered in a highly repetitive manner¹⁸⁴. The backbone conformation in a poly-Pro II helix is restricted, thus the binding of the backbone is not accompanied by a negative entropic term¹⁸⁵. Any binding domains can thus interact favorable with the poly-Pro II helix. There are numerous protein domains known that are able to bind poly-Pro II helices^{186–189}. Each of these domains binds to the poly-Pro II helix by the exposure of aromatic side chains to the solvent, predominately tryptophan and tyrosine residues. The poly-Pro II helix is additionally not only found in the native state of proteins, but also in the unfolded state. The extended conformation allows hydrogen bond formation of the backbone carbonyl group with the solvent, regardless of the proline content^{190,191}.

The role of the poly-Pro II structure in the N-terminal region of foldon is probed by both single-point mutations of the two prolines at position 4 and 7 and by a double-mutation of both positions. An increase in stability of $12.3 \text{ kJ} \cdot \text{mol}^{-1}$ of foldon P4T compared to foldon wild-type does not lead to a significant increase in the refolding kinetics in both GdmCl and urea. The hydrophobic cluster is only slightly perturbed in the burst-phase intermediate, as seen from the decreased fluorescence of the burst-phase intermediate of foldon P4T, but nevertheless can still induce correct and fast association of the foldon monomers into the native trimer. The local structure around residue 4 is possibly preserved by the conserved intra- and intermolecular interactions of the surrounding residues of the foldon N-terminal region. The peptide bond between the residues 3 and 4 is in the *trans* conformation in the native state of both foldon P4T and foldon wild-type, but the isomerization of this peptide bond does not

5. Discussion

have an impact on the overall refolding kinetics. This is supported by the finding of Habazettl et al. where 10-14 % of the foldon monomers are in the *cis* conformation at the Ile3-Pro4 peptide bond¹⁵¹. However, it cannot be ruled out that in fibrin the proline 4 has to be in the *trans* conformation to allow the formation of the fibrin coiled-coil structure. Loss of the rigid backbone conformation by a mutation of the proline 7 has a large effect on the association reaction and stability. The proline 7 in each monomer has to be in the *trans* conformation to form the native foldon trimer. Formation of the dimeric and trimeric intermediate are nevertheless possible even if the proline 7 residue of one or more monomers is in the *cis* conformation. This indicates that the N-terminal region does not influence the association reaction by direct interactions during the association, what is expected, as the association surfaces are located on the other side of the foldon monomer.

The double-mutant P4T P7Nva is destabilized by $19.0 \text{ kJ} \cdot \text{mol}^{-1}$ compared to foldon wild-type, what is, within error range, the sum of the effects of the single mutants on the folding free enthalpy. The double-mutant's refolding rate constants are further decreased by one order of magnitude compared to foldon P7Nva. Though proline 4 is not responsible for the observable *cis-trans* isomerization, its rigidifying effect on the N-terminal poly-Pro II structure is necessary for efficient association during the folding process. The importance of the prolines in the N-terminal region of foldon is possibly not the stabilization of the poly-Pro II structure in the folded, but rather in the unfolded and the intermediate state to allow for a quick formation of the hydrophobic core and efficient dimer and trimer formation. With three of the six amino acids of the hydrophobic cluster located within this region, it is a major nucleation point for the folding of the monomer. If the N-terminal region has a native-like conformation in the unfolded state, which is induced by the two proline residues, the N-terminal residue could act as a scaffold for the initial folding reaction of the unfolded state towards the burst-phase intermediate. This could explain the extremely fast folding into the monomeric structure of the burst-phase intermediate^{95,151}. With both prolines missing in the P4T P7Nva mutant, the gained conformational freedom of the N-terminus could lead to a less-prominent poly-Pro II structure. This would lead to both a decreased rate constant of burst-phase intermediate formation and result in a loss of the scaffold for efficient dimerization. Two monomers could effectively form an encounter complex even if the monomers would not meet at the dimer binding interfaces. The mechanism could be a temporary and weak binding of the poly-Pro II structure with one of the exposed aromatic side chains of the other foldon monomer. The poly-proline II structure would thus act as an additional temporary binding interface during

association. The sharp peak observed in the fluorescence spectrum for the burst-phase intermediate of the foldon P4T P7Nva variant supports this hypothesis. The peak could be caused by a tremendously changed burst-phase intermediate structure, which causes a change of the local environment of the fluorescent tryptophan. The increased fluorescence of tryptophan could arise from a rigid environment of the indole side chain group. The loss of the poly-Pro II structure could explain the low dimerization rate constant of foldon P4T P7Nva by the lowered encounter complex formation probability. Upon folding into the native structure, the closed packing is disrupted and the fluorescence intensity is decreased again. This proposed mechanism is one way of the hydrophobic cluster to influence the association reaction.

Dimerization is influenced strongly by the hydrophobic cluster

The largest impact on the individual folding and association steps of foldon refolding caused by the inserted mutations is seen on the dimerization rate constant, regardless of the used denaturant. The dimerization rate constant is lowered by up to four orders of magnitude. Dimerization is the slowest step of the association reaction in foldon wild-type, with the only exception for large foldon monomer concentrations. The impact of the mutations on the dimerization rate constant is strongly dependent on the amino acid position. An exchange at position 4 has a negligible effect on dimerization in GdmCl, while the largest detectable impact of a single-point mutation is detected at position 6. The A6V variant shows the lowest dimerization rate constant, with $k_D = 7.3 \cdot 10^3 \text{ M}^{-1} \text{ s}^{-1}$. The other two mutants at position 6 are only reduced by one order of magnitude. An additional effect is caused by the change of the reaction mechanism by the mutations at position 6. All other single-point variants of foldon show a decrease of the dimerization rate constant by two orders of magnitude. All refolding traces determined in GdmCl strongly resemble the foldon wild-type refolding traces, despite the strong effects on dimerization. Although these results strongly point towards dimer formation as the crucial step to be influenced by the mutations, a clear result is only obtained by measurements in urea. Any effects on the dimerization rate constant are even more prominent in urea, as the dimer is strongly stabilized.

The effects on dimerization in urea differ from the effects in GdmCl. The characteristic refolding traces of foldon wild-type in urea with the separation of the dimer and trimer formation are not detected for any foldon variant. This separation of the kinetic phases of foldon wild-type is caused by the fact that the dimerization rate constant is larger than the trimerization

5. Discussion

rate constant, resulting in a stabilization of the dimeric intermediate under refolding conditions in urea¹⁵⁰. This separation of the dimerization and trimerization step of foldon cannot be observed for any foldon variant in urea. Even foldon P4T, which shows an almost identical behavior to foldon wild-type when refolded in GdmCl, shows a different shape of the refolding traces. Dimerization of foldon P4T is decreased by a factor of five compared to foldon wild-type at identical foldon monomer concentrations, while the trimerization rate constant is larger in foldon P4T than in foldon wild-type ($6.8 \cdot 10^5 \text{ M}^{-1} \text{ s}^{-1}$ in P4T compared to $4.5 \cdot 10^5 \text{ M}^{-1} \text{ s}^{-1}$ in foldon wild-type). This effect is even enlarged for the variants at position 6 studied in urea, where the trimerization rate constant is increased by a factor of 2.5 compared to foldon wild-type for both mutants. Dimerization is decelerated by a factor of 20 for these mutants, and the stability of the dimeric intermediate is further reduced by an increase of the dissociation rate constant k_{-D} by a factor of three for foldon A6Abu and a factor of six for foldon A6S. The trimerization rate constant k_T and dimer dissociation rate constant k_{-D} are strongly coupled in the evaluation by a global fit of all refolding traces in urea. Therefore, absolute numbers cannot be determined for these rate constants with high fidelity, but the tendency for a higher trimerization rate constant in contrast to the dimerization rate constant is clearly observable from the fits. Otherwise, the overall shape of the fit traces would resemble the ones of foldon wild-type in urea. The lowest dimerization rate constant in urea is observed for foldon P7Nva, which is almost a factor of 10,000 lower than in foldon wild-type. The dimerization rates of foldon Y13F and L22F are very similar to each other and decreased by three orders of magnitude compared to foldon wild-type. Insertion of any mutation into the hydrophobic cluster leads to a stronger destabilization of the dimeric intermediate in urea, compared to foldon wild-type.

Normally, effects on the dimerization of proteins are observed when amino acids within the dimer binding interface are mutated, while mutations in other regions show little to no effect on the association¹¹³. While this is true for the foldon variants at position 20, where the refolding kinetics cannot be determined, all other amino acids of the hydrophobic cluster are not located within the dimer interface (with the exception of one alanine at position 6). The hydrophobic cluster thus influences dimerization by other means. The fluorescence spectra for almost all foldon variants of the burst-phase intermediates show a decrease of the fluorescence intensity. A change of the local structure around the fluorophore of the burst-phase intermediate could induce this change of the fluorescence properties. Changes in structure could strongly affect the ability of the foldon monomers to dimerize, as the intermediate structure would become less defined, and thus partially disrupts the association interfaces. This would likely be accompa-

nied by a decrease of the burst-phase intermediate concentration resulting in a shift of the fast equilibrium between the unfolded state U and the burst-phase intermediate I . A shift towards the unfolded state U reduces the active concentration of monomers which can productively form dimers. Similar effects had been observed for several other proteins¹¹⁶⁻¹²⁰, and changes of the fraction of preformed structure can either increase or decrease the association rate constant by up to two orders of magnitude^{116,117}. Despite the good agreement of this assumption with the determined data of most foldon variants, the foldon variants A6Abu and A6S seem to contradict this proposed mechanism. The burst-phase intermediate fractions of these foldon variant seems to be unchanged when considering the burst-phase intermediate fluorescence spectra, and yet a decrease of the dimerization rate constant is observed for these variants. Thus a change of the intermediate structure is not necessarily accompanied by a change of the local environment of the tryptophan. Additionally, the change of the reaction mechanism further decreases the observed speed of the association reaction. Due to the *cis* conformation of the Xaa6-Pro7 peptide bond, 10 % of the burst-phase intermediate cannot dimerize any more. This reduces the concentration of monomers that can productively form dimers, resulting in a decrease of the association speed. While the effect of the inserted mutations on dimerization is strong in both denaturants, only marginal influences on the trimerization reaction occur. The largest effect is observed for foldon P7Nva and Y13F in urea and foldon A6V and P7Nva in GdmCl, where the trimerization rate constant is decreased by approximately two orders of magnitude, compared to one order of magnitude determined for all other foldon variants. After dimerization the dimer offers a well-defined binding surface for the last monomer. This could compensate for the less defined intermediate structure, thus the effect on the second binding step would be less pronounced. Additionally, binding of the last monomer could be preceded by an induced transition of the unfolded state to the burst-phase intermediate state, as observed for intrinsically unfolded proteins or the S-protein/S-peptide system^{114,115}. Binding of the unstructured monomer could be induced by the formed poly-proline II structure to the exposed aromatic side chains, as described before, followed by folding of the unfolded monomer and association into the trimeric intermediate.

Transition state

Mutations within the hydrophobic cluster could not only change the rate constants, but also the position of the transition state on the reaction coordinate. Any changes of the transition state are determined by rate-equilibrium free energy relationships (REFERs). Any shift of the transition state location can be probed by the determination of the α_D -values in both GdmCl and urea. It is seen for both denaturant agents that no significant shift of the transition state is observable, with the only exception of foldon P4T P7Nva. In this variant the α_D -value is shifted from approximately 0.6 to 0.91, indicating a strong shift of the transition state towards the native structure. This indicates a high compaction of the transition state, accompanied by a low susceptibility to the denaturant. A high similarity of the transition state to the native structure can thus be assumed. The lack of the stabilizing poly-proline II structure in the intermediate necessitate many more native-like interactions in the transition state to allow the formation of the native structure, thus shifting the transition state position on the reaction coordinate towards the folded state.

For all foldon variants, an increase of the m_{eq} -value in both denaturants is observed. This increase can be attributed to the ground-state effect⁷⁴: Most foldon variants show an increase of the m_f -value compared to foldon wild-type while the m_u -value is almost unchanged. There are only two exceptions in GdmCl: The m_u -value of foldon L22F is decreased by approximately 30 % and the one of foldon P4T P7NVa decreases by 70 %, resulting in a significantly increased α_D for these variants. However, these variants also show a significant decrease of the unfolding rate constant k_{-N} despite a destabilizing effect of the two mutants on the native state. Thus it is obvious that the effects of the amino acid exchanges on stability are caused by changes of the kinetics.

Another commonly used probe to detect linear REFERs is the effect of mutations on both the free activation enthalpy and the free folding enthalpy with respect to the wild-type protein. This leads to the α_S or ϕ_f -value¹⁷³. It is defined by:

$$\phi_f = \frac{\Delta\Delta G^{0\ddagger}}{\Delta\Delta G^0} \quad (5.1)$$

In general, a ϕ_f -value of 1 states that the effect of the mutation on the transition state is the same as the effect on the native state. The amino acid thus has native-like interactions in the transition state. A ϕ_f -value of 0 states that the mutation has no effect on the transition

state, thus no native-like interactions of the amino acid are present in the transition state. This method probes for folding nuclei, as these should show high ϕ_f -values. In general, ϕ_f -values between 0 and 0.4 are determined for most amino acids and most positions in almost all studied proteins¹⁹². ϕ_f -values larger than 0.4 are uncommon and interpreted as nucleation sites for the folding reaction, as the interactions of these amino acids in the transition state are larger than average¹⁷³. Reliable ϕ_f -value determination requires several mutants at the same position. It was shown that ϕ_f -values lower than 0 or higher than 1 are often obtained if only a two-point analysis with the wild-type and one mutant is performed, as evaluation of only two data points leads to a large error¹⁹³. This becomes even more prominent if the stability change upon mutation is less than 6-7 kJ · mol⁻¹¹⁹⁴. The ϕ_f -values of the foldon variants are not determined in this study as the refolding reactions were not probed in a denaturant-dependent manner and more data points would be needed to obtain a robust fit for the denaturant-dependent unfolding experiments. Furthermore, a ϕ_f -value analysis requires in most cases a global k_f -value, which cannot be determined for any foldon variant due to the missing phase of the burst-phase intermediate formation. It is on the other hand obvious from the measured data that the foldon P4T variant has a non-canonical ϕ_f -value. The slight increase of the free activation enthalpy, causing the slightly decelerated refolding kinetics, is accompanied by a stabilization of the native state. According to equation 5.1, this leads to a $\phi_f < 0$. A ϕ_f -value equal to zero is seen as a measure that no native state interactions of the amino acid side chain are present in the transition state. The pyrrolidine group of Pro4 does however only have contacts with the solvent, therefore a value around zero is not surprising. A ϕ_f -value below zero indicates that the amino acid forms non-native interactions in the transition state. In case of residue 4 of foldon, any interaction of the side chain with other side chains of the protein in the transition state is a non-native interaction, thereby causing the change of the ϕ_f -value. If this side chain is able to make contact to other regions of foldon, a significantly changed conformation of the N-terminal region must be present in the transition state to allow an orientation of the residue 4 side chain towards the core of the protein. In the native state, as seen from the crystal structure, the threonine side chain of the P4T mutant points into the solvent and does not have any contact to other parts of the folded protein. The non-native interactions formed during the refolding reaction cause a deceleration of the association reaction, as they have to be broken prior to completion of the folding reaction. This explains the deceleration of foldon P4T refolding despite a higher stability of this foldon variant.

Outlook and conclusions

Future experiments on the hydrophobic cluster of foldon could include additional mutations at the positions 4, 13 and 22. These could be insertions of large, bulky amino acids at position 4, e.g. tyrosine or phenylalanine. An incorporation of isoleucine at position 22 would show if the hydrophobic interactions between the residues 13 and 22 were preserved. To probe the effect of the hydrophobic interactions of the tyrosine at position 13, insertion of the non-natural amino acid 6-hydroxy-(L)-norleucine should be feasible. This amino acid lacks the aromatic ring of tyrosine, but should be able to locate the hydroxyl group in a similar position as the tyrosine, thus sustaining the hydrogen bonding network with the first type-II-turn. To gain a better understanding of the burst-phase intermediate and the proposed rapid equilibrium of this intermediate and the unfolded state, NMR structure determination of foldon variants in the monomeric form could be a feasible tool. As all peaks in a ^{15}N -edited HSQC spectrum of the respective foldon variants could be compared to the already assigned peaks of the foldon E5R, the changes of peak intensity should allow a determination of the fraction of formed monomeric structure. The structure of the foldon P4T P7Nva with the additional E5R mutation would further allow the determination of any formed secondary structure of the N-terminal region of the foldon monomer which could account for the changed fluorescence properties of the burst-phase intermediate and the strongly decreased association kinetics. The determination of the free activation enthalpy of both unfolding and refolding of several foldon variants, especially foldon L22F and foldon P4T P7Nva would lead to a ϕ -value analysis and a more detailed characterization of the transition state. This would allow a detailed evaluation of the non-intuitive effect on unfolding observed by the before-mentioned foldon variants.

Monod, Wyman and Changeux stated already in the 1960s when proposing their allosteric model that homo-oligomeric proteins are subject to a higher evolutionary pressure¹⁹⁵. Each mutation within a homo-oligomeric protein complex is amplified by the number of subunits, as is every deletion or addition of amino acids. In most cases, even small modifications will lead to a ruinous effect on the overall stability and catalytic activity of a protein, while some rare events will lead to a large beneficial change in the functional characteristics. The same phenomenon is observed throughout this study: Almost any mutation in foldon, which can be seen as a catalyst for the fibritin association, has tremendous effects on both stability and folding. It seems that all advantageous mutations have already been applied to foldon by evolution, making foldon an almost perfect folding and association domain.

6. Summary

The foldon domain of bacteriophage T4 fibritin is an excellent model system to study the folding and association of small multimeric proteins. The free folding stability of the foldon domain was known from denaturant transitions in both GdmCl and urea, and refolding experiments had been carried out for foldon wild-type which could determine the rate constants of the majority of the folding and association steps during the foldon folding reaction. A wealth of structural data had been resolved, showing foldon wild-type in the native trimeric and an acid-denatured state, as well as the foldon E5R variant in a stable monomeric state. Fluorescence spectra of the burst-phase intermediate of foldon wild-type during refolding and foldon E5R under native conditions show a large similarity. The structural data of the monomeric foldon variant revealed the existence of a hydrophobic cluster. Additionally, the foldon E5R indicated towards an association mechanism that involves the charged surfaces of the foldon monomers, which show a clear separation of positive and negative charges on opposing surfaces of the foldon monomer. However, determination of the impact of varying ionic strengths on the refolding kinetics had shown a primary salt effect of a single net charge. This corresponds to the overall net charge of the foldon wild-type domain. Thus hydrophobic interactions seem to influence the association reaction of foldon, as often found for homo-oligomeric proteins. The hydrophobic cluster was hence chosen for further studies on the association mechanism of foldon.

In this study we used mutations within the hydrophobic cluster of foldon to probe the association mechanism. First, the X-ray structures of 10 different foldon variants were determined which show a remarkable similarity with backbone root mean square deviations around 0.3 Å. Although the effect of the mutations on the side chain packing within the hydrophobic cluster is stronger for some foldon variants, this packing effects cannot account for the wide range of stabilities observed for the different foldon variants. The determined stabilities range from $-101.5 \text{ kJ} \cdot \text{mol}^{-1}$ for the only stabilizing mutation foldon P4T to $-46.5 \text{ kJ} \cdot \text{mol}^{-1}$ for foldon P7F. The two mutations at position 20 are so strongly destabilized that the stabilities could

6. Summary

only be determined in the presence of 500 mM sodium sulfate. The difference of the free folding enthalpy compared to foldon wild-type under these conditions is more than $40 \text{ kJ} \cdot \text{mol}^{-1}$.

The fluorescence properties of the different states of foldon during the assembly process are probed prior to the kinetic measurements. The fluorescence spectra of the unfolded and native state as well as the burst-phase intermediate show a high resemblance of their shapes to the foldon wild-type spectra, while the fluorescence intensity of all states is in most cases strongly decreased. For the native state, this is very likely caused by a lower population of natively folded protein compared to foldon wild-type due to the lower stability of the mutant. The reasons are less clear for the unfolded state and the burst-phase intermediate. In the unfolded state, the fluorescence properties could be changed due to the change of residual structure in the region of the fluorescent tryptophan. The same could hold true for the burst-phase intermediate, however, another reason could be a reduced fraction of I due to a shift of the equilibrium towards the unfolded state caused by the inserted mutations. This is supported by the fact that the burst-phase intermediate I shows a stronger loss of fluorescence intensity than the unfolded state fluorescence in the foldon variants P4T, Y13F and L22F.

The refolding rate constants for several foldon variants are determined in both GdmCl and urea by global fitting of refolding traces at varying foldon monomer concentrations. Unfolding experiments and determination of the initial slope increase the fidelity of the fit as these methods allow an independent determination of the dissociation rate constant k_{-T} and the dimerization rate constant k_D . It is obvious from the global fits that the main effect of the mutations is exerted on the dimer association reaction. This rate constant can be decreased by a factor of up to 10,000, while the trimerization reaction is only decreased by one order of magnitude in most cases. The inserted mutations likely cause a weakening of the hydrophobic interaction within the hydrophobic cluster, thereby hindering the association reaction. Additionally, a shift of the equilibrium between the unfolded state U and the burst-phase intermediate I towards the unfolded state causes a decrease of the association reaction speed. For all variants at position 6, the deceleration of the refolding velocity is additionally influenced by the inability of the monomeric states to form dimers even if the Xaa6-Pro7 peptide bond is in the *cis* conformation. This causes a decrease of the active foldon monomer concentration by approximately 10 %, further decreasing the refolding speed.

For foldon P4T, the effect on the association rate constants is only marginal compared to other variants, and the stability of foldon P4T is increased compared to foldon wild-type . On the other hand, mutations at position 7 cause the strongest deceleration observed for foldon variants, and the double mutant at both positions has the slowest dimer association rate constant observed for any foldon variant in this study. The loss of the rigid backbone structure, which is normally prevailed by the prolines, is a likely source for this behavior. The poly-proline type II structure induced by the prolines could act as a scaffold for the folding reaction and even enhance the encounter complex formation. Loss of this structure, perhaps accompanied by the formation of non-native secondary structure, can be the source for the large impacts these mutations have on the association reaction. The strong effects on the burst-phase intermediate fluorescence spectra of foldon P4T P7Nva strongly point in this direction.

A. Appendix

A.1. Crystal structure data

	resolution (Å)	r.m.s.d.s (Å)		resolution (Å)	r.m.s.d.s (Å)
P4T	1.10	0.29			
A6S	0.90	0.30	A6Abu	1.10	0.33
P7F	1.20	0.34	P7Nva	1.10	0.28
P4T P7Nva	1.10	0.31			
Y13F	1.00	0.30			
W20H	0.90	0.34	W20Nal	0.90	0.33
L22F	1.10	0.29			
wild-type	1.40	-			

Table A.1.: Resolutions and backbone r.m.s.d.s compared to the foldon wild-type crystal structure of all foldon crystal structures.

A. Appendix

	Foldon P4T	Foldon A6Abu	Foldon A6S
Crystal parameter			
Space group	P6 ₃	P2 ₁ 3	P6 ₃
Cell dimensions			
<i>a</i> , <i>b</i> , <i>c</i> (Å)	27.8; 27.8; 63.9	46.3; 46.3; 46.3	28.0; 28.0; 63.9
α , β , γ (°)	90.0; 90.0; 120.0	90.0; 90.0; 90.0	90.0; 90.0; 120.0
Molecules per AU	1	1	1
Data collection			
Beam line	CuK α	CuK α	CuK α
Wavelength (Å)	1.5418	1.5418	1.5418
Resolution range (Å) ^b	64.0 - 1.1 (1.2 - 1.1)	46.3 - 1.1 (1.2 - 1.1)	63.9 - 0.9 (1.0 - 0.9)
Unique reflections ^c	11,079	12,952	19,621
Completeness (%) ^b	99.60 (95.33)	99.99 (100.00)	99.49 (93.09)
R_{merge} (%) ^{b,d}	7.1 (29.0)	6.0 (65.0)	4.9 (25.1)
$I / \sigma(I)$ ^b	22.2 (4.9)	23.7 (2.0)	21.2 (4.2)
Refinement			
Resolution (Å)	10.0 - 1.1	10.0 - 1.1	10.0 - 0.9
R_{work} / R_{free} ^e	0.165 / 0.176	0.134 / 0.174	0.135 / 0.160
No. atoms			
Protein	218	221	220
Waters	29	58	66
<i>B</i> -factors	7.897	12.965	9.440
R.m.s. deviations ^f			
Bond lengths (Å)	0.026	0.027	0.027
Bond angles (°)	2.048	2.017	2.475
Ramachandran (%) ^g	96.0 / 4.0 / 0.0	95.5 / 4.5 / 0.0	96.0 / 4.0 / 0.0

Table A.2.: Data collection and refinement statistics for the foldon variants P4T, A6Abu, and A6S.

^a Asymmetric unit

^b Values in parenthesis of resolution range, completeness, R_{merge} , and $I / \sigma(I)$ correspond to the last resolution shell

^c Friedel pairs were treated as identical reflections

^d $R_{merge}(I) = \sum_{hkl} \sum_j |I(hkl)_j - I(hkl)| / [\sum_{hkl} I_{hkl}]$, where $I(hkl)_j$ is the j th measurement of the intensity of reflection hkl and $\langle I(hkl) \rangle$ is the average intensity

^e $R = \sum_{hkl} ||F_{obs}| - |F_{calc}|| / \sum_{hkl} |F_{obs}|$, where R_{free} is calculated without a sigma cutoff for a randomly chosen 5 % of reflections, which were not used for structure refinement, and R_{work} is calculated for the remaining reflections

^f Deviations from ideal bond lengths/angles

^g Percentage of residues in favored region / allowed region / outlier region

	Foldon P7Nva	Foldon P7F	Foldon P4T P7Nva
Crystal parameter			
Space group	P6 ₃	P6 ₃	P6 ₃
Cell dimensions			
<i>a</i> , <i>b</i> , <i>c</i> (Å)	28.0; 28.0; 63.7	27.8; 27.8; 63.9	28.1; 28.1; 64.3
α , β , γ (°)	90.0; 90.0; 120.0	90.0; 90.0; 120.0	90.0; 90.0; 120.0
Molecules per AU ^a	1	1	1
Data collection			
Beam line	CuK α	CuK α	CuK α
Wavelength (Å)	1.5418	1.5418	1.5418
Resolution range (Å) ^b	63.7 - 1.1 (1.2 - 1.1)	63.9 - 1.1 (1.2 - 1.1)	24.3 - 1.1 (1.2 - 1.1)
Unique reflections ^c	10,918	8,315	10,940
Completeness (%) ^b	100.00 (100.00)	99.35 (97.99)	98.48
R_{merge} (%) ^{b,d}	4.6 (11.7)	4.6 (11.7)	5.2 (34.1)
$I / \sigma(I)$ ^b	15.0 (4.0)	19.1 (4.3)	15.6 (6.2)
Refinement			
Resolution (Å)	8.0 - 1.1	10.0 - 1.2	10.0 - 1.1
R_{work} / R_{free} ^e	0.128 / 0.152	0.149 / 0.197	0.119 / 0.141
No. atoms			
Protein	219	223	220
Waters	72	55	68
<i>B</i> -factors	8.426	8.783	8.962
R.m.s. deviations ^f			
Bond lengths (Å)	0.024	0.028	0.027
Bond angles (°)	2.060	2.242	2.537
Ramachandran (%) ^g	95.5 / 4.5 / 0.0	96.0 / 4.0 / 0.0	95.2 / 4.8 / 0.0

Table A.3.: Data collection and refinement statistics for the foldon variants P7Nva, P7F, and P4T P7Nva.

^a Asymmetric unit

^b Values in parenthesis of resolution range, completeness, R_{merge} , and $I / \sigma(I)$ correspond to the last resolution shell

^c Friedel pairs were treated as identical reflections

^d $R_{merge}(I) = \sum_{hkl} \sum_j |I(hkl)_j - \langle I(hkl) \rangle| / [\sum_{hkl} I(hkl)]$, where $I(hkl)_j$ is the j th measurement of the intensity of reflection hkl and $\langle I(hkl) \rangle$ is the average intensity

^e $R = \sum_{hkl} ||F_{obs}| - |F_{calc}|| / \sum_{hkl} |F_{obs}|$, where R_{free} is calculated without a sigma cutoff for a randomly chosen 5 % of reflections, which were not used for structure refinement, and R_{work} is calculated for the remaining reflections

^f Deviations from ideal bond lengths/angles

^g Percentage of residues in favored region / allowed region / outlier region

A. Appendix

	Foldon Y13F	Foldon W20NaI	Foldon W20H
Crystal parameter			
Space group	P6 ₃	P6 ₃	
Cell dimensions			
<i>a, b, c</i> (Å)	28.0; 28.0; 64.0	27.8; 27.8; 64.3	
<i>α, β, γ</i> (°)	90.0; 90.0; 120.0	90.0; 90.0; 120.0	
Molecules per AU	1	1	1
Data collection			
Beam line	CuK _α	CuK _α	CuK _α
Wavelength (Å)	1.5418	1.5418	1.5418
Resolution range (Å) ^b	64.0 - 1.0 (1.1 - 1.0)	64.3 - 0.9 (1.0 - 0.9)	
Unique reflections ^c	14,351	19,847	
Completeness (%) ^b	98.43 (94.71)	99.99 (99.93)	
<i>R</i> _{merge} (%) ^{b,d}	6.4 (41.3)	4.8 (41.9)	
<i>I</i> / <i>σ</i> (<i>I</i>) ^b	14.3 (4.3)	27.4 (6.1)	
Refinement			
Resolution (Å)	8.0 - 1.0	10.0 - 0.9	
<i>R</i> _{work} / <i>R</i> _{free} ^e	0.138 / 0.167	0.120 / 0.141	
No. atoms			
Protein	218	220	
Waters	79	68	
<i>B</i> -factors	9.524	8.136	
R.m.s. deviations ^f			
Bond lengths (Å)	0.028	0.023	
Bond angles (°)	2.199	1.864	
Ramachandran (%) ^g	96.0 / 4.0 / 0.0	95.5 / 4.5 / 0	

Table A.4.: Data collection and refinement statistics for the foldon variants Y13F, W20NaI, and W20H.

^a Asymmetric unit

^b Values in parenthesis of resolution range, completeness, *R*_{merge}, and *I* / *σ* (*I*) correspond to the last resolution shell

^c Friedel pairs were treated as identical reflections

^d $R_{merge}(I) = \frac{\sum_{hkl} \sum_j |I(hkl)_j - \langle I(hkl) \rangle|}{\sum_{hkl} I(hkl)}$, where *I*(*hkl*)_{*j*} is the *j*th measurement of the intensity of reflection *hkl* and $\langle I(hkl) \rangle$ is the average intensity

^e $R = \frac{\sum_{hkl} ||F_{obs} - F_{calc}||}{\sum_{hkl} |F_{obs}|}$, where *R*_{free} is calculated without a sigma cutoff for a randomly chosen 5 % of reflections, which were not used for structure refinement, and *R*_{work} is calculated for the remaining reflections

^f Deviations from ideal bond lengths/angles

^g Percentage of residues in favored region / allowed region / outlier region

Foldon L22F	
Crystal parameter	
Space group	P6 ₃
Cell dimensions	
<i>a</i> , <i>b</i> , <i>c</i> (Å)	27.9; 27.9; 63.9
α, β, γ (°)	90.0; 90.0; 120.0
Molecules per AU	1
Data collection	
Beam line	CuK _α
Wavelength (Å)	1.5418
Resolution range (Å) ^b	62.2 - 1.1 (1.2 - 1.1)
Unique reflections ^c	5,922
Completeness (%) ^b	100.0 (100.0)
<i>R</i> _{merge} (%) ^{b,d}	6.8 (42.4)
<i>I</i> / σ (<i>I</i>) ^b	16.4 (5.2)
Refinement	
Resolution (Å)	10.0 - 1.1
<i>R</i> _{work} / <i>R</i> _{free} ^e	0.129 / 0.169
No. atoms	
Protein	220
Waters	68
<i>B</i> -factors	12.475
R.m.s. deviations ^f	
Bond lengths (Å)	0.029
Bond angles (°)	2.120
Ramachandran (%) ^g	96.0 / 4.0 / 0.0

Table A.5.: Data collection and refinement statistics for the foldon variant L22F.^a Asymmetric unit^b Values in parenthesis of resolution range, completeness, *R*_{merge}, and *I* / σ (*I*) correspond to the last resolution shell^c Friedel pairs were treated as identical reflections^d $R_{merge}(I) = \sum_{hkl} \sum_j |I(hkl)_j - \langle I(hkl) \rangle| / [\sum_{hkl} I(hkl)]$, where $I(hkl)_j$ is the *j*th measurement of the intensity of reflection *hkl* and $\langle I(hkl) \rangle$ is the average intensity^e $R = \sum_{hkl} ||F_{obs}| - |F_{calc}|| / \sum_{hkl} |F_{obs}|$, where *R*_{free} is calculated without a sigma cutoff for a randomly chosen 5 % of reflections, which were not used for structure refinement, and *R*_{work} is calculated for the remaining reflections^f Deviations from ideal bond lengths/angles^g Percentage of residues in favored region / allowed region / outlier region

A.2. GdmCl transitions of all foldon variants

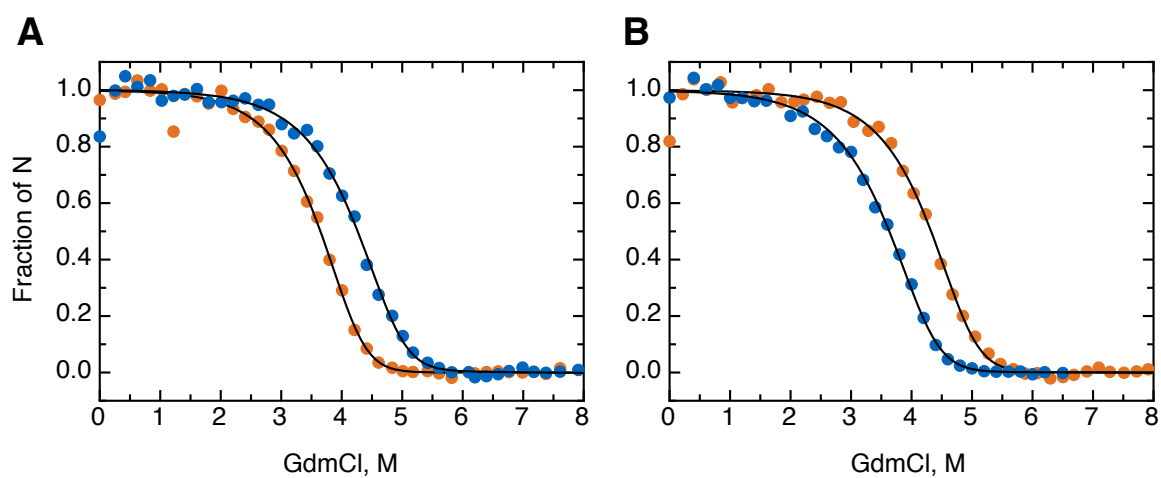


Figure A.1.: A) Normalized GdmCl transition curves for 5 μ M (orange) and 30 μ M (blue) monomer concentrations of foldon P4T. B) Comparison of the 30 μ M foldon wild-type (blue) and 30 μ M foldon P4T (orange) normalized GdmCl transition curves, measured by changes in intrinsic tryptophan fluorescence.

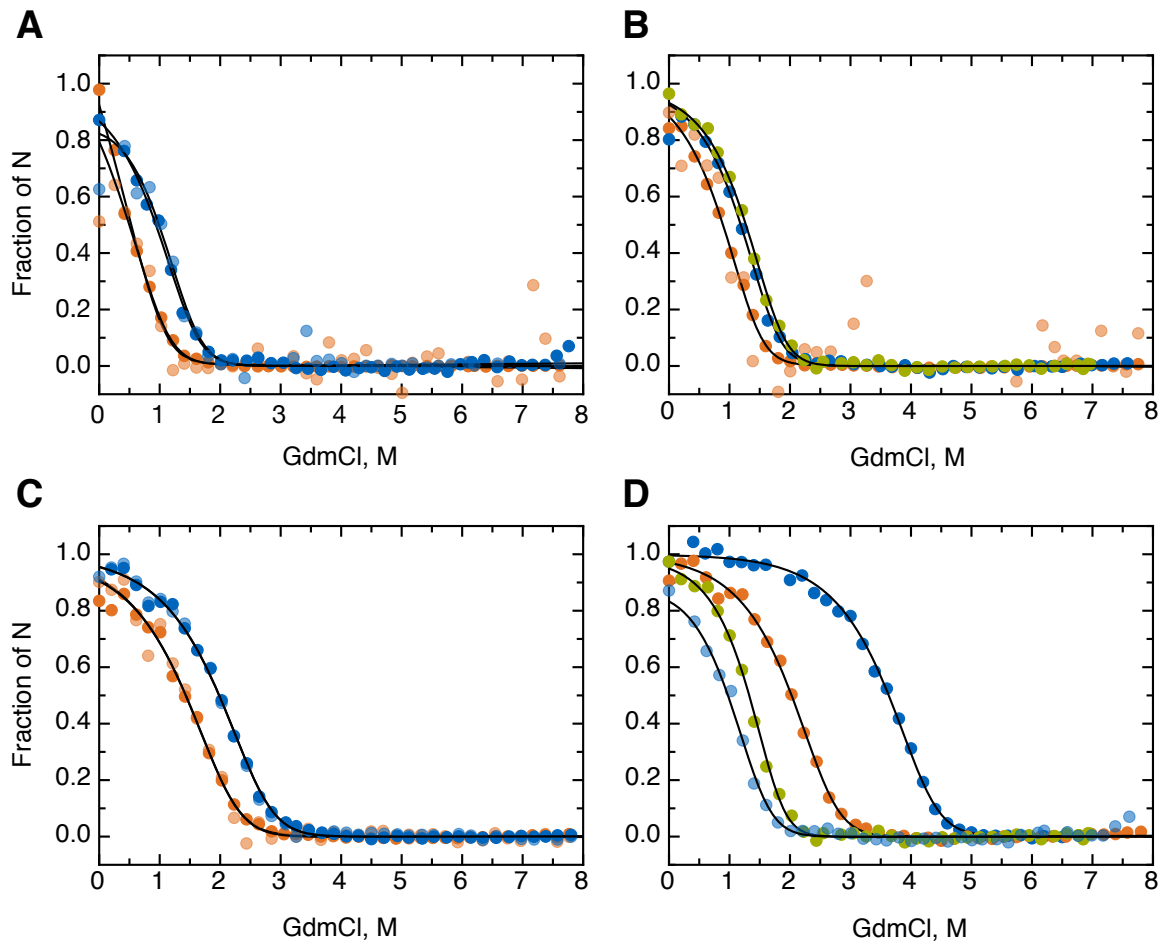


Figure A.2.: Normalized GdmCl transition curves of all three alanine 6 substitutions and comparison with foldon wild-type. A) Normalized GdmCl transition curves for 10 μM (orange) and 30 μM (blue) monomer concentrations of foldon A6V, measured by intrinsic tryptophan fluorescence change (opaque points) or change in circular dichroism absorption at 228 nm (translucent points). B) Normalized GdmCl transition curves for 10 μM (orange), 25 μM (blue) and 30 μM (green) monomer concentrations of foldon A6S, measured by intrinsic tryptophan fluorescence change (opaque points) or change in circular dichroism absorption at 228 nm (translucent points). C) Normalized GdmCl transition curves for 10 μM (orange) and 30 μM (blue) monomer concentrations of foldon A6Abu, measured by intrinsic tryptophan fluorescence change (opaque points) or change in circular dichroism absorption at 228 nm (translucent points). D) Comparison of the foldon A6V (light blue), A6S (green) and A6Abu (orange) variants at position 6 with foldon wild-type (blue) using 30 μM monomer concentration for all variants. All concentrations apply to monomer concentrations.

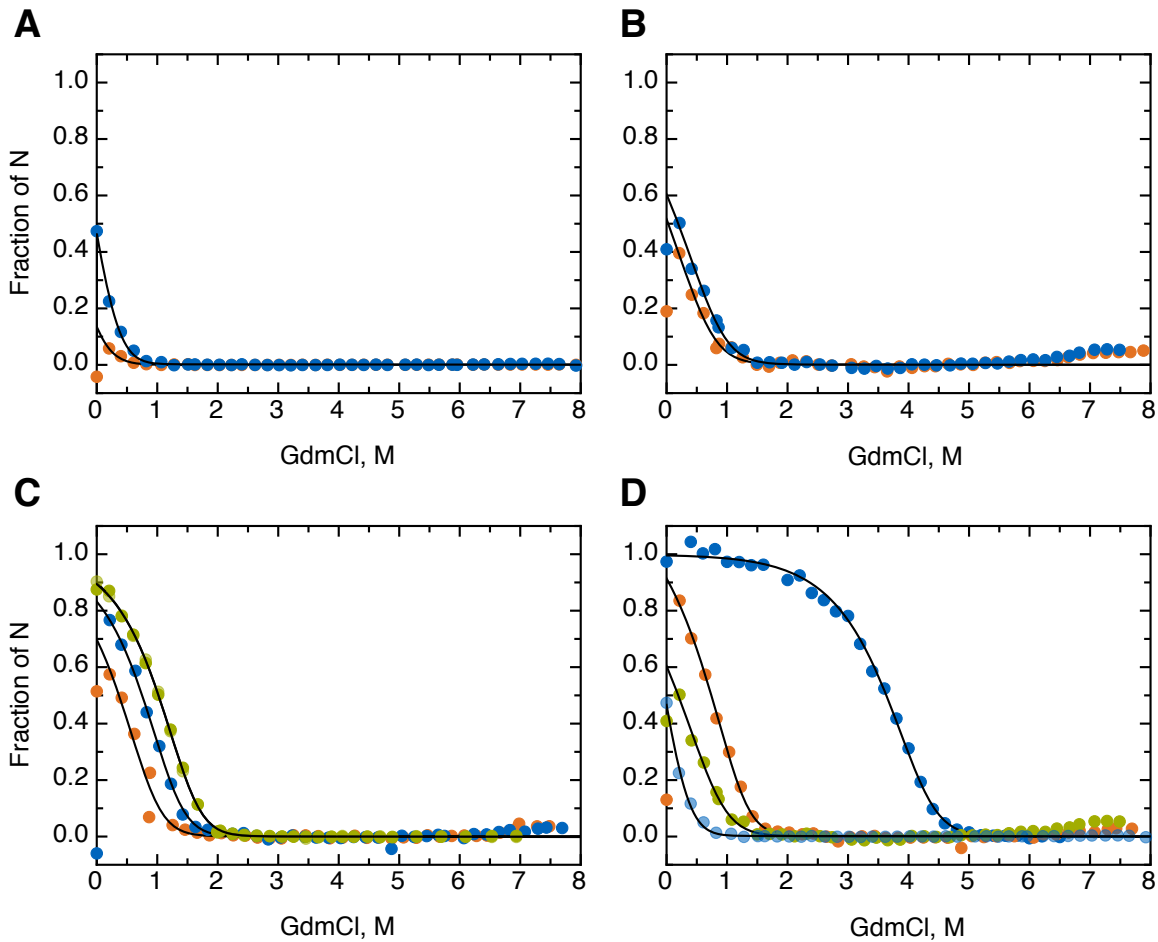


Figure A.3.: Normalized GdmCl transition curves of all three proline 7 substitutions and comparison with foldon wild-type. A) Normalized GdmCl transition curves for 20 μ M (orange) and 50 μ M (blue) monomer concentrations of foldon P7F, measured by intrinsic tryptophan fluorescence change. B) Normalized GdmCl transition curves for 20 μ M (orange) and 30 μ M (blue) monomer concentrations of foldon P7V, measured by intrinsic tryptophan fluorescence change. C) Normalized GdmCl transition curves for 20 μ M (orange), 30 μ M (blue) and 90 μ M (green) monomer concentrations of foldon P7Nva, measured by intrinsic tryptophan fluorescence change (opaque points) or change in circular dichroism absorption at 228 nm (translucent points). D) Comparison of the 30 μ M foldon P7F (light blue), 50 μ M foldon P7V (green), and 50 μ M foldon P7Nva (orange) with 30 μ M foldon wild-type (blue). All concentrations apply to monomer concentrations.

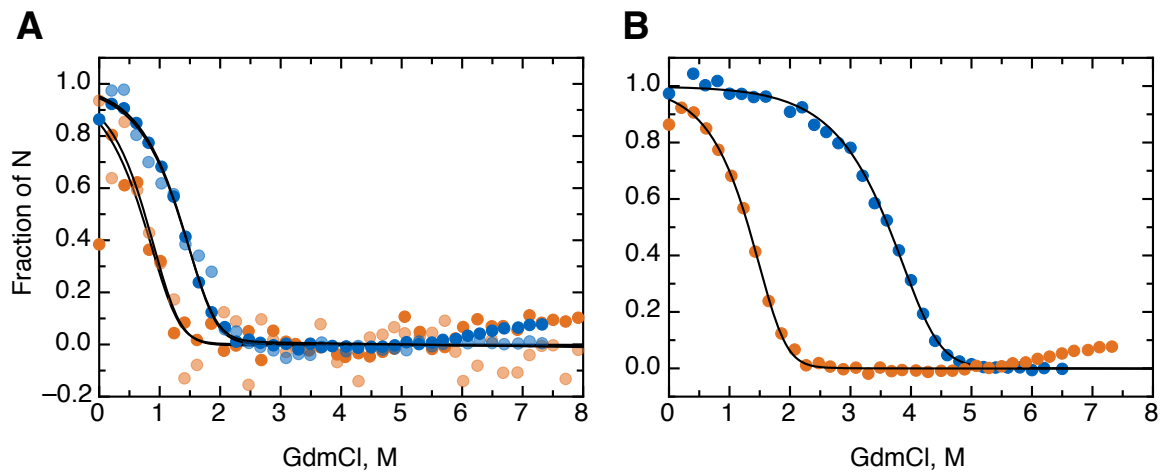


Figure A.4.: A) Normalized GdmCl transition curves for 5 μM (orange) and 30 μM (blue) monomer concentrations of foldon P4T P7Nva, measured by intrinsic tryptophan fluorescence change (opaque points) or change in circular dichroism absorption at 228 nm (translucent points). B) Comparison of the foldon wild-type (blue) and foldon P4T P7Nva (orange) normalized GdmCl transition curves at 30 μM monomer concentration, measured by changes in intrinsic tryptophan fluorescence.

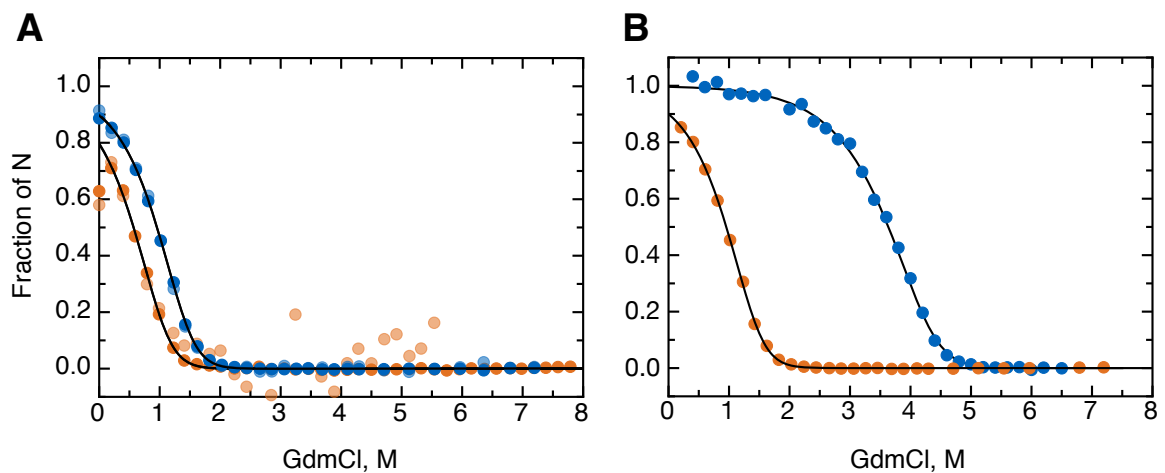


Figure A.5.: A) Normalized GdmCl transition curves for 5 μM (orange) and 30 μM (blue) monomer concentrations of foldon Y13F, measured by intrinsic tryptophan fluorescence change (opaque points) or change in circular dichroism absorption at 228 nm (translucent points). B) Comparison of the foldon wild-type (blue) and foldon Y13F (orange) normalized GdmCl transition curves at 30 μM monomer concentration, measured by changes in intrinsic tryptophan fluorescence.

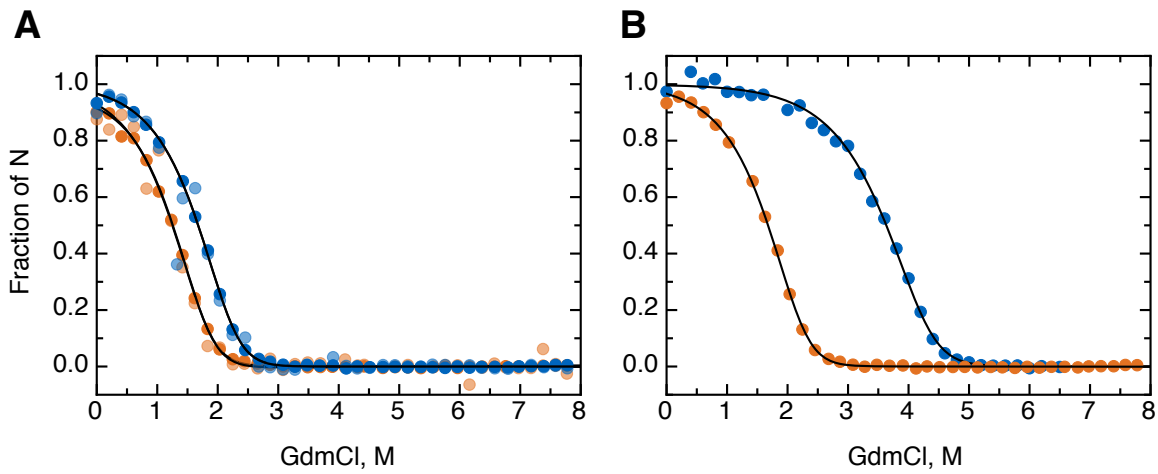


Figure A.6.: A) Normalized GdmCl transition curves for 5 μM (orange) and 30 μM (blue) monomer concentrations of foldon L22F, measured by intrinsic tryptophan fluorescence change (opaque points) or change in circular dichroism absorption at 228 nm (translucent points). B) Comparison of the foldon wild-type (blue) and foldon L22F (orange) normalized GdmCl transition curves at 30 μM monomer concentration, measured by changes in intrinsic tryptophan fluorescence.

A.3. Urea transitions of foldon variants

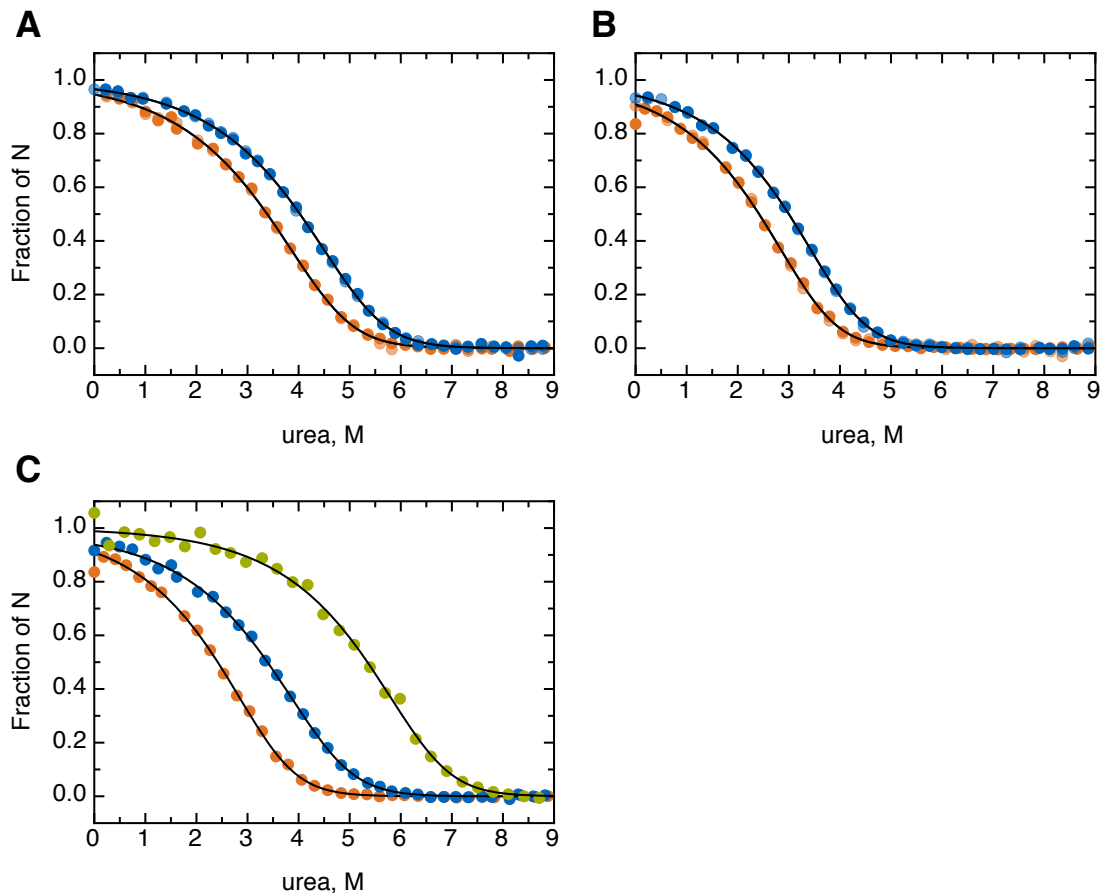


Figure A.7.: Normalized urea transition curves of all alanine 6 substitutions, A) Normalized urea transition curves for 30 μM (orange) and 60 μM (blue) monomer concentrations of foldon A6Abu, measured by intrinsic tryptophan fluorescence change (opaque points) or change in circular dichroism absorption at 228 nm (translucent points). B) Normalized urea transition curves for 30 μM (orange) and 60 μM (blue) monomer concentrations of foldon A6S, measured by intrinsic tryptophan fluorescence change (opaque points) or change in circular dichroism absorption at 228 nm (translucent points). C) Comparison of the foldon A6S (orange) and A6Abu (blue) variants at position 6 at 30 μM monomer concentration with foldon wild-type (green) using 5 μM monomer concentration.

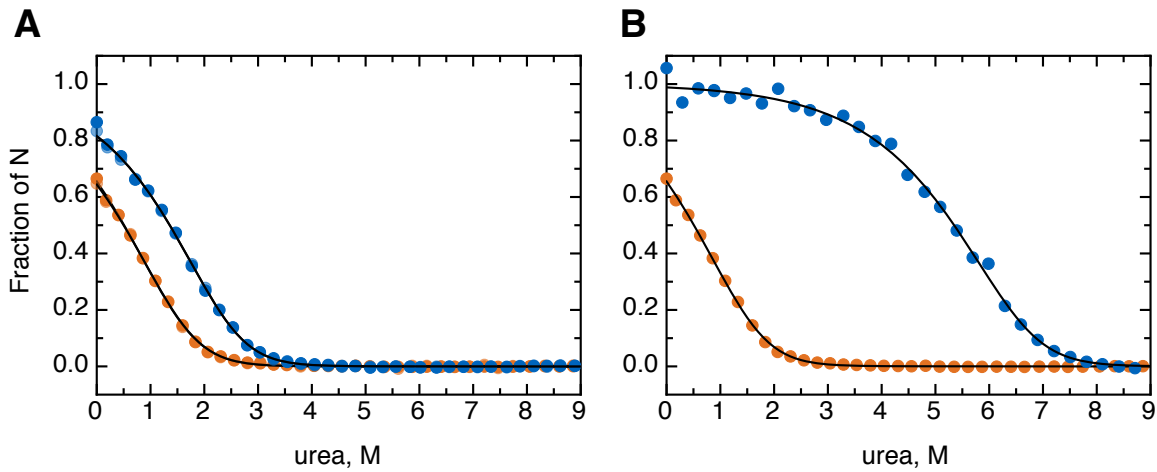


Figure A.8.: A) Normalized urea transition curves for 30 μM (orange) and 90 μM (blue) monomer concentrations of foldon P7Nva, measured by intrinsic tryptophan fluorescence change (opaque points) or change in circular dichroism absorption at 228 nm (translucent points). B) Comparison of a 5 μM foldon wild-type monomer concentration (blue) and a 30 μM foldon P7Nva monomer concentration (orange) normalized urea transition curves, measured by change in intrinsic tryptophan fluorescence.

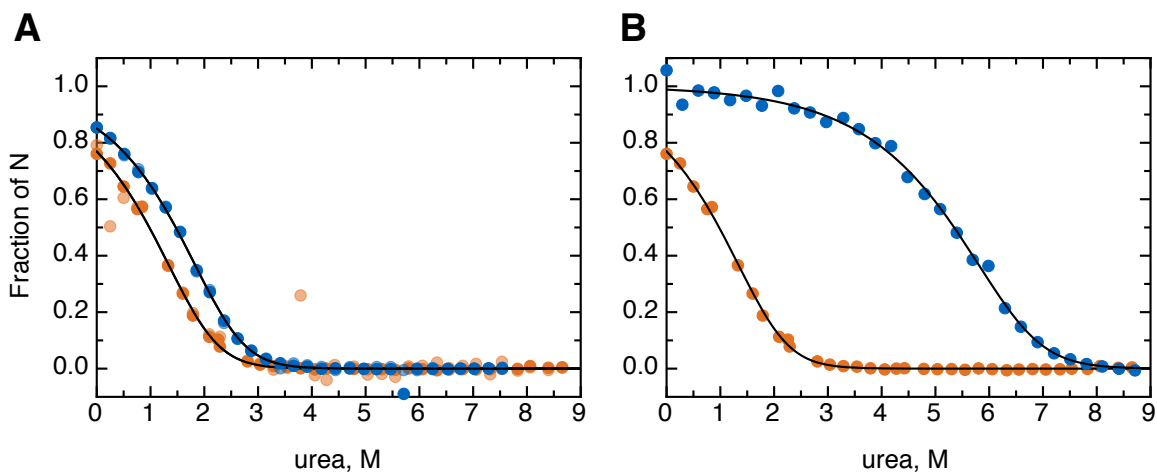


Figure A.9.: A) Normalized urea transition curves for 30 μM (orange) and 60 μM (blue) monomer concentrations of foldon Y13F, measured by intrinsic tryptophan fluorescence change (opaque points) or change in circular dichroism absorption at 228 nm (translucent points). B) Comparison of a 5 μM foldon wild-type monomer concentration (blue) and a 30 μM foldon Y13F monomer concentration (orange) normalized urea transition curves, measured by change in intrinsic tryptophan fluorescence.

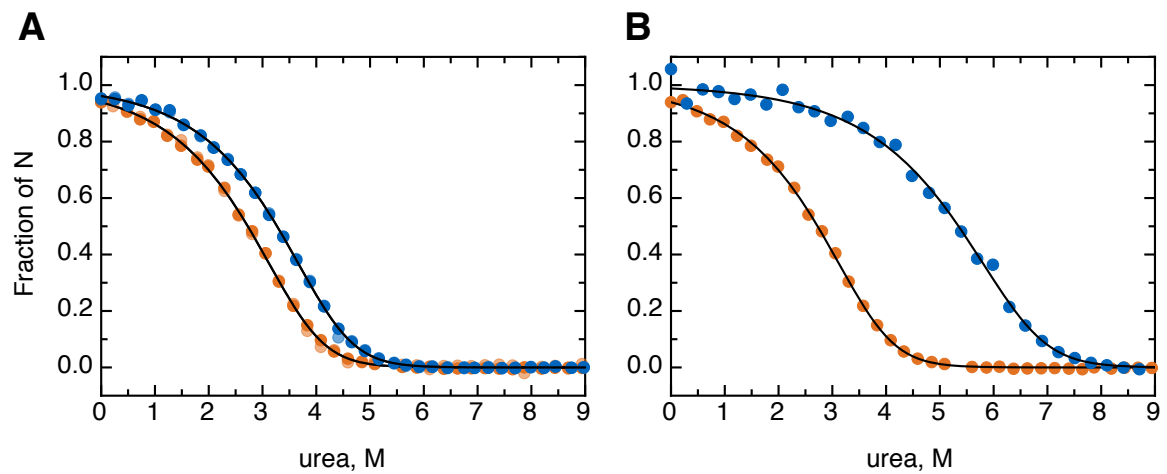


Figure A.10.: A) Normalized urea transition curves for 30 μM (orange) and 60 μM (blue) monomer concentrations of foldon L22F, measured by intrinsic tryptophan fluorescence change (opaque points) or change in circular dichroism absorption at 228 nm (translucent points). B) Comparison of a 5 μM foldon wild-type monomer concentration (blue) and a 30 μM foldon L22F monomer concentration (orange) normalized urea transition curves, measured by change in intrinsic tryptophan fluorescence.

A.4. GdmCl unfolding kinetics

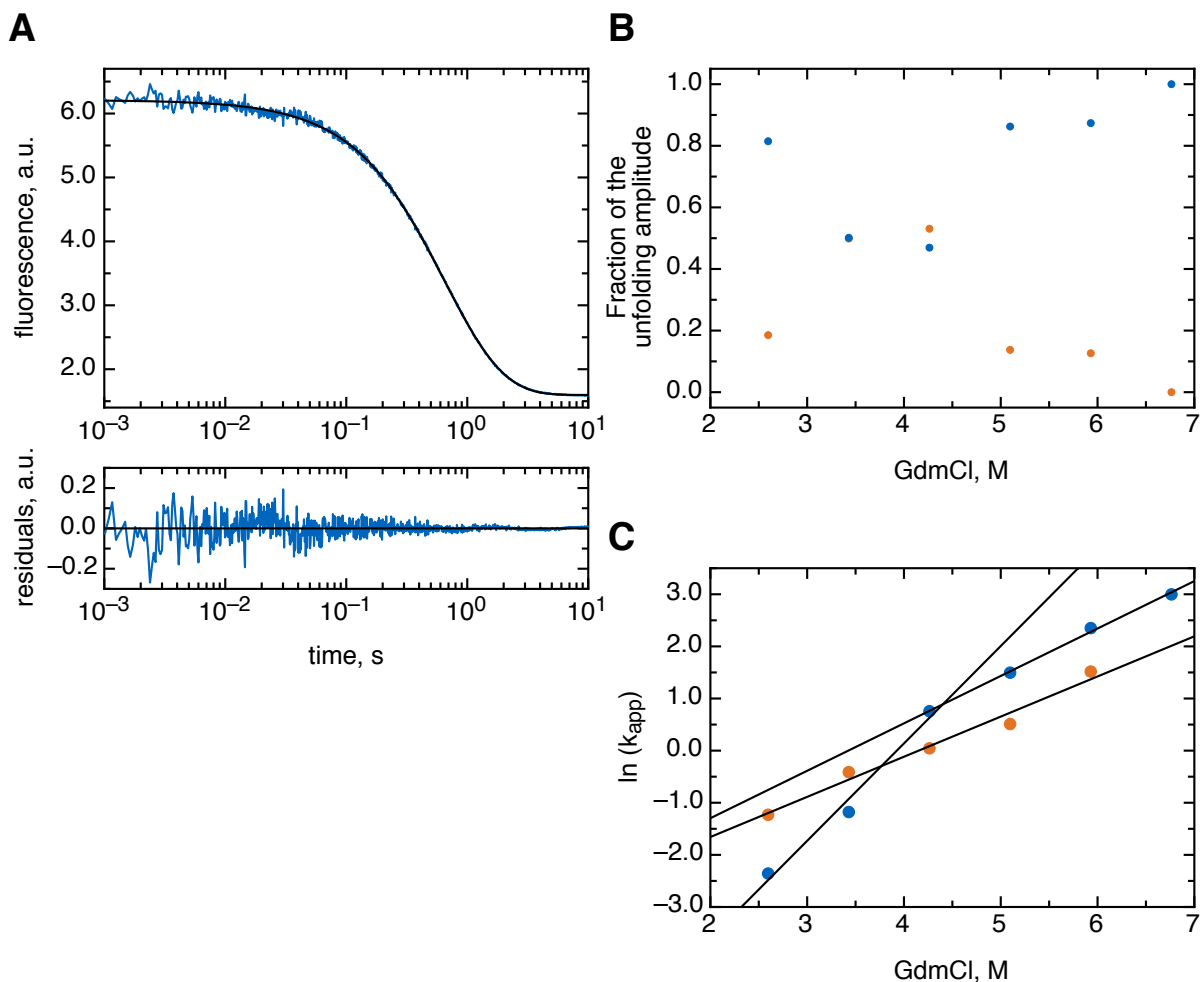


Figure A.11.: Foldon A6Abu GdmCl unfolding kinetics. A) Unfolding of foldon A6Abu in 0.58 M GdmCl by dilution with a 6fold excess of a 6 M urea buffer, fitted with a two-exponential equation, and the corresponding residuals B) Fraction of the amplitudes of the one or two phases from one- or two-exponential fitting of the unfolding data at varying final GdmCl concentrations. C) Linear unfolding plot of the common logarithms of the two phases of the foldon A6Abu GdmCl unfolding kinetics against the corresponding GdmCl concentrations.

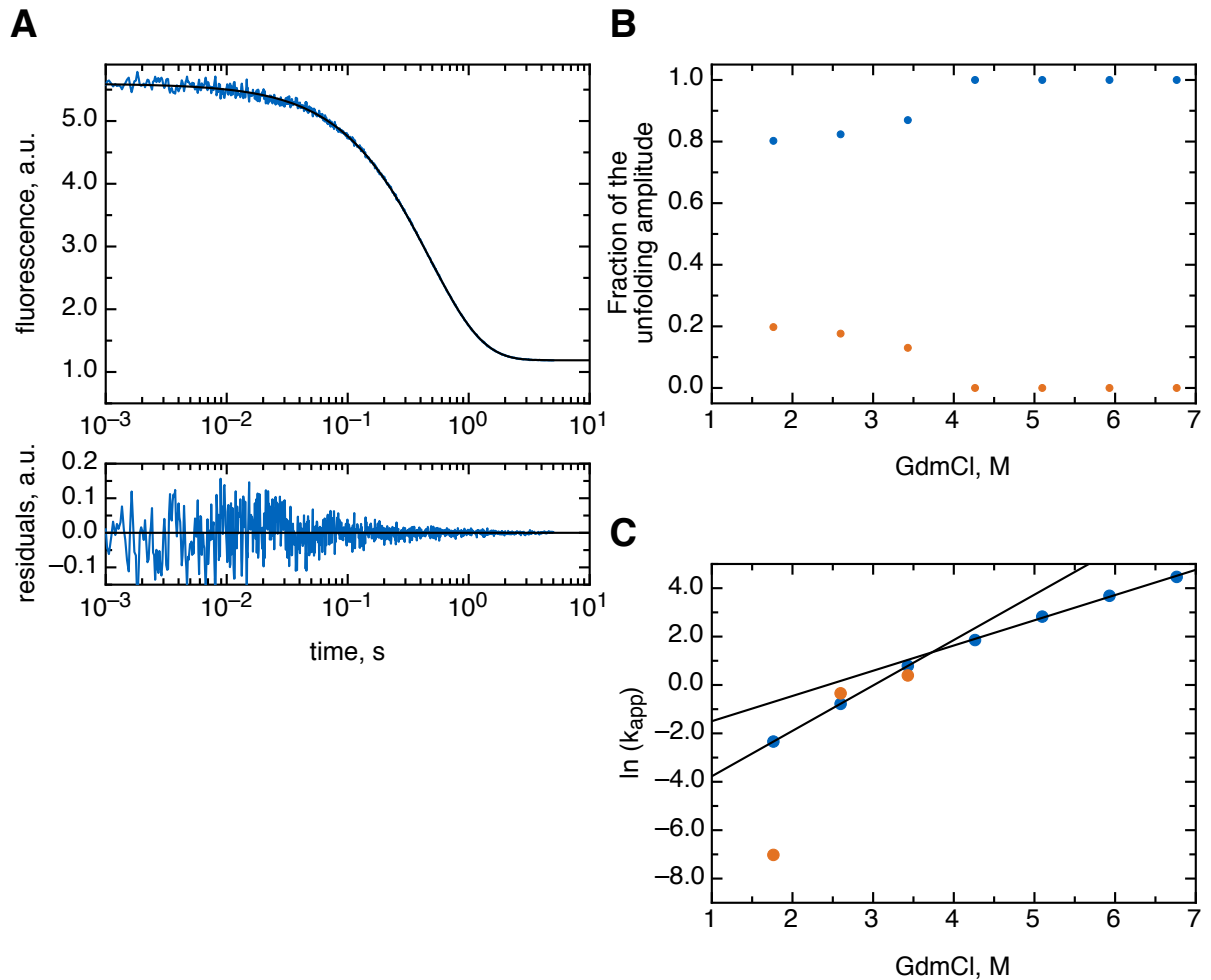


Figure A.12.: Foldon A6S GdmCl unfolding kinetics. A) Unfolding of foldon A6S in 0.58 M GdmCl by dilution with a 6fold excess of a 4 M urea buffer, fitted with a two-exponential equation, and the corresponding residuals. B) Fraction of the amplitudes of the one or two phases from one- or two-exponential fitting of the unfolding data at varying final GdmCl concentrations. C) Linear unfolding plot of the common logarithms of the two phases of the foldon A6S GdmCl unfolding kinetics against the corresponding GdmCl concentrations.

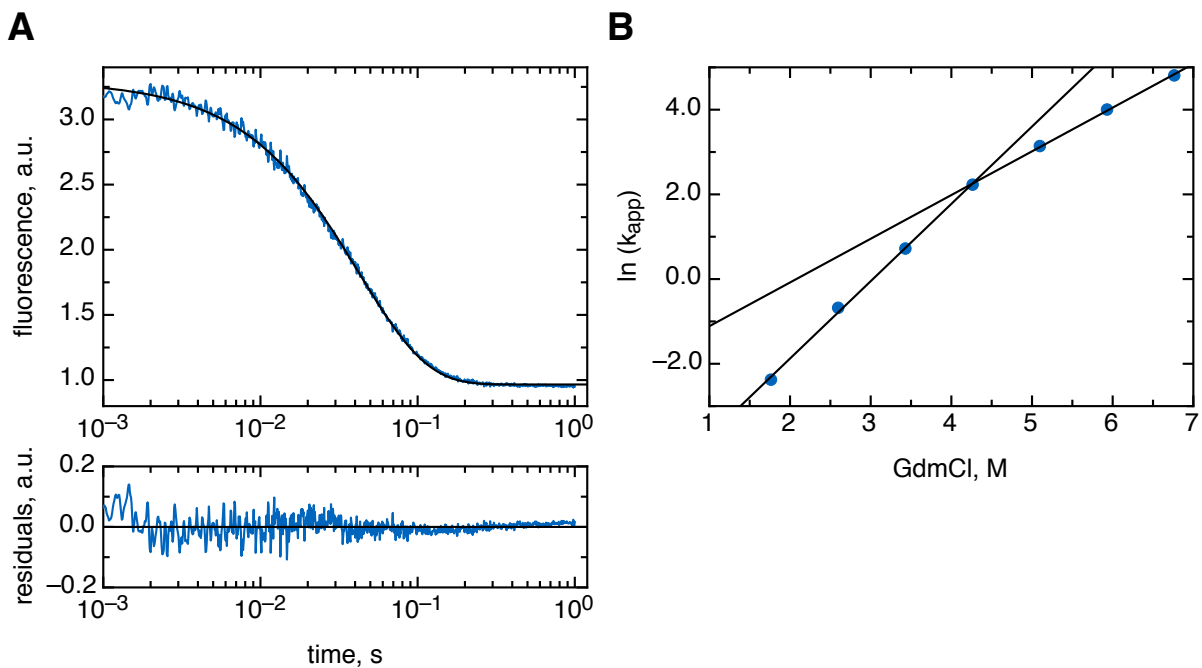


Figure A.13.: Foldon A6V GdmCl unfolding kinetics. A) Unfolding of foldon A6V in 0.58 M GdmCl by dilution with a 6fold excess of a 6 M urea buffer, fitted with a single-exponential equation, and the corresponding residuals B) Plot of the common logarithms of the rate constants of the foldon A6V GdmCl unfolding kinetics against the corresponding GdmCl concentrations.

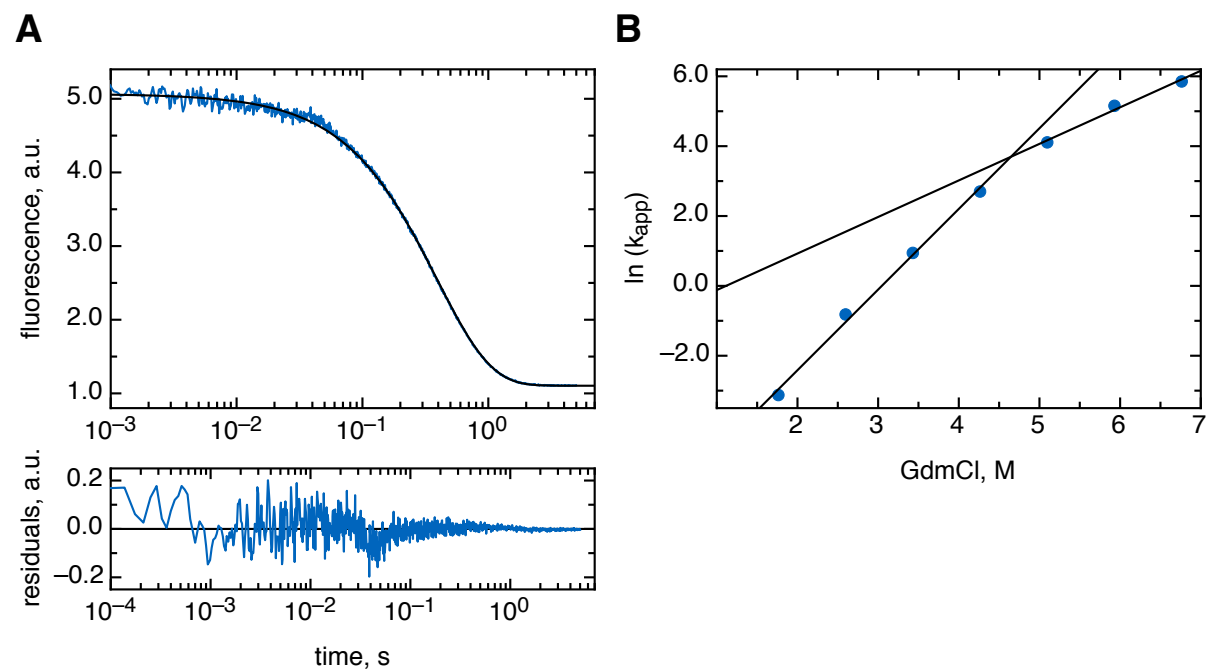


Figure A.14.: Foldon P7Nva GdmCl unfolding kinetics. A) Unfolding of foldon P7Nva in 0.58 M GdmCl by dilution with a 6fold excess of a 4 M urea buffer, fitted with a single-exponential equation, and the corresponding residuals. B) Plot of the common logarithms of the rate constants of foldon P7Nva GdmCl unfolding kinetics against the corresponding GdmCl concentrations.

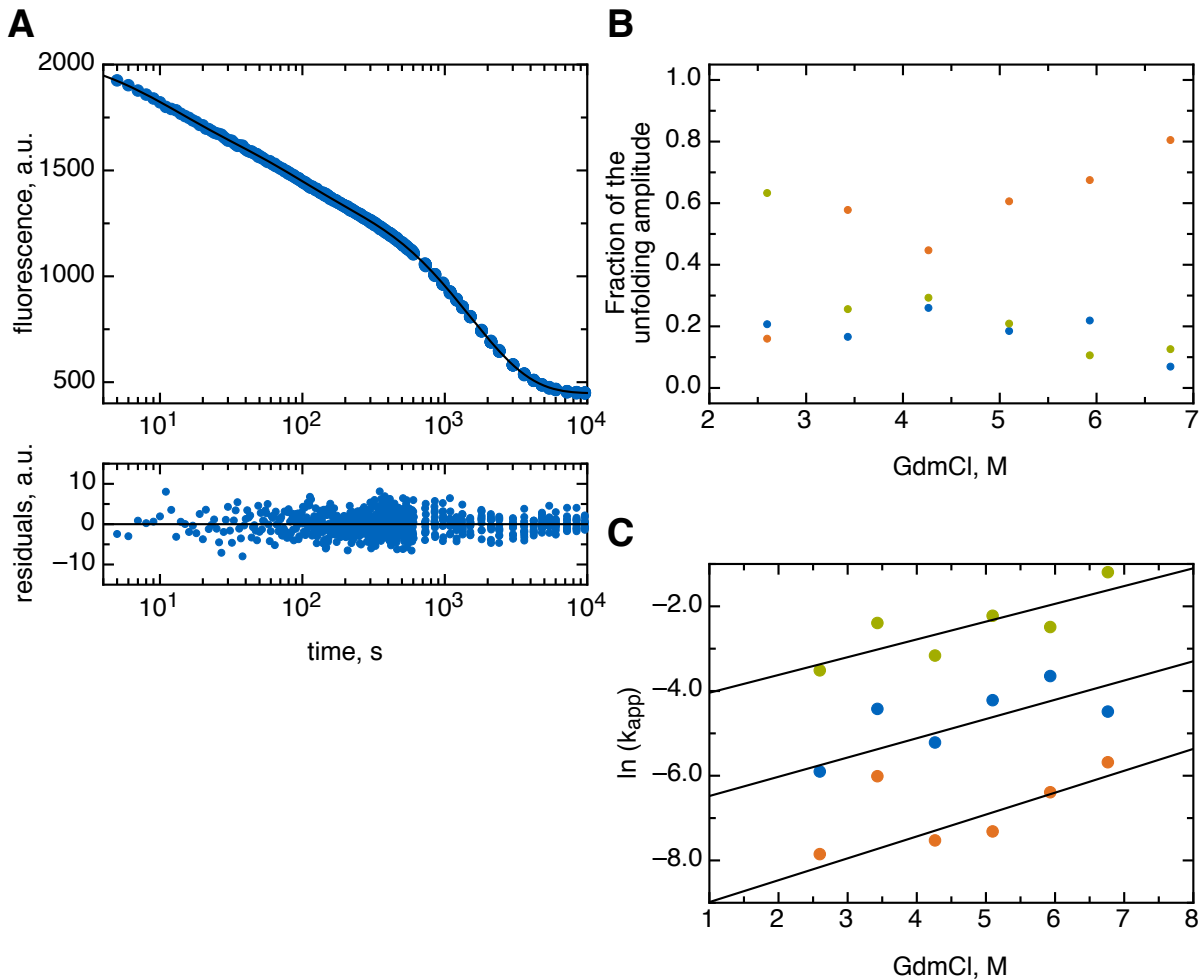


Figure A.15.: Foldon P4T P7Nva GdmCl unfolding kinetics. A) Unfolding of foldon P4T P7Nva in 0.58 M GdmCl by dilution with a 6fold excess of a 3 M urea buffer, fitted with a three-exponential equation, and the corresponding residuals. B) Fraction of the amplitudes of the all phases from three-exponential fitting of the unfolding data at varying final GdmCl concentrations. C) Linear unfolding plot of the common logarithms of the slow (orange), medium (blue) and fast (green) phases of the foldon P4T P7Nva GdmCl unfolding kinetics against the corresponding GdmCl concentrations.

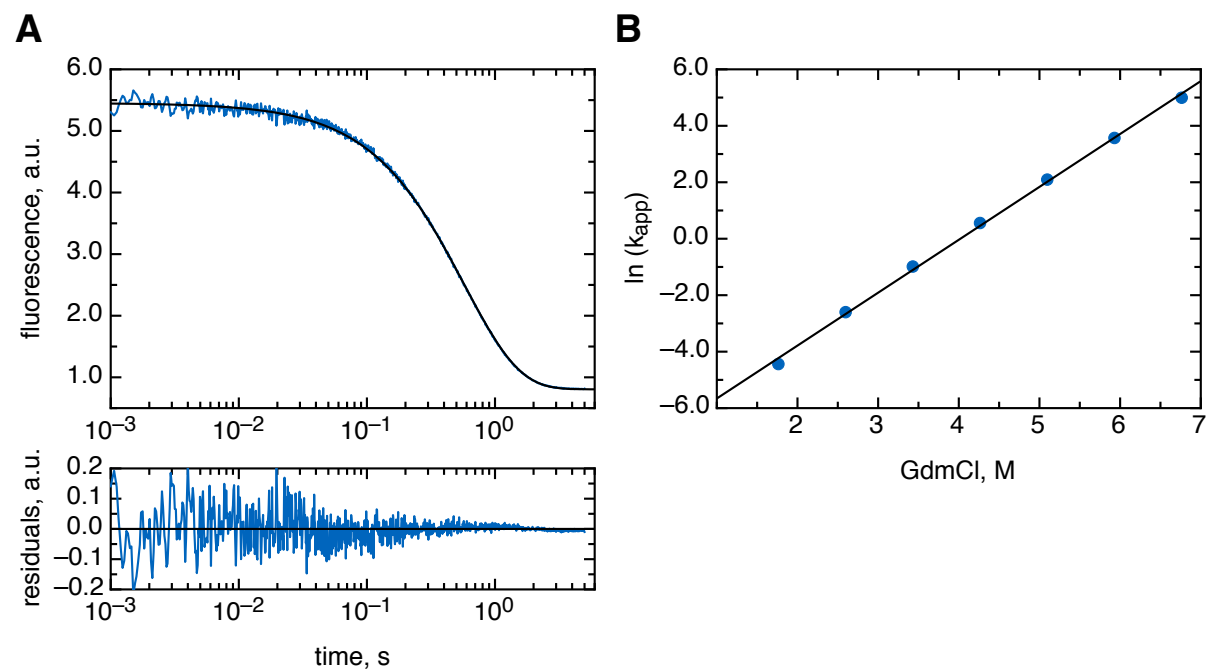


Figure A.16.: Foldon Y13F GdmCl unfolding kinetics. A) Unfolding of foldon Y13F in 0.58 M GdmCl by dilution with a 6fold excess of a 5 M urea buffer, fitted with a single-exponential equation, and the corresponding residuals. B) Plot of the common logarithms of the rate constants of foldon Y13F GdmCl unfolding kinetics against the corresponding GdmCl concentrations.

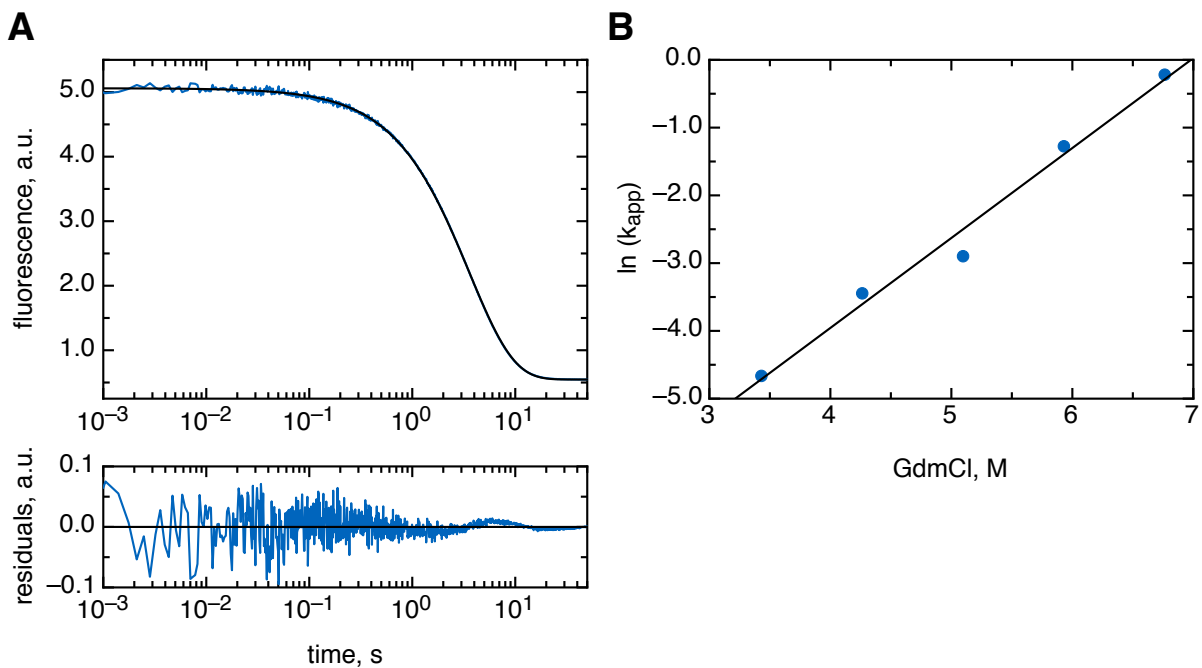


Figure A.17.: Foldon L22F GdmCl unfolding kinetics. A) Unfolding of foldon L22F in 0.58 M GdmCl by dilution with a 6fold excess of a 7 M urea buffer, fitted with a single-exponential equation, and the corresponding residuals. B) Plot of the common logarithm of the apparent unfolding rate constants of foldon L22F GdmCl unfolding kinetics against the corresponding GdmCl concentrations.

A.5. Urea unfolding kinetics

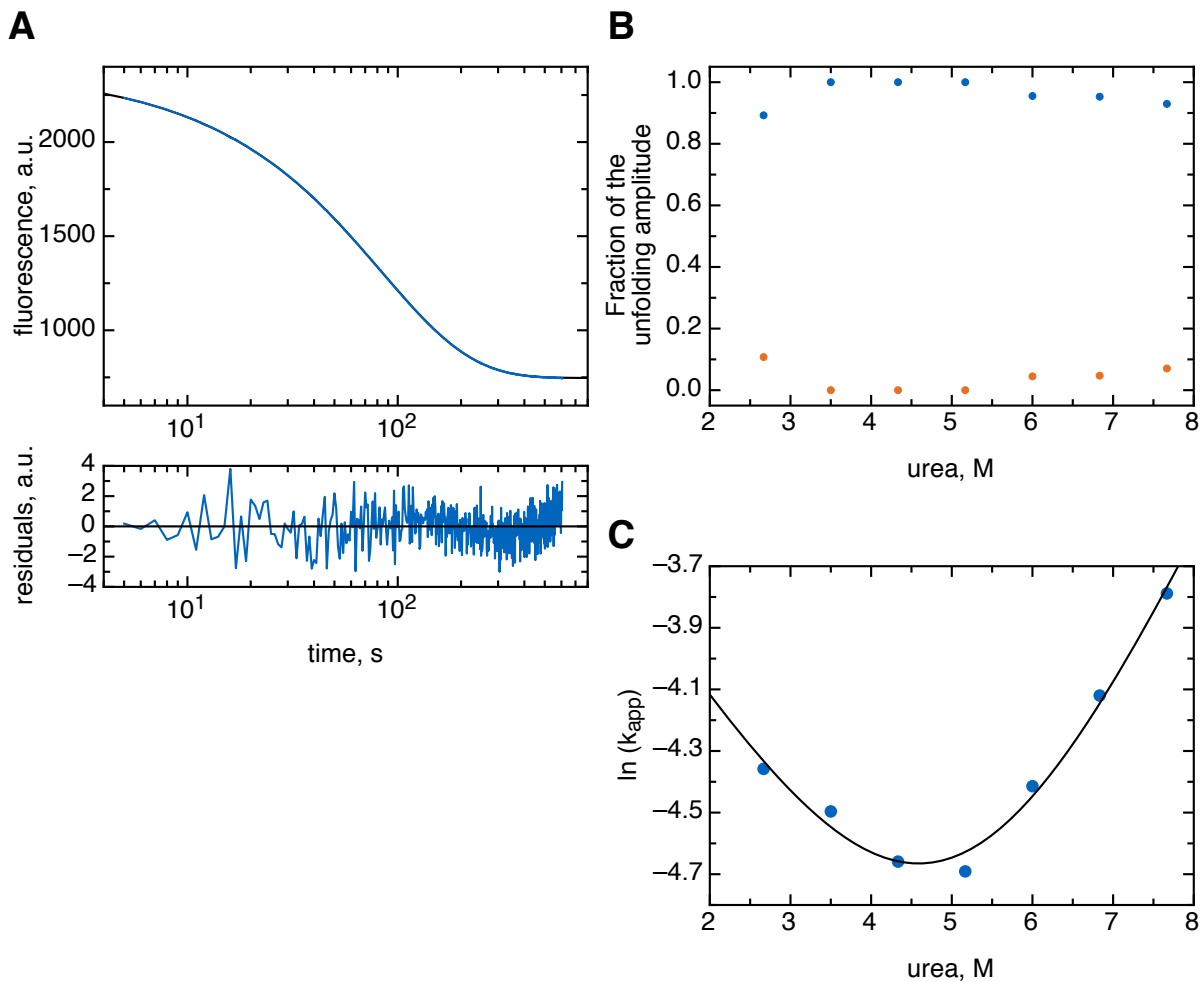


Figure A.18.: Foldon A6Abu urea unfolding kinetics. A) Unfolding of foldon A6Abu in 1 M urea by dilution with a 6fold excess of a 7 M urea buffer, fitted with a two-exponential equation, and the corresponding residuals. B) Fraction of the amplitudes of the two phases from two-exponential fitting of the unfolding data at varying final urea concentrations. C) Chevron plot of the common logarithms of the major phase of the foldon A6Abu urea unfolding kinetics at varying urea concentrations.

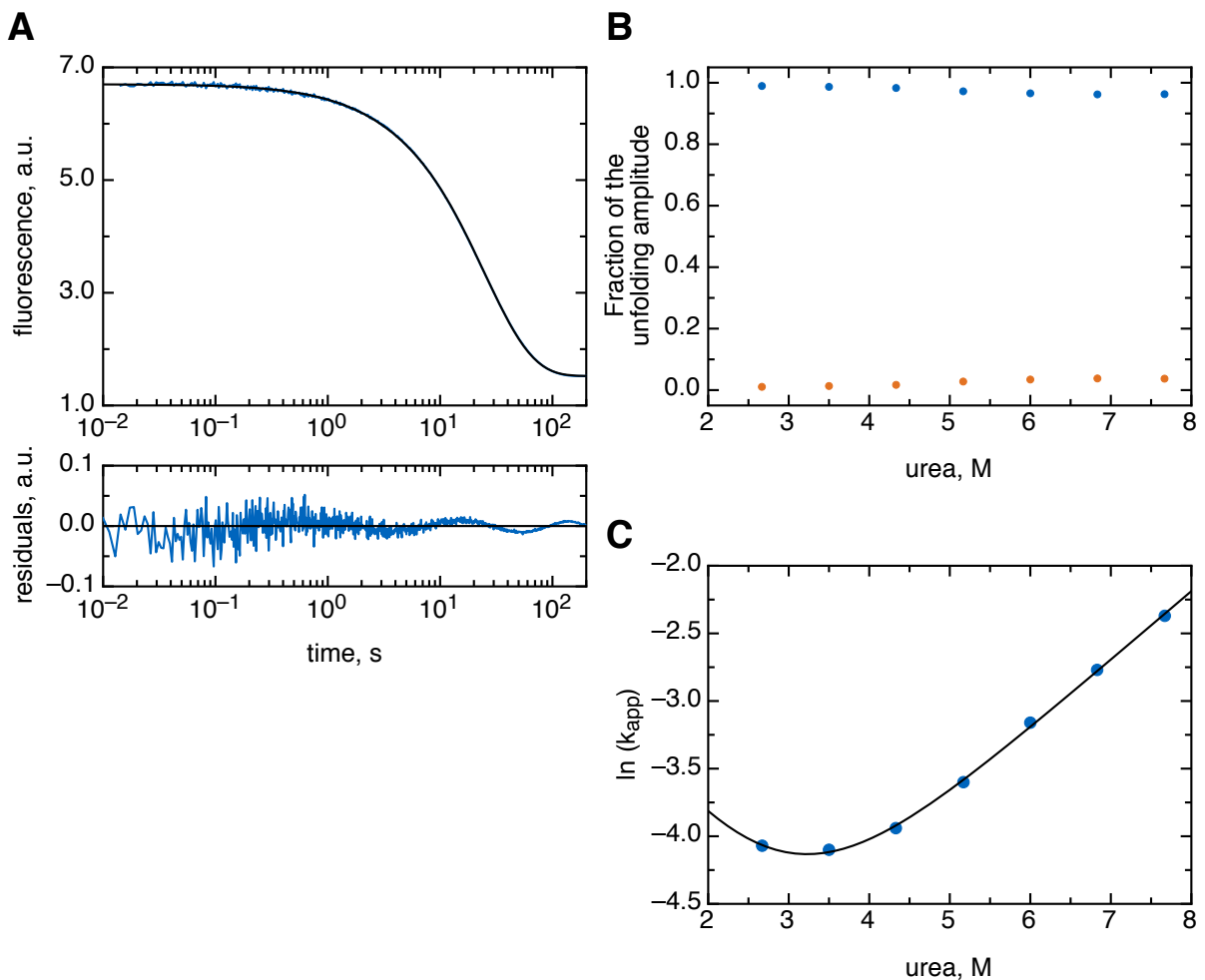


Figure A.19.: Foldon A6S urea unfolding kinetics. A) Unfolding of foldon A6S in 1 M urea by dilution with a 6fold excess of a 7 M urea buffer, fitted with a two-exponential equation, and the corresponding residuals. B) Fraction of the amplitudes of the two phases from two-exponential fitting of the unfolding data at varying final urea concentrations. C) Chevron plots of the common logarithms of the major phase of the foldon A6S urea unfolding kinetics at varying urea concentrations.

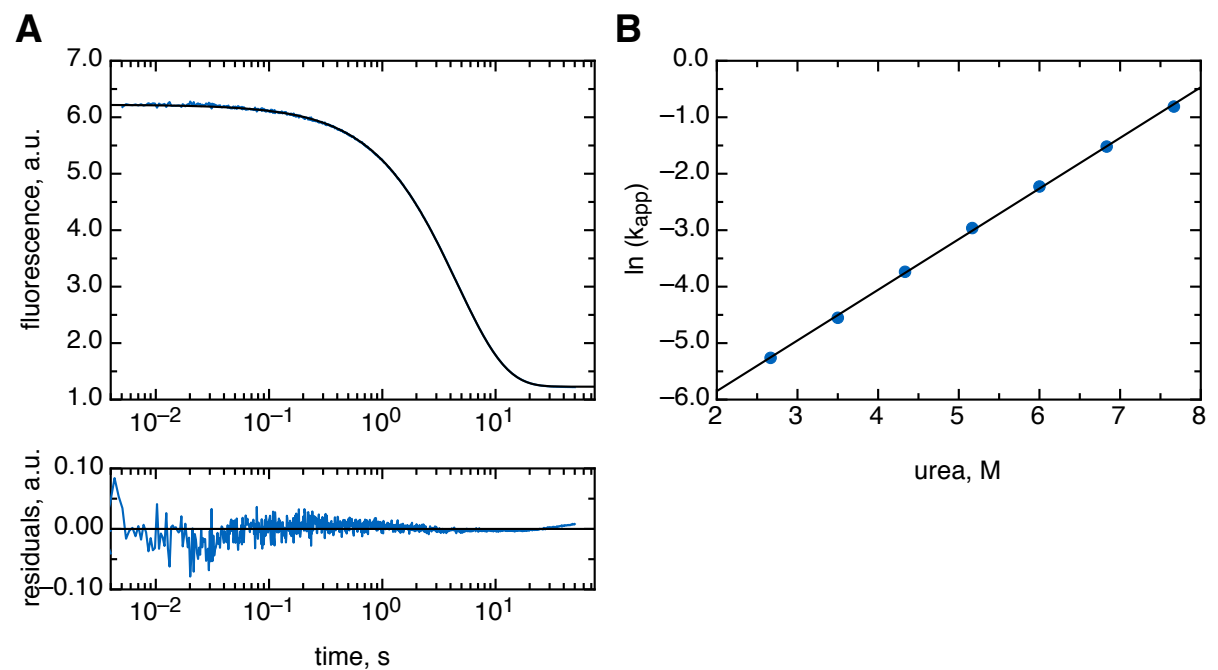


Figure A.20.: Foldon P7Nva urea unfolding kinetics. A) Unfolding of foldon P7Nva in 1 M urea by dilution with a 6fold excess of a 8 M urea buffer, fitted with a single-exponential equation, and the corresponding residuals. B) Plot of the common logarithms of the rate constants of foldon P7Nva urea unfolding kinetics against the corresponding urea concentrations.

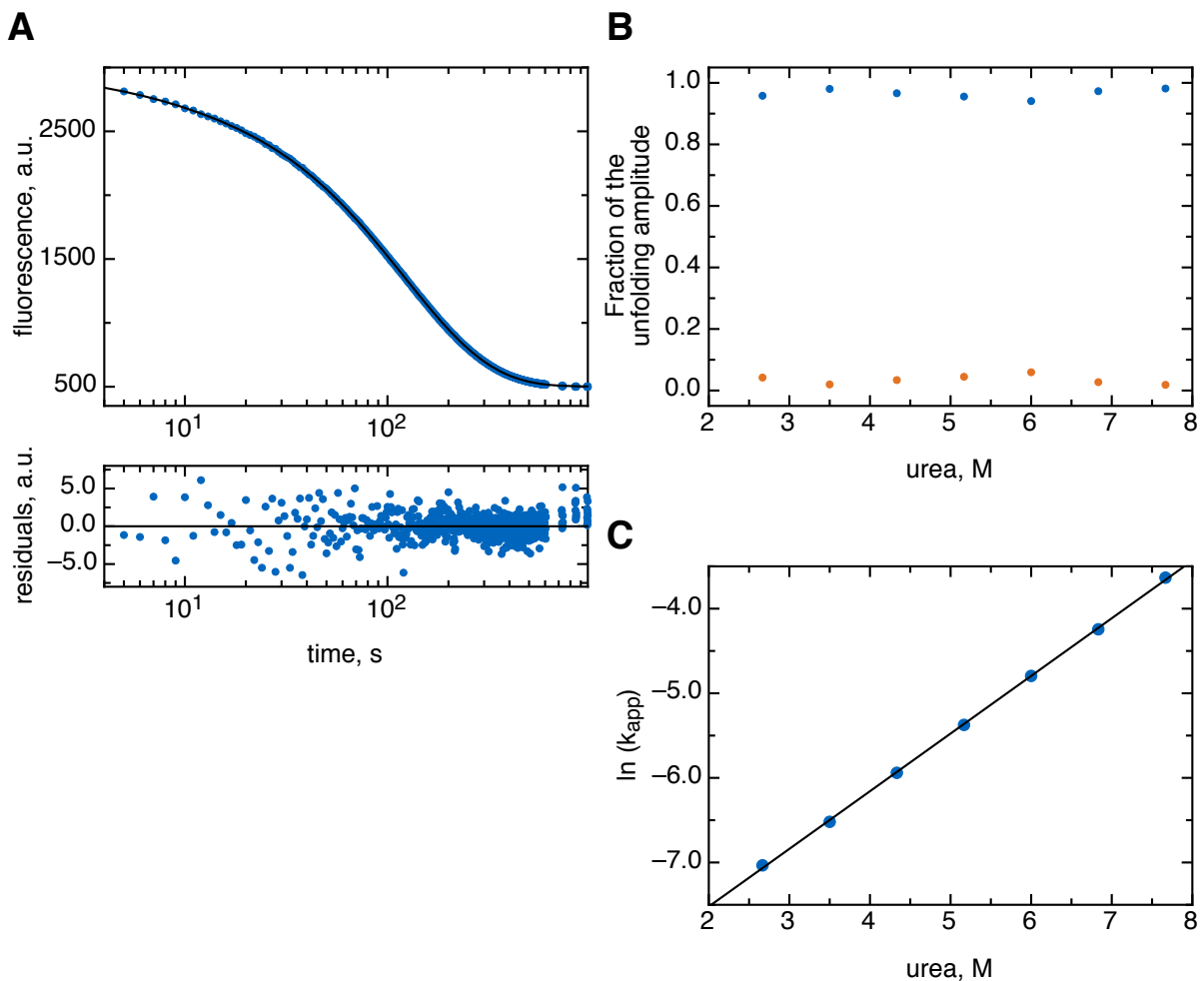


Figure A.21.: Foldon Y13F urea unfolding kinetics. A) Unfolding of foldon Y13F in 1 M urea by dilution with a 6fold excess of a 7 M urea buffer, fitted with a two-exponential equation, and the corresponding residuals. B) Fraction of the amplitudes of the two phases from two-exponential fitting of the unfolding data at varying final urea concentrations. C) Linear unfolding plot of the common logarithms of the major phase of the foldon Y13F urea unfolding kinetics at varying urea concentrations.

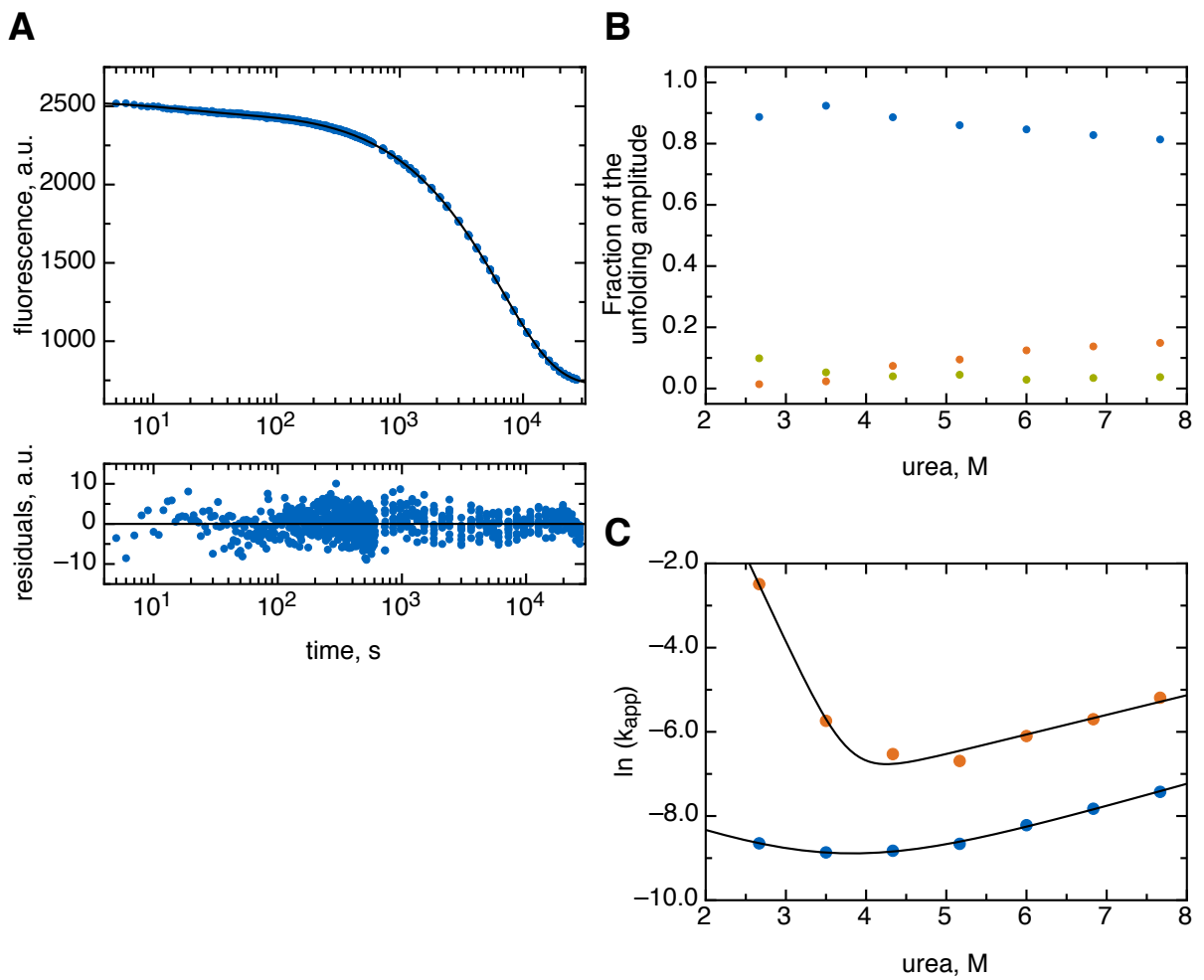


Figure A.22.: Foldon L22F urea unfolding kinetics. A) Unfolding of foldon Y13F in 1 M urea by dilution with a 6fold excess of a 5 M urea buffer, fitted with a three-exponential equation, and the corresponding residuals. B) Fraction of the amplitudes of the three phases from three-exponential fitting of the unfolding data at varying final urea concentrations. C) Chevron plots of the common logarithms of the two major phases of the foldon L22F urea unfolding kinetics against varying urea concentrations.

A.6. Reaction order from initial slope determination

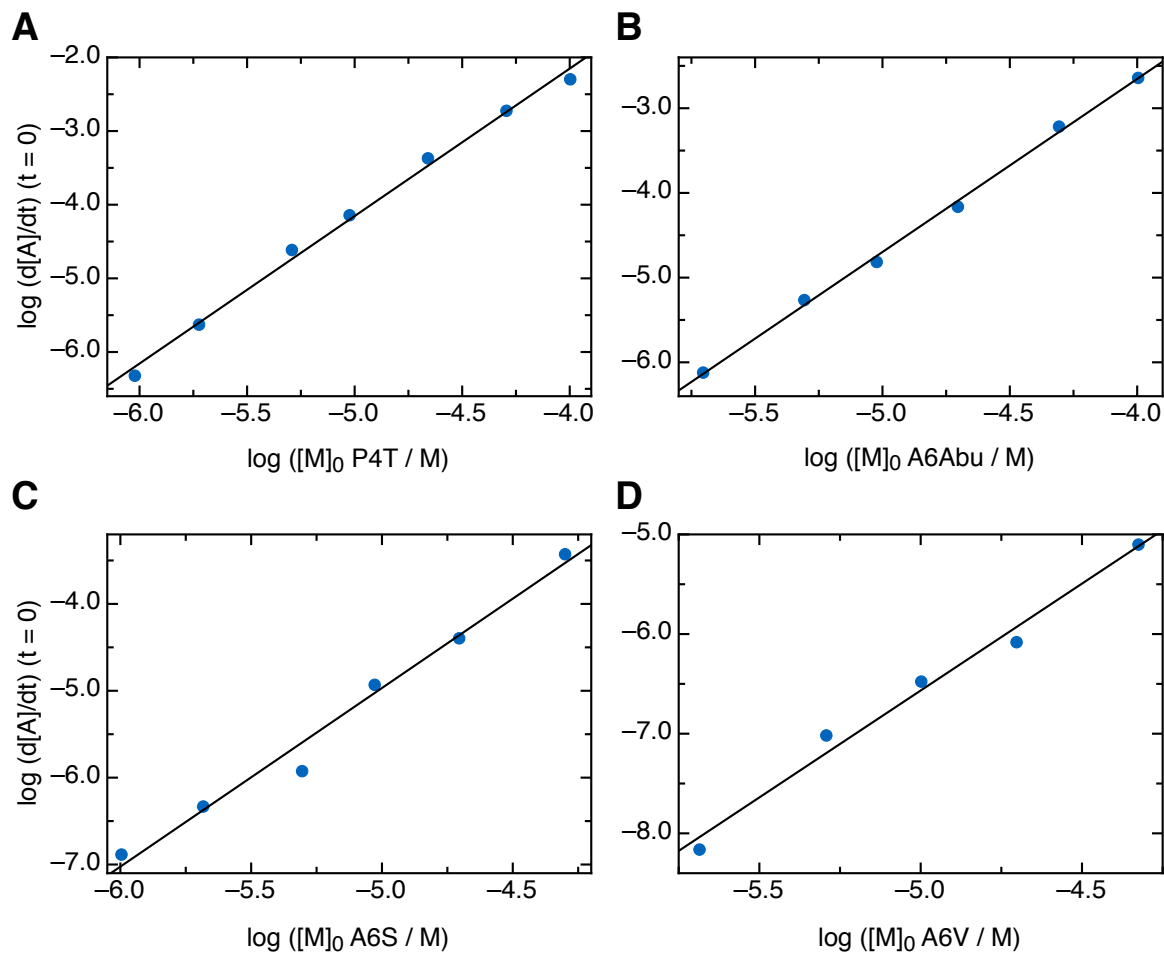


Figure A.23.: Reaction order determination of the GdmCl refolding experiments from the initial slope for the foldon variants A) foldon P4T, B) foldon A6Abu, C) foldon A6S, D) foldon A6V.

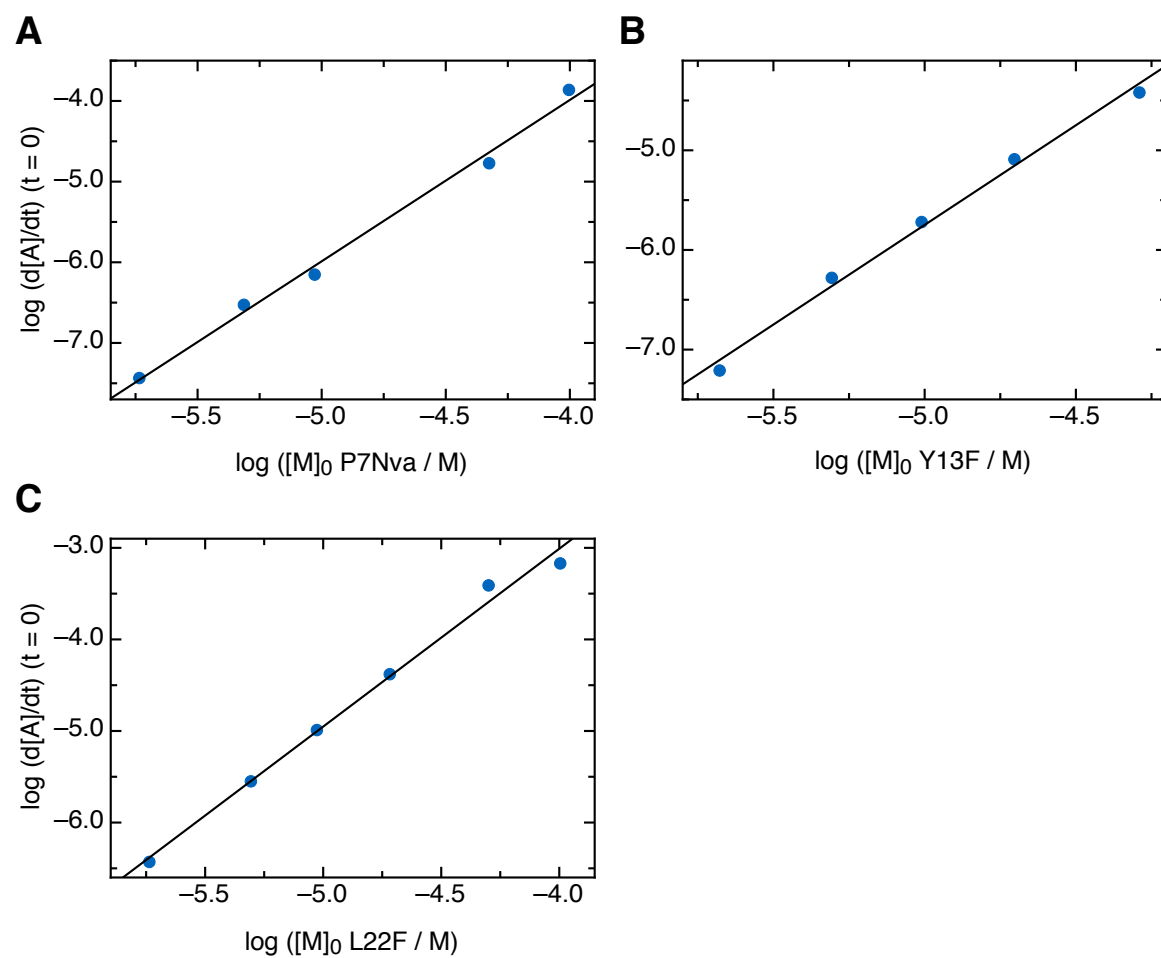


Figure A.24.: Reaction order determination of the GdmCl refolding experiments from the initial slope for the foldon variants A) foldon P7Nva, B) foldon Y13F, C) foldon L22F.

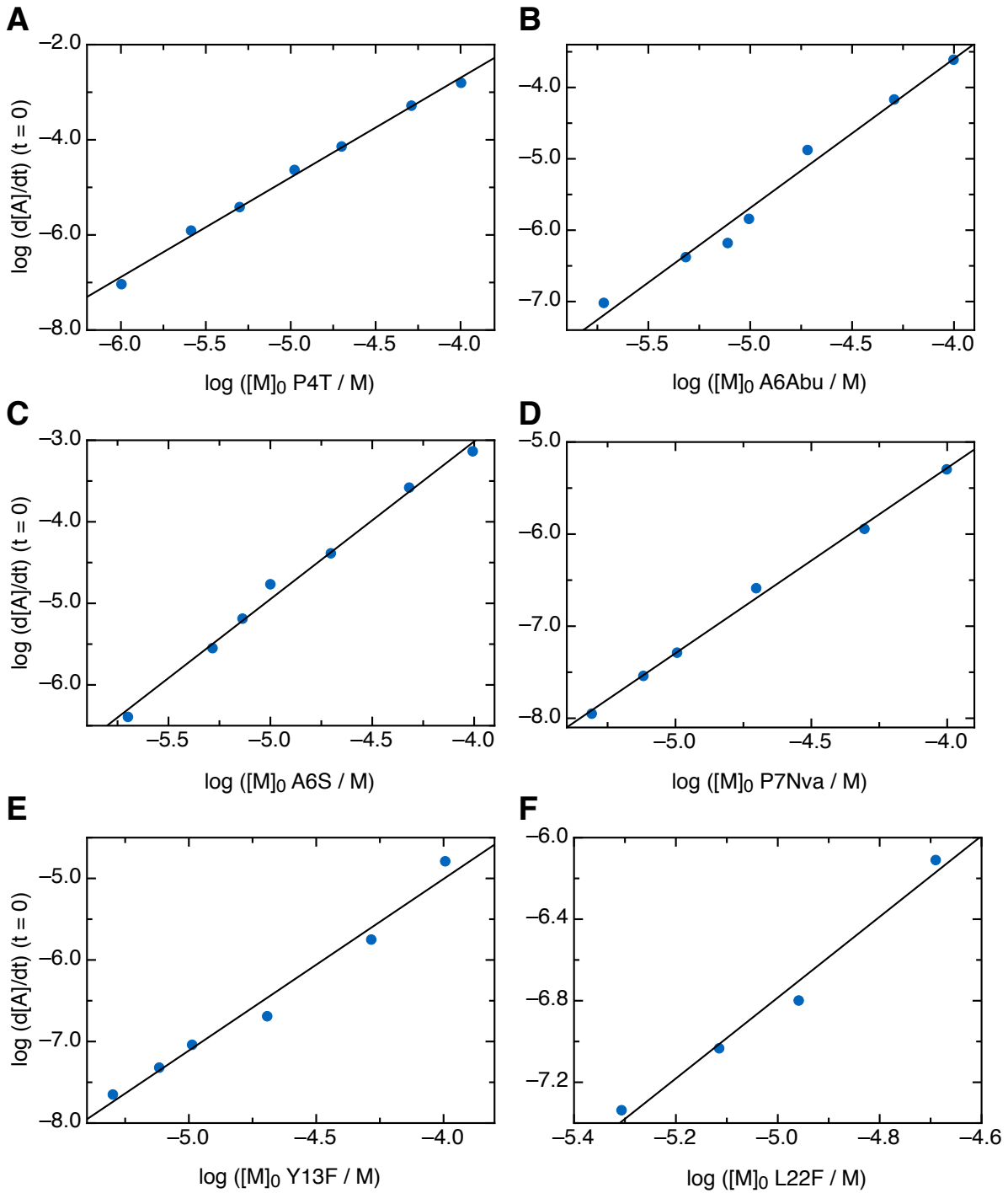


Figure A.25.: Reaction order determination of all urea refolding experiments by the initial slope determination method

A.7. Reaction order determination from the half-life times

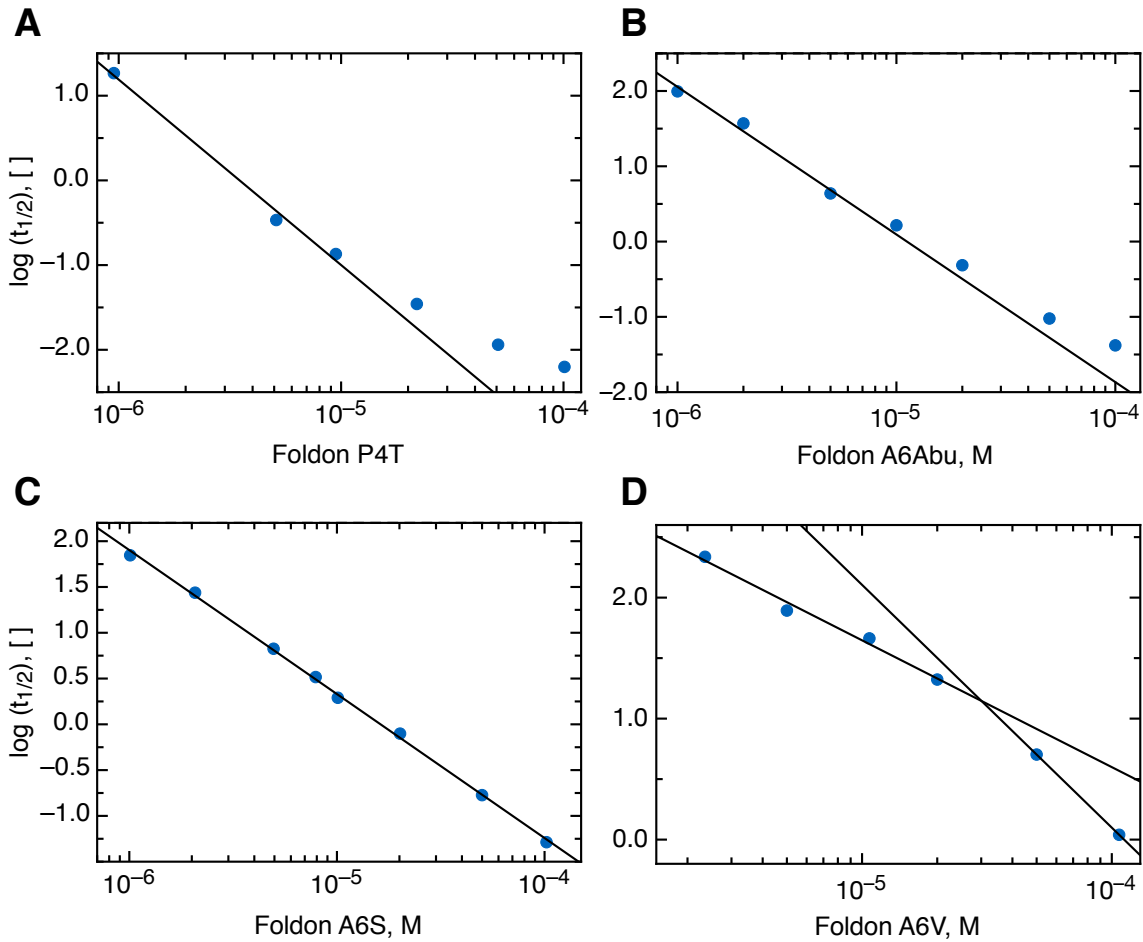


Figure A.26.: Reaction order determination of the GdmCl refolding experiments from the half-life times for the foldon variants A) P4T, B) A6Abu, C) A6S and D) A6V.

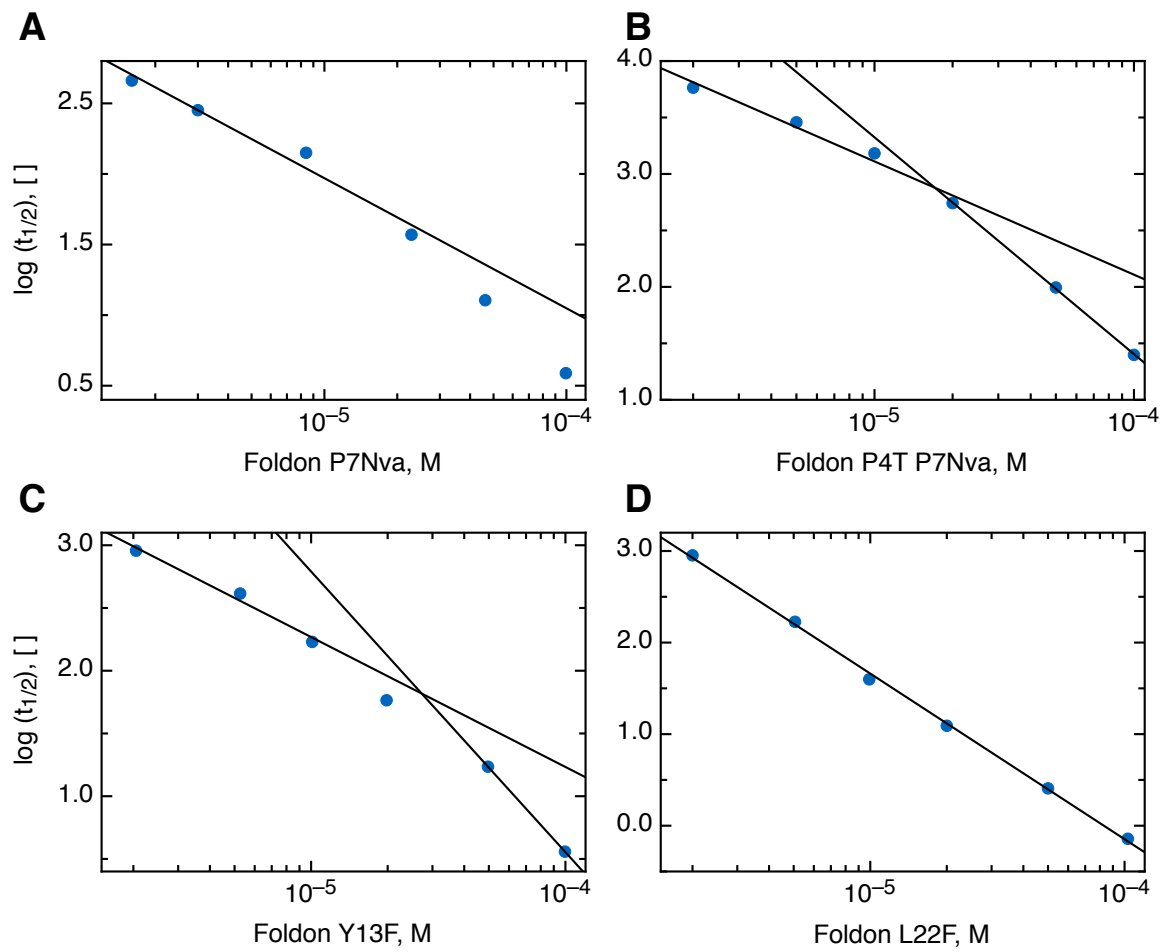


Figure A.27.: Reaction order determination of the GdmCl refolding experiments from the half-life times for the foldon variants A) P7Nva, B) P4T P7Nva, C) Y13F and D) L22F.

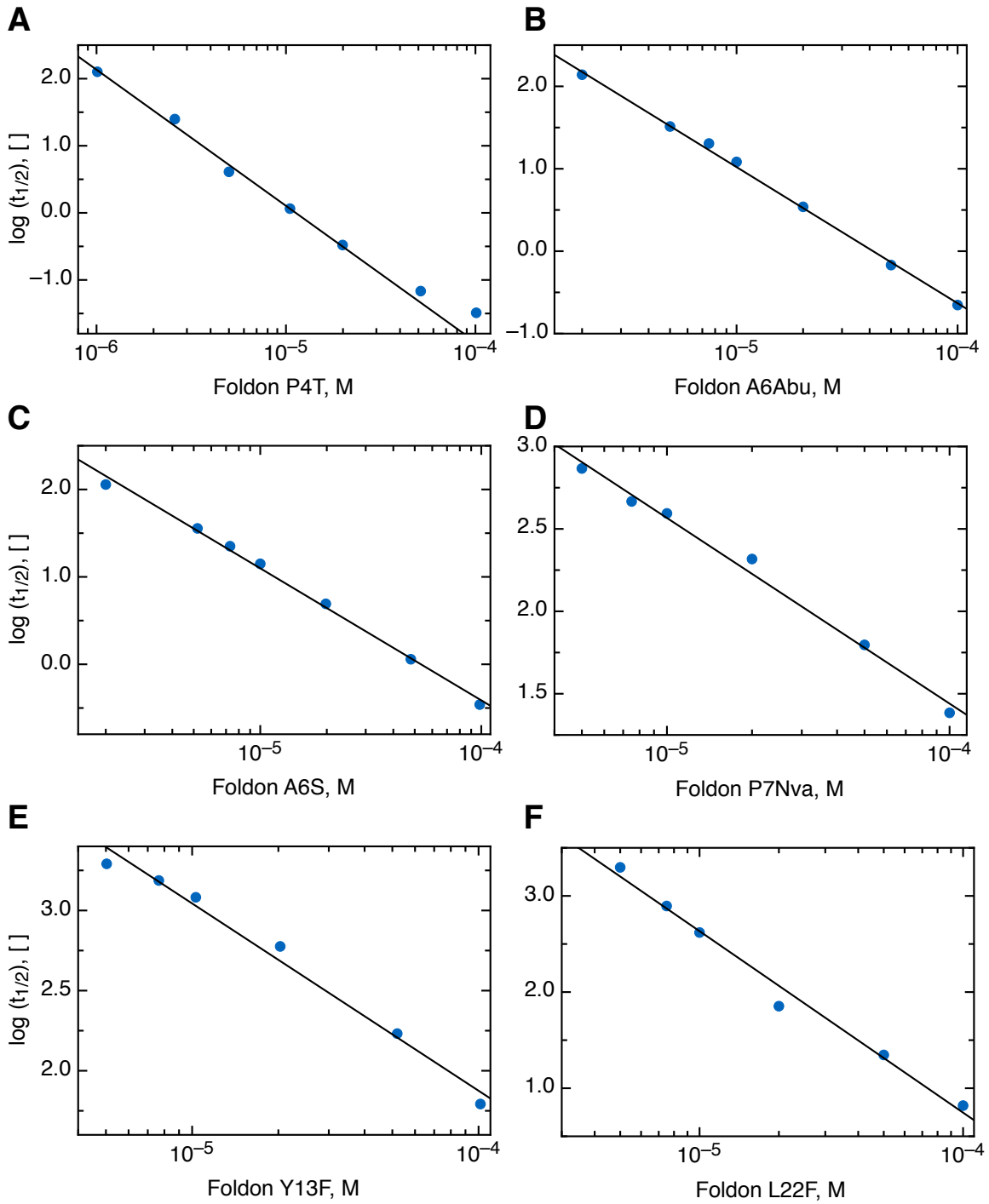


Figure A.28.: Reaction order determination of all urea refolding experiments from the half-life times

List of Figures

1.1. NMR structures of foldon wild-type	13
1.2. Foldon wild-type refolding and interrupted refolding kinetic traces	15
1.3. The foldon assembly pathway in GdmCl	15
1.4. The foldon assembly pathway in urea	16
3.1. Chemical structures of the used non-natural amino acids	22
3.2. Foldon association mechanisms	43
4.1. The hydrophobic cluster in foldon wild-type	48
4.2. Crystal structure of foldon P4T	52
4.3. Foldon P4T fluorescence spectra and GdmCl transition	53
4.4. Comparison of foldon wild-type and foldon P4T refolding traces	54
4.5. Global fit of foldon P4T GdmCl refolding kinetics	56
4.6. Global fit of foldon P4T urea refolding kinetics	57
4.7. Crystal structure of foldon A6Abu	59
4.8. Fluorescence spectra, GdmCl transition, and unfolding plots in GdmCl and urea of foldon A6Abu	60
4.9. Comparison of foldon wild-type and foldon A6Abu refolding traces	64
4.10. Global fit of foldon A6Abu GdmCl refolding kinetics	65
4.11. Global fit of foldon A6Abu urea refolding kinetics	66
4.12. Crystal structure of foldon A6S	68
4.13. Fluorescence spectra, GdmCl transition, and unfolding plots in GdmCl and urea of foldon A6S	70
4.14. Comparison of foldon wild-type and foldon A6S refolding traces	71
4.15. Global fit of foldon A6S GdmCl refolding kinetics	72
4.16. Global fit of foldon A6S urea refolding kinetics	73
4.17. Fluorescence spectra, GdmCl transition, unfolding and refolding plots in GdmCl of foldon A6V	75

List of Figures

4.18. Global fit of foldon A6V GdmCl refolding kinetics	77
4.19. Crystal structure of foldon P7F	79
4.20. GdmCl transition of foldon P7V and foldon P7F.	80
4.21. Crystal structure of foldon P7Nva	82
4.22. Fluorescence spectra, GdmCl transition, and unfolding plots in GdmCl and urea of foldon P7Nva	83
4.23. Comparison of foldon wild-type and foldon P7Nva refolding traces	84
4.24. Global fit of foldon P7Nva GdmCl refolding kinetics	85
4.25. Global fit of foldon P7Nva urea refolding kinetics	86
4.26. Crystal structure of foldon P4T P7Nva	88
4.27. Fluorescence spectra, GdmCl transition, unfolding and refolding plots in GdmCl of foldon P4T P7Nva	90
4.28. Global fit of foldon P4T P7Nva GdmCl refolding kinetics	91
4.29. Crystal structure of foldon Y13F	93
4.30. Fluorescence spectra, GdmCl transition, and unfolding plots in GdmCl and urea of foldon Y13F	95
4.31. Comparison of foldon wild-type and foldon Y13F refolding traces	96
4.32. Global fit of foldon Y13F GdmCl refolding kinetics	97
4.33. Global fit of foldon Y13F urea refolding kinetics	98
4.34. Crystal structure of foldon W20Nal	100
4.35. Foldon W20Nal fluorescence spectra and GdmCl unfolding experiments	101
4.36. Crystal structure of foldon W20H	103
4.37. Foldon W20H fluorescence spectra and GdmCl unfolding experiments	104
4.38. Normalized GdmCl transition curves of foldon wild-type and the Trp20 mutations in the presence of sodium sulfate	105
4.39. Normalized urea transition curves of the Trp20 mutations in the presence of sodium sulfate	107
4.40. Comparison of foldon W20H GdmCl transitions in the presence of sodium sulfate at varying pH-values	108
4.41. Crystal structure of foldon L22F	110
4.42. Fluorescence spectra, GdmCl transition, and unfolding plots in GdmCl and urea of foldon L22F	112
4.43. Comparison of foldon wild-type and foldon L22F refolding traces	113
4.44. Global fit of foldon L22F GdmCl refolding kinetics	114

4.45. Global fit of foldon L22F urea refolding kinetics	116
4.46. <i>N</i> -test of foldon L22F	117
4.47. Equilibrium <i>N</i> -test of foldon variants	122
5.1. Hydrogen bonds in the hydrophobic cluster	141
A.1. Normalized GdmCl transition curve of foldon P4T	161
A.2. Normalized GdmCl transition curves of all three alanine 6 substitutions	162
A.3. Normalized GdmCl transition curves of all three proline 7 substitutions	163
A.4. Normalized GdmCl transition curve of foldon P4T P7Nva	164
A.5. Normalized GdmCl transition curve of foldon Y13F	164
A.6. Normalized GdmCl transition curve of foldon L22F	165
A.7. Normalized urea transition curves of the alanine 6 substitutions	166
A.8. Normalized urea transition curve of foldon P7Nva	167
A.9. Normalized urea transition curve of foldon Y13F	167
A.10. Normalized urea transition curve of foldon L22F	168
A.11. Foldon A6Abu GdmCl unfolding kinetics	169
A.12. Foldon A6S GdmCl unfolding kinetics	170
A.13. Foldon A6V GdmCl unfolding kinetics	171
A.14. Foldon P7Nva GdmCl unfolding kinetics	172
A.15. Foldon P4T P7Nva GdmCl unfolding kinetics	173
A.16. Foldon Y13F GdmCl unfolding kinetics	174
A.17. Foldon Y13F GdmCl unfolding kinetics	175
A.18. Foldon A6Abu urea unfolding kinetics	176
A.19. Foldon A6Su urea unfolding kinetics	177
A.20. Foldon P7Nva urea unfolding kinetics	178
A.21. Foldon Y13F urea unfolding kinetics	179
A.22. Foldon L22F urea unfolding kinetics	180
A.23. GdmCl refolding reaction order determination from the initial slope I	181
A.24. GdmCl refolding reaction order determination from the initial slope II	182
A.25. Urea refolding reaction order determination from the initial slope	183
A.26. GdmCl refolding reaction order determination from the half-life times I	185
A.27. GdmCl refolding reaction order determination from the half-life times II	186
A.28. Urea refolding reaction order determination from the half-life times	187

List of Tables

1.1. Foldon wild-type amino-acid sequence	12
3.1. Foldon primer sequences	24
3.2. Mixture for the quick-change mutagenesis PCR	25
3.3. PCR temperature program	25
3.4. Restriction enzyme test pipetting scheme	26
3.5. pH-values of the different stock solutions for crystallization of the foldon variants	28
3.6. Crystallization conditions of all obtained crystals for all foldon variants . . .	29
3.7. Emission wavelengths used for burst-phase intermediate spectrum determination	31
3.8. Fluorimeter refolding data point time intervals	37
4.1. All studied foldon variants	47
4.2. All free folding enthalpies of the foldon W20 variants	106
4.3. Folding stabilities of all foldon variants determined by a guanidinium chloride transition	118
4.4. Folding stabilities of foldon variants determined by urea transitions	119
4.5. Comparison of the determined free folding enthalpies in GdmCl and urea . .	120
4.6. Folding stabilities of foldon variants determined by global fitting of all transition curves	120
4.7. GdmCl unfolding experiments	123
4.8. Urea unfolding experiments	125
4.9. α_D -value determination in GdmCl	127
4.10. α_D -value determination in urea	127
4.11. Initial slope values for all GdmCl refolding experiments	128
4.12. Initial slope values for all urea refolding experiments	129
4.13. Reaction order for all GdmCl refolding experiments determined from the half-life times	130

List of Tables

4.14. Reaction order for all urea refolding experiments determined from the half-life times	130
4.15. Rate constants for GdmCl refolding experiments	133
4.16. Rate constants for urea refolding experiments	134
4.17. The k_T/k_{-D} ratio of GdmCl refolding global fits	135
4.18. The k_T/k_{-D} ratio of urea refolding global fits	135
A.1. Resolutions and backbone r.m.s.d.s of all foldon crystal structures	155
A.2. Data collection and refinement statistics I	156
A.3. Data collection and refinement statistics II	157
A.4. Data collection and refinement statistics III	158
A.5. Data collection and refinement statistics IV	159

Bibliography

- [1] Voet, D. and Voet, J. G. (2004) *Biochemistry*. John Wiley & Sons, Inc., 3rd edn.
- [2] Lewin, B. (2004) *Genes VIII*. Pearson Prentice Hall, international edn.
- [3] Ramakrishnan, V. and Moore, P. B. (2001) Atomic structures at last: The ribosome in 2000. *Current Opinion in Structural Biology*, **11**, 144–154.
- [4] Cabrita, L. D., Dobson, C. M., and Christodoulou, J. (2010) Protein folding on the ribosome. *Current Opinion in Structural Biology*, **20**, 33–45.
- [5] Wilson, D. N. and Beckmann, R. (2011) The ribosomal tunnel as a functional environment for nascent polypeptide folding and translational stalling. *Current Opinion in Structural Biology*, **21**, 274–282.
- [6] Anfinsen, C. B., Redfield, R. R., Choate, W. L., Page, J., and Carroll, W. R. (1954) Studies on the gross structure, cross-linkages, and terminal sequences in ribonuclease. *Journal of Biological Chemistry*, **207**, 201–210.
- [7] Pauling, L., Corey, R. B., and Branson, H. R. (1951) The structure of proteins. *Proceedings of the National Academy of Sciences*, **37**, 205–211.
- [8] Pauling, L. and Corey, R. B. (1951) Configurations of polypeptide chains with favored orientations around single bonds. *Proceedings of the National Academy of Sciences*, **37**, 729–740.
- [9] Leszczynski, J. and Rose, G. (1986) Loops in globular proteins: A novel category of secondary structure. *Science*, **234**, 849–855.
- [10] Branden, C. and Tooze, J. (1999) *Introduction to Protein Structure*. Garland Publishing, 2nd edn.

- [11] Scholtz, J. M., Qian, H., Robbins, V. H., and Baldwin, R. L. (1993) The energetics of ion-pair and hydrogen-bonding interactions in a helical peptide. *Biochemistry*, **32**, 9668–9676.
- [12] Kauzmann, W. (1959) Some factors in the interpretation of protein denaturation. *Advances in Protein Chemistry*, **Volume 14**, 1–63.
- [13] Goodsell, D. S. and Olson, A. J. (1993) Soluble proteins: Size, shape and function. *Trends in Biochemical Sciences*, **18**, 65–68.
- [14] Tao, Y., Strelkov, S. V., Mesyanzhinov, V. V., and Rossmann, M. G. (1997) Structure of bacteriophage T4 fibritin: a segmented coiled coil and the role of the C-terminal domain. *Structure*, **5**, 789–98.
- [15] Argos, P. (1988) An investigation of protein subunit and domain interfaces. *Protein Engineering*, **2**, 101–113.
- [16] Creighton, T. E. (1989) *Protein Function. A Practical Approach*. IRL Press.
- [17] Harding, J. J. and Crabbe, M. J. C. (eds.) (1992) *Post-Translational Modifications of Proteins*. CRC Press.
- [18] Uversky, V. N. (2011) Intrinsically disordered proteins from a to z. *The International Journal of Biochemistry & Cell Biology*, **43**, 1090–1103.
- [19] Neurath, H. (1984) Evolution of proteolytic enzymes. *Science*, **224**, 350–357.
- [20] Bochtler, M., Ditzel, L., Groll, M., Hartmann, C., and Huber, R. (1999) The proteasome. *Annual Review of Biophysics and Biomolecular Structure*, **28**, 295–317.
- [21] Carrell, R. W. and Gooptu, B. (1998) Conformational changes and disease – serpins, prions and alzheimer's. *Current Opinion in Structural Biology*, **8**, 799–809.
- [22] Prusiner, S. B. (1998) Prions. *Proceedings of the National Academy of Sciences of the United States of America*, **95**, 13363–13383.
- [23] Levinthal, C. (1968) Are there pathways for protein folding? *Journal de Chimie Physique et de Physico-Chimie Biologique*, **65**, 44–45.

-
- [24] Pace, C. N. (1975) The stability of globular proteins. *Critical Reviews in Biochemistry and Molecular Biology*, **3**, 1–43.
- [25] Stickle, D. F., Presta, L. G., Dill, K. A., and Rose, G. D. (1992) Hydrogen bonding in globular proteins. *Journal of Molecular Biology*, **226**, 1143–1159.
- [26] Waldburger, C. D., Jonsson, T., and Sauer, R. T. (1996) Barriers to protein folding: Formation of buried polar interactions is a slow step in acquisition of structure. *Proceedings of the National Academy of Sciences*, **93**, 2629–2634.
- [27] Makhatadze, G. I. and Privalov, P. L. (1996) On the entropy of protein folding. *Protein Science*, **5**, 507–510.
- [28] Atkins, P. W. (1998) *Physical Chemistry*. Oxford University Press, sixth edn.
- [29] Greenstein, J. P. (1938) Sulfhydryl groups in proteins. *Journal of Biological Chemistry*, **125**, 501–513.
- [30] Simpson, R. B. and Kauzmann, W. (1953) The kinetics of protein denaturation. I. The behavior of the optical rotation of ovalbumin in urea solutions. *Journal of the American Chemical Society*, **75**, 5139–5152.
- [31] Aune, K. C. and Tanford, C. (1969) Thermodynamics of the denaturation of lysozyme by guanidine hydrochloride. II. dependence on denaturant concentration at 25 °. *Biochemistry*, **8**, 4586–4590.
- [32] Eisenberg, H. (1976) *Biological Macromolecules and Polyelectrolytes in Solution*. Clarendon Press.
- [33] Schellman, J. A. (1987) Selective binding and solvent denaturation. *Biopolymers*, **26**, 549–559.
- [34] Schellman, J. A. (1990) A simple model for solvation in mixed solvents: Applications to the stabilization and destabilization of macromolecular structures. *Biophysical Chemistry*, **37**, 121–140.
- [35] Pace, C. N. and Tanford, C. (1968) Thermodynamics of the unfolding of β -lactoglobulin A in aqueous urea solutions between 5 and 55. *Biochemistry*, **7**, 198–208.

- [36] Pace, C. (1986) Determination and analysis of urea and guanidine hydrochloride denaturation curves. *Methods Enzymol.*, **131**, 266–280.
- [37] Greene, R. F. and Pace, C. N. (1974) Urea and guanidine hydrochloride denaturation of ribonuclease, lysozyme, α -chymotrypsin, and β -lactoglobulin. *Journal of Biological Chemistry*, **249**, 5388–5393.
- [38] Tanford, C. (1968) Protein denaturation. *Advances in Protein Chemistry*, **23**, 121–282.
- [39] Pace, C. N. and Alexander, S. S. (1971) Comparison of the denaturation of bovine β -lactoglobulins A and B and goat β -lactoglobulin. *Biochemistry*, **10**, 2738–2743.
- [40] Santoro, M. M. and Bolen, D. W. (1988) Unfolding free energy changes determined by the linear extrapolation method. 1. Unfolding of phenylmethanesulfonyl α -chymotrypsin using different denaturants. *Biochemistry*, **27**, 8063–8068.
- [41] Myers, J. K., Nick Pace, C., and Martin Scholtz, J. (1995) Denaturant m values and heat capacity changes: Relation to changes in accessible surface areas of protein unfolding. *Protein Science*, **4**, 2138–2148.
- [42] Makhatadze, G. I. (1999) Thermodynamics of protein interactions with urea and guanidinium hydrochloride. *The Journal of Physical Chemistry B*, **103**, 4781–4785.
- [43] Santoro, M. M. and Bolen, D. W. (1992) A test of the linear extrapolation of unfolding free energy changes over an extended denaturant concentration range. *Biochemistry*, **31**, 4901–4907.
- [44] Bolen, D. W. and Yang, M. (2000) Effects of guanidine hydrochloride on the proton inventory of proteins: Implications on interpretations of protein stability. *Biochemistry*, **39**, 15208–15216.
- [45] Broering, J. M. and Bommarius, A. S. (2005) Evaluation of Hofmeister effects on the kinetic stability of proteins. *The Journal of Physical Chemistry B*, **109**, 20612–20619.
- [46] Hofmeister, F. (1888) Zur Lehre von der Wirkung der Salze. *Archiv für experimentelle Pathologie und Pharmakologie*, **24**, 247–260.

- [47] Kunz, W., Henle, J., and Ninham, B. W. (2004) "Zur Lehre von der Wirkung der Salze" (about the science of the effect of salts): Franz Hofmeister's historical papers. *Current Opinion in Colloid & Interface Science*, **9**, 19–37.
- [48] Carbonnaux, C., Ries-Kautt, M., and Ducruix, A. (1995) Relative effectiveness of various anions on the solubility of acidic *Hypoderma lineatum* collagenase at pH 7.2. *Protein Science*, **4**, 2123–2128.
- [49] Boström, M., Tavares, F. W., Finet, S., Skouri-Panet, F., Tardieu, A., and Ninham, B. W. (2005) Why forces between proteins follow different Hofmeister series for pH above and below pI. *Biophysical Chemistry*, **117**, 217–224.
- [50] Baldwin, R. L. (1996) How hofmeister ion interactions affect protein stability. *Biophysical Journal*, **71**, 2056–2063.
- [51] Geisler, M., Pirzer, T., Ackerschott, C., Lud, S., Garrido, J., Scheibel, T., and Hugel, T. (2007) Hydrophobic and hofmeister effects on the adhesion of spider silk proteins onto solid substrates: An afm-based single-molecule study. *Langmuir*, **24**, 1350–1355.
- [52] Gurau, M. C., Lim, S.-M., Castellana, E. T., Albertorio, F., Kataoka, S., and Cremer, P. S. (2004) On the mechanism of the Hofmeister effect. *Journal of the American Chemical Society*, **126**, 10522–10523.
- [53] Zhang, Y., Furyk, S., Bergbreiter, D. E., and Cremer, P. S. (2005) Specific ion effects on the water solubility of macromolecules: PNIPAM and the Hofmeister series. *Journal of the American Chemical Society*, **127**, 14505–14510.
- [54] Omta, A. W., Kropman, M. F., Woutersen, S., and Bakker, H. J. (2003) Negligible effect of ions on the hydrogen-bond structure in liquid water. *Science*, **301**, 347–349.
- [55] Batchelor, J. D., Olteanu, A., Tripathy, A., and Pielak, G. J. (2004) Impact of protein denaturants and stabilizers on water structure. *Journal of the American Chemical Society*, **126**, 1958–1961.
- [56] Kiefhaber, T. (1995) *Methods in Molecular Biology, Vol. 40: Protein Stability and Folding Protocols*, pp. 313–341. Humana Press.
- [57] Buchner, J. and Kiefhaber, T. (eds.) (2005) *Protein Folding Handbook*. Wiley-VCH, 1st edn.

- [58] Ferguson, J. K. W. and Roughton, F. J. W. (1934) The direct chemical estimation of carbamino compounds of co₂ with hæmoglobin. *The Journal of Physiology*, **83**, 68–86.
- [59] IKAI, A. and TANFORD, C. (1971) Kinetic evidence for incorrectly folded intermediate states in the refolding of denatured proteins. *Nature*, **230**, 100–102.
- [60] Cowgill, R. W. (1964) Fluorescence and the structure of proteins: Iii. effects of denaturation on fluorescence of insulin and ribonuclease. *Archives of Biochemistry and Biophysics*, **104**, 84–92.
- [61] Hopkins, T. R. and Spikes, J. D. (1967) Denaturation of proteins in 8 m urea as monitored by tryptophan fluorescence: Chymotrypsin, chymotrypsinogen and some derivatives. *Biochemical and Biophysical Research Communications*, **28**, 480–484.
- [62] Greenfield, N., Davidson, B., and Fasman, G. D. (1967) The use of computed optical rotatory dispersion curves for the evaluation of protein conformation. *Biochemistry*, **6**, 1630–1637.
- [63] Quadrifoglio, F. and Urry, D. W. (1968) Ultraviolet rotatory properties of polypeptides in solution. i. helical poly-l-alanine. *Journal of the American Chemical Society*, **90**, 2755–2760.
- [64] Quadrifoglio, F. and Urry, D. W. (1968) Ultraviolet rotatory properties of polypeptides in solution. ii. poly-l-serine. *Journal of the American Chemical Society*, **90**, 2760–2765.
- [65] Fersht, A. R. (1999) *Structure and Mechanism in Protein Science: A Guide to Enzyme Catalysis and Protein Folding*. W. H. Freeman and Company.
- [66] Roder, H. and Wüthrich, K. (1986) Protein folding kinetics by combined use of rapid mixing techniques and NMR observation of individual amide protons. *Proteins: Structure, Function, and Bioinformatics*, **1**, 34–42.
- [67] Jackson, S. E. (1998) How do small single-domain proteins fold? *Folding and Design*, **3**, R81–R91.
- [68] Ikai, A., Fish, W. W., and Tanford, C. (1973) Kinetics of unfolding and refolding of proteins: Ii. results for cytochrome c. *Journal of Molecular Biology*, **73**, 165–184.

- [69] Matthews, C. R. and Ray Wu, L. G. (1987) *Effect of point mutations of the folding of globular proteins*, vol. Volume 154, pp. 498–511. Academic Press.
- [70] Jackson, S. E. and Fersht, A. R. (1991) Folding of chymotrypsin inhibitor 2. 1. evidence for a two-state transition. *Biochemistry*, **30**, 10428–10435.
- [71] Tanford, C. (1970) Protein denaturation. Part C. Theoretical models for the mechanism of denaturation. *Advances in Protein Chemistry*, **24**, 1–95.
- [72] Matouschek, A. and Fersht, A. R. (1993) Application of physical organic chemistry to engineered mutants of proteins: Hammond postulate behavior in the transition state of protein folding. *Proceedings of the National Academy of Sciences*, **90**, 7814–7818.
- [73] Jonsson, T., Waldburger, C. D., and Sauer, R. T. (1996) Nonlinear free energy relationships in arc repressor unfolding imply the existence of unstable, native-like folding intermediates. *Biochemistry*, **35**, 4795–4802.
- [74] Sanchez, I. E. and Kiefhaber, T. (2003) Hammond behavior versus ground state effects in protein folding: Evidence for narrow free energy barriers and residual structure in unfolded states. *Journal of Molecular Biology*, **327**, 867–884.
- [75] Bieri, O., Wirz, J., Hellrung, B., Schutkowski, M., Drewello, M., and Kiefhaber, T. (1999) The speed limit for protein folding measured by triplet–triplet energy transfer. *Proceedings of the National Academy of Sciences*, **96**, 9597–9601.
- [76] Krieger, F., Fierz, B., Bieri, O., Drewello, M., and Kiefhaber, T. (2003) Dynamics of unfolded polypeptide chains as model for the earliest steps in protein folding. *Journal of Molecular Biology*, **332**, 265–274.
- [77] Garel, J.-R. and Baldwin, R. L. (1973) Both the fast and slow refolding reactions of ribonuclease a yield native enzyme. *Proceedings of the National Academy of Sciences*, **70**, 3347–3351.
- [78] Colón, W., Wakem, L. P., Sherman, F., and Roder, H. (1997) Identification of the predominant non-native histidine ligand in unfolded cytochrome c. *Biochemistry*, **36**, 12535–12541.

- [79] Brandts, J. F., Halvorson, H. R., and Brennan, M. (1975) Consideration of the possibility that the slow step in protein denaturation reactions is due to *cis-trans* isomerism of proline residues. *Biochemistry*, **14**, 4953–4963.
- [80] Balbach, J. and Schmid, F. X. (2000) *Proline isomerization and its catalysis in protein folding*. Oxford University Press.
- [81] Reimer, U., Scherer, G., Drewello, M., Kruber, S., Schutkowski, M., and Fischer, G. (1998) Side-chain effects on peptidyl-prolyl *cis/trans* isomerisation. *Journal of Molecular Biology*, **279**, 449–460.
- [82] Grathwohl, C. and Wüthrich, K. (1981) NMR studies of the rates of proline *cis-trans* isomerization in oligopeptides. *Biopolymers*, **20**, 2623–2633.
- [83] Lang, K., Schmid, F. X., and Fischer, G. (1987) Catalysis of protein folding by prolyl isomerase. *Nature*, **329**, 268–270.
- [84] Kiefhaber, T., Grunert, H. P., Hahn, U., and Schmid, F. X. (1990) Replacement of a *cis* proline simplifies the mechanism of ribonuclease T1 folding. *Biochemistry*, **29**, 6475–6480.
- [85] Schmid, F. X. and Blaschek, H. (1981) A native-like intermediate on the ribonuclease A folding pathway. *European Journal of Biochemistry*, **114**, 111–117.
- [86] Arnold, U., Hinderaker, M. P., Köditz, J., Golbik, R., Ulbrich-Hofmann, R., and Raines, R. T. (2003) Protein prosthesis: A nonnatural residue accelerates folding and increases stability. *Journal of the American Chemical Society*, **125**, 7500–7501.
- [87] Plaxco, K. W., Spitzfaden, C., Campbell, I. D., and Dobson, C. M. (1996) Rapid re-folding of a proline-rich all-beta-sheet fibronectin type III module. *Proceedings of the National Academy of Sciences*, **93**, 10703–10706.
- [88] Hammond, G. S. (1955) A correlation of reaction rates. *Journal of the American Chemical Society*, **77**, 334–338.
- [89] Kim, P. S. and Baldwin, R. L. (1982) Specific intermediates in the folding reactions of small proteins and the mechanism of protein folding. *Annual Review of Biochemistry*, **51**, 459–489.

- [90] Baldwin, R. L. (1996) On-pathway versus off-pathway folding intermediates. *Folding and Design*, **1**, R1–R8.
- [91] Sosnick, T. R., Mayne, L., Hiller, R., and Englander, S. W. (1994) The barriers in protein folding. *Nat Struct Mol Biol*, **1**, 149–156.
- [92] Krantz, B. A., Mayne, L., Rumbly, J., Englander, S. W., and Sosnick, T. R. (2002) Fast and slow intermediate accumulation and the initial barrier mechanism in protein folding. *Journal of Molecular Biology*, **324**, 359–371.
- [93] Schmid, F. X. (1983) Mechanism of folding of ribonuclease A. Slow refolding is a sequential reaction via structural intermediates. *Biochemistry*, **22**, 4690–4696.
- [94] Bieri, O., Wildegger, G., Bachmann, A., Wagner, C., and Kiefhaber, T. (1999) A salt-induced kinetic intermediate is on a new parallel pathway of lysozyme folding. *Biochemistry*, **38**, 12460–12470.
- [95] Güthe, S., Kapinos, L., Möglich, A., Meier, S., Grzesiek, S., and Kiefhaber, T. (2004) Very fast folding and association of a trimerization domain from bacteriophage T4 fibritin. *Journal of Molecular Biology*, **337**, 905–915.
- [96] Fasman, G. D. (2010) *Circular Dichroism and the Conformational analysis of Biomolecules*. Springer-Verlag New York Inc.
- [97] Semisotnov, G. V., et al. (1996) Protein globularization during folding. A study by synchrotron small-angle X-ray scattering. *Journal of Molecular Biology*, **262**, 559–574.
- [98] Matagne, A. and Dobson, C. M. (1998-04-19) The folding process of hen lysozyme: a perspective from the ‘new view’. *Cellular and Molecular Life Sciences*, **54**, 363–371.
- [99] Milla, M. E. and Sauer, R. T. (1994) P22 arc repressor: Folding kinetics of a single-domain, dimeric protein. *Biochemistry*, **33**, 1125–1133.
- [100] Zitzewitz, J. A., Bilsel, O., Luo, J., Jones, B. E., and Matthews, C. R. (1995) Probing the folding mechanism of a leucine zipper peptide by stopped-flow circular dichroism spectroscopy. *Biochemistry*, **34**, 12812–12819.

- [101] Munson, M., Anderson, K. S., and Regan, L. (1997) Speeding up protein folding: Mutations that increase the rate at which rop folds and unfolds by over four orders of magnitude. *Folding and Design*, **2**, 77–87.
- [102] Schreiber, G. and Fersht, A. R. (1996) Rapid, electrostatically assisted association of proteins. *Nat Struct Mol Biol*, **3**, 427–431.
- [103] Kiel, C., Selzer, T., Shaul, Y., Schreiber, G., and Herrmann, C. (2004) Electrostatically optimized Ras-binding Ral guanine dissociation stimulator mutants increase the rate of association by stabilizing the encounter complex. *Proceedings of the National Academy of Sciences of the United States of America*, **101**, 9223–9228.
- [104] Selzer, T., Albeck, S., and Schreiber, G. (2000) Rational design of faster associating and tighter binding protein complexes. *Nat Struct Mol Biol*, **7**, 537–541.
- [105] Ward, W. H., Jones, D. H., and Fersht, A. R. (1986) Protein engineering of homodimeric tyrosyl-trna synthetase to produce active heterodimers. *Journal of Biological Chemistry*, **261**, 9576–9578.
- [106] Cole, N. and Ralston, G. B. (1992) The effects of ionic strength on the self-association of human spectrin. *Biochimica et Biophysica Acta (BBA) - Protein Structure and Molecular Enzymology*, **1121**, 23–30.
- [107] Murray, D., McLaughlin, S., and Honig, B. (2001) The role of electrostatic interactions in the regulation of the membrane association of g protein $\beta\gamma$ heterodimers. *Journal of Biological Chemistry*, **276**, 45153–45159.
- [108] Radić, Z., Kirchhoff, P. D., Quinn, D. M., McCammon, J. A., and Taylor, P. (1997) Electrostatic influence on the kinetics of ligand binding to acetylcholinesterase. *Journal of Biological Chemistry*, **272**, 23265–23277.
- [109] Wallis, R., Moore, G. R., James, R., and Kleantous, C. (1995) Protein-protein interactions in colicin E9 DNase-immunity protein complexes. 1. Diffusion-controlled association and femtomolar binding for the cognate complex. *Biochemistry*, **34**, 13743–13750.
- [110] Wendt, H., Leder, L., Härmä, H., Jelesarov, I., Baici, A., and Bosshard, H. R. (1997) Very rapid, ionic strength-dependent association and folding of a heterodimeric leucine zipper. *Biochemistry*, **36**, 204–213.

- [111] Cheng, S., Sultana, S., Goss, D. J., and Gallie, D. R. (2008) Translation initiation factor 4B homodimerization, rna binding, and interaction with poly(A)-binding protein are enhanced by zinc. *Journal of Biological Chemistry*, **283**, 36140–36153.
- [112] Zheng, X., Mueller, G. A., Cuneo, M. J., DeRose, E. F., and London, R. E. (2010) Homodimerization of the p51 subunit of HIV-1 reverse transcriptase. *Biochemistry*, **49**, 2821–2833.
- [113] Schreiber, G., Haran, G., and Zhou, H. X. (2009) Fundamental aspects of protein-protein association kinetics. *Chemical Reviews*, **109**, 839–860.
- [114] Dyson, H. J. and Wright, P. E. (2005) Intrinsically unstructured proteins and their functions. *Nat Rev Mol Cell Biol*, **6**, 197–208.
- [115] Bachmann, A., Wildemann, D., Praetorius, F., Fischer, G., and Kiefhaber, T. (2011) Mapping backbone and side-chain interactions in the transition state of a coupled protein folding and binding reaction. *Proceedings of the National Academy of Sciences*.
- [116] Zitzewitz, J. A., Ibarra-Molero, B., Fishel, D. R., Terry, K. L., and Robert Matthews, C. (2000) Preformed secondary structure drives the association reaction of GCN4-p1, a model coiled-coil system. *Journal of Molecular Biology*, **296**, 1105–1116.
- [117] Dürr, E., Jelesarov, I., and Bosshard, H. R. (1999) Extremely fast folding of a very stable leucine zipper with a strengthened hydrophobic core and lacking electrostatic interactions between helices. *Biochemistry*, **38**, 870–880.
- [118] Gloss, L. M. and Matthews, C. R. (1998) Mechanism of folding of the dimeric core domain of *Escherichia coli* Trp repressor: A nearly diffusion-limited reaction leads to the formation of an on-pathway dimeric intermediate. *Biochemistry*, **37**, 15990–15999.
- [119] Hornby, J. A. T., Luo, J.-K., Stevens, J. M., Wallace, L. A., Kaplan, W., Armstrong, R. N., and Dirr, H. W. (2000) Equilibrium folding of dimeric class μ glutathione transferases involves a stable monomeric intermediate. *Biochemistry*, **39**, 12336–12344.
- [120] Topping, T. B. and Gloss, L. M. (2004) Stability and folding mechanism of mesophilic, thermophilic and hyperthermophilic archaeal histones: The importance of folding intermediates. *Journal of Molecular Biology*, **342**, 247–260.

- [121] Jaenicke, R. (1998) Protein self-organization in vitro and in vivo: Partitioning between physical biochemistry and cell biology. *Biological Chemistry*, **379**, 237–282.
- [122] Larsen, T. A., Olson, A. J., and Goodsell, D. S. (1998) Morphology of protein-protein interfaces. *Structure (London, England : 1993)*, **6**, 421–427.
- [123] Huang, C.-W., Chen, Y.-H., Chen, Y.-H., Tsai, Y.-C., and Lee, H.-J. (2009) The interaction of Glu294 at the subunit interface is important for the activity and stability of goose δ -crystallin. *Molecular Vision*, **15**, 2358–2363.
- [124] Espinoza-Fonseca, L. M., Wong-Ramírez, C., and Trujillo-Ferrara, J. (2010) Tyr74 is essential for the formation, stability and function of *Plasmodium falciparum* triosephosphate isomerase dimer. *Archives of Biochemistry and Biophysics*, **494**, 46–57.
- [125] Boice, J. A., Dieckmann, G. R., DeGrado, W. F., and Fairman, R. (1996) Thermodynamic analysis of a designed three-stranded coiled coil. *Biochemistry*, **35**, 14480–14485.
- [126] Backmann, J., Schäfer, G., Wyns, L., and Bönisch, H. (1998) Thermodynamics and kinetics of unfolding of the thermostable trimeric adenylate kinase from the archaeon *Sulfolobus acidocaldarius*. *Journal of Molecular Biology*, **284**, 817–833.
- [127] Bjelić, S., Karshikoff, A., and Jelesarov, I. (2006) Stability and folding/unfolding kinetics of the homotrimeric coiled coil lpp-56. *Biochemistry*, **45**, 8931–8939.
- [128] Simler, B. R., Doyle, B. L., and Matthews, C. R. (2004) Zinc binding drives the folding and association of the homo-trimeric γ -carbonic anhydrase from *Methanosarcina thermophila*. *Protein Engineering Design and Selection*, **17**, 285–291.
- [129] Alley, S. C., Shier, V. K., Abel-Santos, E., Sexton, D. J., Soumillion, P., and Benkovic, S. J. (1999) Sliding clamp of the bacteriophage T4 polymerase has open and closed subunit interfaces in solution. *Biochemistry*, **38**, 7696–7709.
- [130] Hlodan, R. and Pain, R. (1995) The folding and assembly pathway of tumor necrosis factor tnf alpha, a globular trimeric protein. *European Journal of Biochemistry*, **231**, 381–387.
- [131] Jelesarov, I. and Lu, M. (2001) Thermodynamics of trimer-of-hairpins formation by the siv gp41 envelope protein. *Journal of Molecular Biology*, **307**, 637–656.

-
- [132] Marti, D. N., Bjelic, S., Lu, M., Bosshard, H. R., and Jelesarov, I. (2004) Fast folding of the HIV-1 and SIV gp41 six-helix bundles. *Journal of Molecular Biology*, **336**, 1–8.
- [133] Dürr, E. and Bosshard, H. R. (2000) Folding of a three-stranded coiled coil. *Protein Science*, **9**, 1410–1415.
- [134] Steinbacher, S., Seckler, R., Miller, S., Steipe, B., Huber, R., and Reinemer, P. (1994) Crystal structure of p22 tailspike protein: interdigitated subunits in a thermostable trimer. *Science*, **265**, 383–386.
- [135] Miller, S., Schuler, B., and Seckler, R. (1998) Phage p22 tailspike protein: Removal of head-binding domain unmasks effects of folding mutations on native-state thermal stability. *Protein Science*, **7**, 2223–2232.
- [136] Miller, S., Schuler, B., and Seckler, R. (1998) A reversibly unfolding fragment of p22 tailspike protein with native structure: The isolated β -helix domain. *Biochemistry*, **37**, 9160–9168.
- [137] Gage, M. J. and Robinson, A. S. (2003) C-terminal hydrophobic interactions play a critical role in oligomeric assembly of the p22 tailspike trimer. *Protein Science*, **12**, 2732–2747.
- [138] Gage, M. J., Zak, J. L., and Robinson, A. S. (2005) Three amino acids that are critical to formation and stability of the p22 tailspike trimer. *Protein Science*, **14**, 2333–2343.
- [139] Efimov, V. P., et al. (1994) Fibrin encoded by bacteriophage T4 gene *wac* has a parallel triple-stranded α -helical coiled-coil structure. *Journal of Molecular Biology*, **242**, 470–486.
- [140] Sobolev, B. N. and Mesyanzhinov, V. V. (1991) The *wac* gene product of bacteriophage T4 contains coiled-coil structure patterns. *Journal of Structure and Dynamics*, **8**, 953–65.
- [141] Cohen, C. and Parry, D. A. (1990) Alpha-helical coiled coils and bundles: How to design an alpha-helical protein. *Proteins*, **7**, 1–15.
- [142] Strelkov, S. V., Tao, Y., Shneider, M. M., Mesyanzhinov, V. V., and Rossmann, M. G. (1998) Structure of bacteriophage T4 fibrin M: a troublesome packing arrangement. *Acta Crystallographia D Biological Crystallography*, **54**, 805–16.

- [143] Frank, S., Kammerer, R. A., Mechling, D., Schulthess, T., Landwehr, R., Bann, J., Guo, Y., Lustig, A., Bächinger, H. P., and Engel, J. (2001) Stabilization of short collagen-like triple helices by protein engineering. *Journal of Molecular Biology*, **308**, 1081–1089.
- [144] Panchenko, A. R., Luthey-Schulten, Z., and Wolynes, P. G. (1996) Foldons, protein structural modules, and exons. *Proceedings of the National Academy of Sciences of the United States of America*, **93**, 2008–2013.
- [145] Letarov, A. V., Londer, Y. Y., Boudko, S. P., and Mesyanzhinov, V. V. (1999) The carboxy-terminal domain initiates trimerization of bacteriophage T4 fibrin. *Biochemistry (Moscow)*, **64**, 817–823.
- [146] Yang, X., Lee, J., Mahony, E. M., Kwong, P. D., Wyatt, R., and Sodroski, J. (2002) Highly stable trimers formed by human immunodeficiency virus type 1 envelope glycoproteins fused with the trimeric motif of T4 bacteriophage fibrin. *The Journal of Virology*, **76**, 4634–4642.
- [147] Bhardwaj, A., Walker-Kopp, N., Wilkens, S., and Cingolani, G. (2008) Foldon-guided self-assembly of ultra-stable protein fibers. *Protein Science*, **17**, 1475–1485.
- [148] Papanikolopoulou, K., Teixeira, S., Belrhali, H., Forsyth, V. T., Mitraki, A., and van Raaij, M. J. (2004) Adenovirus fibre shaft sequences fold into the native triple β -spiral fold when N-terminally fused to the bacteriophage T4 fibrin foldon trimerisation motif. *Journal of Molecular Biology*, **342**, 219–227.
- [149] Meier, S., Güthe, S., Kiefhaber, T., and Grzesiek, S. (2004) Foldon, the natural trimerization domain of T4 fibrin, dissociates into a monomeric A-state form containing a stable β -hairpin: Atomic details of trimer dissociation and local β -hairpin stability from residual dipolar couplings. *Journal of Molecular Biology*, **344**, 1051–1069.
- [150] Joder, K. (2011) *Intramolecular and intermolecular diffusion processes in protein folding and assembly*. Ph.D. thesis, Technische Universität München.
- [151] Habazettl, J., Reiner, A., and Kiefhaber, T. (2009) NMR structure of a monomeric intermediate on the evolutionarily optimized assembly pathway of a small trimerization domain. *Journal of Molecular Biology*, **389**, 103–114.

- [152] Atherton, E., Fox, H., Harkiss, D., and Sheppard, R. C. (1978) Application of polyamide resins to polypeptide synthesis: An improved synthesis of beta-endorphin using fluorenylmethoxycarbonylamino-acids. *Journal of the Chemical Society, Chemical Communications*, pp. 539–540.
- [153] Mutter, M., Nefzi, A., Sato, T., Sun, X., Wahl, F., and Wöhr, T. (1995) Pseudo-prolines (ψ Pro) for accessing "inaccessible" peptides. *Peptide Research*, **8**, 145–153.
- [154] Wöhr, T., Wahl, F., Nefzi, A., Rohwedder, B., Sato, T., Sun, X., and Mutter, M. (1996) Pseudo-prolines as a solubilizing, structure-disrupting protection technique in peptide synthesis. *Journal of the American Chemical Society*, **118**, 9218–9227.
- [155] Lauer, J. L., Fields, C. G., and Fields, G. B. (1994) Sequence dependence of aspartimide formation during 9-fluorenylmethoxycarbonyl solid-phase peptide synthesis. *Letters of Peptide Science*, **1**, 197–205.
- [156] Zahariev, S., Guarnaccia, C., Zanuttin, F., Pintar, A., Esposito, G., Maravicacuta, G., Krust, B., Hovanessian, A. G., and Pongor, S. (2005) Efficient synthesis and comparative studies of the arginine and N,N-dimethylarginine forms of the human nucleolin glycine/arginine rich domain. *Journal of Peptide Science*, **11**, 17–28.
- [157] Quibell, M., Owen, D., Packman, L. C., and Johnson, T. (1994) Suppression of piperidine-mediated side product formation for Asp(OBut)-containing peptides by the use of N-(2-hydroxy-4-methoxybenzyl) (Hmb) backbone amide protection. *Journal of the Chemical Society, Chemical Communications*, pp. 2343–2344.
- [158] Krause, S. (2001) *Folding and Stability of Foldon, the trimeric C-terminal domain from Bacteriophage T4 fibritin*. Master's thesis, Universität Basel.
- [159] Walker, J. M. (ed.) (2005) *The Proteomics Protocols Handbook*. Humana Press.
- [160] Collaborative (1994) The CCP4 suite: programs for protein crystallography. *Acta Crystallographica Section D*, **50**, 760–763.
- [161] Winn, M. D., et al. (2011) Overview of the CCP4 suite and current developments. *Acta Crystallographica Section D*, **67**, 235–242.
- [162] Guex, N. and Peitsch, M. C. (1997) SWISS-MODELL and the Swiss-PdbViewer: An environment for comparative protein modeling. *Electrophoresis*, **18**, 2714–23.

- [163] Schrödinger, L. L. C. (2008) The PyMOL molecular graphics system, version 1.1r1.
- [164] Edelhoch, H. (1967) Spectroscopic determination of tryptophan and tyrosine in proteins. *Biochemistry*, **6**, 1948–1954.
- [165] Johnson, K. A., Simpson, Z. B., and Blom, T. (2009) Global kinetic explorer: A new computer program for dynamic simulation and fitting of kinetic data. *Analytical Biochemistry*, **387**, 20–29.
- [166] Johnson, K. A., Simpson, Z. B., and Blom, T. (2009) Fitspace explorer: An algorithm to evaluate multidimensional parameter space in fitting kinetic data. *Analytical Biochemistry*, **387**, 30–41.
- [167] Leffler, J. E. (1953) Parameters for the description of transition states. *Science*, **117**, 340–341.
- [168] Bachmann, A. and Kiefhaber, T. (2001) Apparent two-state tendamistat folding is a sequential process along a defined route. *Journal of Molecular Biology*, **306**, 375–386.
- [169] Sánchez, I. E. and Kiefhaber, T. (2003) Evidence for sequential barriers and obligatory intermediates in apparent two-state protein folding. *Journal of Molecular Biology*, **325**, 367–376.
- [170] Berman, H. M., Westbrook, J., Feng, Z., Gilliland, G., Bhat, T. N., Weissig, H., Shindyalov, I. N., and Bourne, P. E. (2000) The Protein Data Bank. *Nucleic Acids Research*, **28**, 235–242.
- [171] Baase, W. A., Liu, L., Tronrud, D. E., and Matthews, B. W. (2010) Lessons from the lysozyme of phage T4. *Protein Science*, **19**, 631–641.
- [172] Fersht, A. R. (1987) The hydrogen bond in molecular recognition. *Trends in Biochemical Sciences*, **12**, 301–304.
- [173] Fersht, A. R., Matouschek, A., and Serrano, L. (1992) The folding of an enzyme: I. theory of protein engineering analysis of stability and pathway of protein folding. *Journal of Molecular Biology*, **224**, 771–782.
- [174] Razvi, A. and Scholtz, J. M. (2006) Lessons in stability from thermophilic proteins. *Protein Science*, **15**, 1569–1578.

- [175] Eckhardt, B., Grosse, W., Essen, L.-O., and Geyer, A. (2010) Structural characterization of a β -turn mimic within a protein–protein interface. *Proceedings of the National Academy of Sciences*, **107**, 18336–18341.
- [176] Jeffrey, G. A. and Saenger, W. (1994) *Hydrogen Bonding in Biological Systems*. Springer Verlag.
- [177] Wu, X.-H., Chen, R.-C., Gao, Y., and Wu, Y.-D. (2010) The effect of Asp-His-Ser/Thr-Trp tetrad on the thermostability of WD40-repeat proteins. *Biochemistry*, **49**, 10237–10245.
- [178] Grey, M. J., Tang, Y., Alexov, E., McKnight, C. J., Raleigh, D. P., and Palmer III, A. G. (2006) Characterizing a partially folded intermediate of the villin headpiece domain under non-denaturing conditions: Contribution of His41 to the pH-dependent stability of the N-terminal subdomain. *Journal of Molecular Biology*, **355**, 1078–1094.
- [179] Meng, J. and McKnight, C. J. (2008) Crystal structure of a pH-stabilized mutant of villin headpiece. *Biochemistry*, **47**, 4644–4650.
- [180] Rubin, G. M., et al. (2000) Comparative genomics of the eukaryotes. *Science*, **287**, 2204–2215.
- [181] MacArthur, M. W. and Thornton, J. M. (1991) Influence of proline residues on protein conformation. *Journal of Molecular Biology*, **218**, 397–412.
- [182] Williamson, M. P. (1994) The structure and function of proline-rich regions in proteins. *Biochem. J.*, **297**, 249–260.
- [183] Rath, A., Davidson, A. R., and Deber, C. M. (2005) The structure of "unstructured" regions in peptides and proteins: Role of the polyproline II helix in protein folding and recognition. *Peptide Science*, **80**, 179–185.
- [184] Siligardi, G. and Drake, A. F. (1995) The importance of extended conformations and, in particular, the PII conformation for the molecular recognition of peptides. *Biopolymers*, **37**, 281–292.
- [185] Petrella, E. C., Machesky, L. M., Kaiser, D. A., and Pollard, T. D. (1996) Structural requirements and thermodynamics of the interaction of proline peptides with profilin. *Biochemistry*, **35**, 16535–16543.

- [186] Tanaka, M. and Shibata, H. (1985) Poly(L-proline)-binding proteins from chick embryos are a profilin and a profilactin. *European Journal of Biochemistry*, **151**, 291–297.
- [187] Yu, H., Chen, J. K., Feng, S., Dalgarno, D. C., Brauer, A. W., and Schrelber, S. L. (1994) Structural basis for the binding of proline-rich peptides to SH3 domains. *Cell*, **76**, 933–945.
- [188] Chen, H. I., Einbond, A., Kwak, S.-J., Linn, H., Koepf, E., Peterson, S., Kelly, J. W., and Sudol, M. (1997) Characterization of the WW domain of human Yes-associated protein and its polyproline-containing ligands. *Journal of Biological Chemistry*, **272**, 17070–17077.
- [189] Zarrinpar, A., Bhattacharyya, R. P., and Lim, W. A. (2003) The structure and function of proline recognition domains. *Science Signaling*, **2003**.
- [190] Tiffany, M. L. and Krimm, S. (1968) New chain conformations of poly(glutamic acid) and polylysine. *Biopolymers*, **6**, 1379–1382.
- [191] Shi, Z., Woody, R. W., Kallenbach, N. R., and Rose, G. D. (2002) *Is polyproline II a major backbone conformation in unfolded proteins?*, vol. 62, pp. 163–240. Academic Press.
- [192] Goldenberg, D. P. (1999) Finding the right fold. *Nat Struct Mol Biol*, **6**, 987–990.
- [193] Sánchez, I. E. and Kiefhaber, T. (2003) Origin of unusual ϕ -values in protein folding: Evidence against specific nucleation sites. *Journal of Molecular Biology*, **334**, 1077–1085.
- [194] Cymes, G. D., Grosman, C., and Auerbach, A. (2002) Structure of the transition state of gating in the acetylcholine receptor channel pore: A ϕ -value analysis. *Biochemistry*, **41**, 5548–5555.
- [195] Monod, J., Wyman, J., and Changeux, J.-P. (1965) On the nature of allosteric transitions: A plausible model. *Journal of Molecular Biology*, **12**, 88–118.

Danksagung

Mein besonderer Dank geht an:

Prof. Dr. Thomas Kiefhaber für die wunderbare Möglichkeit, diese Arbeit in seiner Arbeitsgruppe in Basel und München anfertigen zu können, für die großzügig gewährten Freiräume und die hervorragende wissenschaftliche Betreuung.

Prof. Dr. Michael Groll für die Zusammenarbeit beim Lösen der Kristallstrukturen und die Möglichkeit, einen kleinen Einblick in die Welt der Proteinkristallographie zu erhalten.

Meine Kollegen in der Arbeitsgruppe, vor allem Dr. Annett Bachmann für die Unterstützung bei allen kinetischen Messungen an der Stopped-flow, Dr. Karin Stecher für die hilfreichen Ratschläge beim Fitten der Daten, Dr. Andreas Reiner für die Einführung in die Proteinfestphasensynthese, sowie Natalie Merk, Ursula Zinth und insbesondere Christian Nyffenegger für das ausführliche Korrekturlesen dieser Arbeit. Ich danke Priv.-Doz. Dr. Alexander Ogrodnik, Dr. Tobias Aumüller, Dr. Maria Dolores Crespo, Dr. Beat Firz, Dr. Kerstin Hoffmann-Jacobsen, Dr. Kristine Steen Jensen, Dr. Daniel Winter, Richard Kil, Marco Perna, Florian Praetorius und Sabine Rauch sowie allen ehemaligen Mitarbeitern der Arbeitsgruppe für alle Diskussionen und Anregungen. Ich danke Dr. Sarah Güthe, Matthias Berg, Peter Kämmerer, Sabine Kullick, Matthias Stecher, Ursula Seidel und Traudl Wenger für die hervorragende Unterstützung bei allen wissenschaftlichen, technischen und verwaltungstechnischen Problemen. Euch allen vielen Dank für eine wundervolle Zeit, die Freundschaft und Kollegialität.

Die Mitarbeiter der Arbeitsgruppe von Prof. Dr. Michael Groll, besonders Dr. Melissa Gräwert, Andrea Kunfermann und Astrid König für die herzliche Zusammenarbeit und Unterstützung.

Irene Geßler für all die Unterstützung und Zuversicht, sowie ihre Familie.

Meine Eltern und meinen Bruder, ohne deren Hilfe ich viele Möglichkeiten nie gehabt und viele Ziele nie erreicht hätte.

Lebenslauf

Persönliche Daten

Tobias Schümmer

geboren am 4. März 1981 in Greven

ledig

Ausbildung und Tätigkeiten

9/1987 bis 8/1991	Nikolaus-Grundschule Rüthen
9/1991 bis 6/2000	Friedrich-Spee-Gymnasium, Rüthen
6/2000	Allgemeine Hochschulreife
7/2000 bis 5/2001	Wehrdienst
10/2001 bis 9/2003	Grundstudium der Biochemie an der Universität Bielefeld
10/2003 bis 9/2006	Hauptstudium der Biochemie an der Universität Witten-Herdecke
11/2005 bis 9/2006	Diplomarbeit am Lehrstuhl für Physikalische Biochemie von Prof. Dr. M. V. Rodnina: "Role of His84 in Elongation Factor Tu in the mechanism of nucleotide exchange catalyzed by Elongation Factor Ts".
10/2006 bis heute	Dissertationsstudium / Wissenschaftlicher Mitarbeiter am Lehrstuhl für Biophysikalische Chemie von Prof. Dr. T. Kiefhaber, Universität Basel & Technische Universität München

Publikationen

Gromadski, K. B., Schümmer, T., Strømgaard, A., Knudsen, C. R., Kinzy, T. G., and Rodnina, M. V. (2007) Kinetics of the interactions between yeast elongation factors 1A and 1B α , guanine nucleotides, and aminoacyl-tRNA. *Journal of Biological Chemistry*, **282**, 35629-35637.

Schümmer, T., Gromadski, K. B., and Rodnina, M. V. (2007) Mechanism of EF-Ts-catalyzed guanine nucleotide exchange in EF-Tu: Contribution of interactions mediated by helix B of EF-Tu. *Biochemistry*, **46**, 4977-4984.

