Structural Transitions in Helical Peptides

The Influence of Water - Implications for Molecular Recognition and Protein Folding

MARTIN LIGNELL
Dissertation presented at Uppsala University to be publicly examined in Häggsalen, Ångström laboratorium, Lägerhyddsvägen 1 Polacksbacken, Uppsala, Monday, November 30, 2009 at 10:30 for the degree of Doctor of Philosophy. The examination will be conducted in English.

Abstract

Fluctuations in protein structure are vital to function. This contrasts the dominating structure-function paradigm, which connects the well-defined three-dimensional protein structure to its function. However, catalysis is observed in disordered enzymes, which lack a defined structure. Disordered proteins are involved in molecular recognition events as well. The aim of this Thesis is to describe the structural changes occurring in protein structure and to investigate the mechanism of molecular recognition.

Protein architecture is classified in a hierarchical manner, that is, it is categorized into primary, secondary, and tertiary levels. One of the major questions in biology today is how proteins fold into a defined three-dimensional structure. Some protein folding models, like the framework model, suggest that the secondary structure, like α-helices, is formed before the tertiary structure. This Thesis raises two questions: First, are structural fluctuations that occur in the protein related to the folding of the protein structure? Second, is the hierarchic classification of the protein architecture useful to describe said structural fluctuations?

Kinetic studies of protein folding show that important dynamical processes of the folding occur on the microsecond timescale, which is why time-resolved fluorescence spectroscopy was chosen as the principal method for studying structural fluctuations in the peptides. Time-resolved fluorescence spectroscopy offers a number of experimental advantages and is useful for characterizing typical structural elements of the peptides on the sub-microsecond timescale. By observing the fluorescence lifetime distribution of the fluorescent probe, which is a part of the hydrophobic core of a four-helix bundle, it is shown that the hydrophobic core changes hydration state, from a completely dehydrated to a partly hydrated hydrophobic core. These fluctuations are related to the tertiary structure of the four-helix bundle and constitute structural transitions between the completely folded four-helix bundle and the molten globule version. Equilibrium unfolding of the four-helix bundle, using chemical denaturants or increased temperature, shows that the tertiary structure unfolds before the secondary structure, via the molten globule state, which suggests a hierarchic folding mechanism of the four-helix bundle.

Fluctuations of a 12 amino acid long helical segment, without tertiary structure, involve a conformational search of different helical organizations of the backbone.

Binding and recognition of a helix-loop-helix to carbonic anhydrase occurs through a partly folded intermediate before the final tertiary and bimolecular structure is formed between the two biomolecules. This confirms the latest established theory of recognition that the binding and the folding processes are coupled for the binding molecules.

Keywords: protein dynamics, protein folding, molten globule, time-resolved fluorescence spectroscopy, CD spectroscopy, molecular recognition, structure-function paradigm

Martin Lignell, Department of Photochemistry and Molecular Science, Chemical Physics, Box 523, Uppsala University, SE-75120 Uppsala, Sweden

© Martin Lignell 2009

ISSN 1651-6214
urn:nbn:se:uu:diva-109396 (http://urn.kb.se/resolve?urn=urn:nbn:se:uu:diva-109396)
To my parents
List of Papers

This thesis is based on the following papers, which are referred to in the text by their Roman numerals.


III Lignell, M., Becker, H-C. Conformational Switching Between $3_{10}$, $\alpha$ and $\pi$-Helical States, Studied by Time-Resolved Fluorescence and CD Spectroscopy, in a 12 Amino Acid Long Peptide.


Reprints were made with permission from the respective publishers.
Contribution report

In all papers (I-IV), I have significantly contributed to the different research problems, written all manuscripts and performed all experiments.
Contents

1. Introduction........................................................................................................... 11

2. Structural dynamics of the protein....................................................................... 13
   The Protein folding problem .................................................................................. 13
   The framework model – hierarchic folding via partly folded intermediates .......... 14
   The hydrophobic collapse model ........................................................................ 15
   Folding through multiple pathways – the nucleation condensation model .......... 16
   Folding funnels .................................................................................................... 17
   The role of water .................................................................................................. 18

How does the protein structure fluctuate?................................................................. 19
   Definition of the problem .................................................................................... 20
   Structural dynamics of the polypeptide chain ...................................................... 21
   Structure and dynamics of the denatured state .................................................... 21
   Structural Dynamics of the helical backbone ...................................................... 23
   Dynamical properties of the protein-structure mediate recognition and binding to a target molecule ................................................................. 25

3. Choice of methods for studying structural fluctuations in proteins................. 26
   Probing the local water content in protein conformations with Dansyl fluorescence ......................................................................................................................... 27
   Time-resolved fluorescence spectroscopy ............................................................. 30
      Time-correlated single photon counting (TCSPC) ................................................ 31
      Analysis of the fluorescence decay ................................................................... 32
      Measuring the rotational motion of the fluorophore .......................................... 34
      Förster energy transfer – a ruler to measure distances in proteins .................. 35
   Resolving different backbone structures with Circular Dichroism .................... 37
      Singular value decomposition of CD spectra ..................................................... 38

4. Choice of polypeptide systems for studying structural dynamics .................... 39
   The four-helix bundle peptide .............................................................................. 40
   Recognition of the helix-loop-helix to carbonic anhydrase .................................. 41
   The 12 amino acid long helical peptide – WLL12 ............................................. 42

5. Results.................................................................................................................. 43
   The four-helix bundle peptide (paper I) ............................................................... 43
Structural dynamics of the four-helix bundle ........................................... 43
Structural transitions and their dependence on temperature ................. 47
Chemically denatured four-helix bundle conformations (paper II) .......... 50
  Equilibrium denaturation as followed by steady state spectroscopy
  shows sequential unfolding of (KE2D15)$_2$........................................ 51
  Equilibrium denaturation as followed by time-resolved fluorescence
  spectroscopy reveals structural transitions corresponding to folding /
  unfolding steps of (KE2D15)$_2$......................................................... 52
The conformational backbone space of the $\alpha$-helical peptide (Paper III)
.............................................................................................................. 59
  The peptide end-to-end distances reflects a correlated reorganization of
  the WLL12 backbone at different solvent conditions ............................ 59
  The FRET decay of WLL12 and SVD analysis of the CD spectra
  reveals four backbone conformations in equilibrium.......................... 63
A coupled folding and binding funnel leads to recognition between the
  helix-loop-helix and human carbonic anhydrase (Paper IV) ............... 68
  Dansyl probes new and folded tertiary structure as a result of binding
  ............................................................................................................. 68
  The most important interaction conformations of the helix-loop-helix
  can be seen from the HCAII surface through energy transfer.............. 73
6. Conclusions.............................................................................................. 76
  Pathways of structural changes on the tertiary level......................... 76
    Tertiary structural transitions may involve structural changes on the
    secondary level as well, depending on the solvent conditions .......... 76
  Pathways of structural changes on the secondary level .................... 77
  Which folding model is supported by the structural dynamics observed in
  the helical peptides?............................................................................ 78
  The function of the disorganized helix-loop-helix structure in connection
  with molecular recognition .................................................................... 79
7. Acknowledgements.................................................................................. 81
8. Summary in Swedish................................................................................ 82
9. References............................................................................................... 84
## Abbreviations

<table>
<thead>
<tr>
<th>Abbreviation</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>CD</td>
<td>Circular Dichroism</td>
</tr>
<tr>
<td>FRET</td>
<td>Förster resonance energy transfer</td>
</tr>
<tr>
<td>FWHM</td>
<td>Full width half maximum</td>
</tr>
<tr>
<td>HCAII</td>
<td>Human carbonic anhydrase II</td>
</tr>
<tr>
<td>KE2D15</td>
<td>Helix-loop-helix monomer (Karin Enander 2) with dansyl labeled on lysine 15</td>
</tr>
<tr>
<td>KE2D15-8</td>
<td>Helix-loop-helix monomer modified with dansyl on lysine 15 and benzenesulfonamide on lysine 34</td>
</tr>
<tr>
<td>OPA</td>
<td>Optical parametric amplifier</td>
</tr>
<tr>
<td>SVD</td>
<td>Single value decomposition</td>
</tr>
<tr>
<td>TCSPC</td>
<td>Time-correlated single photon counting</td>
</tr>
<tr>
<td>TFE</td>
<td>Trifluoroethanol</td>
</tr>
<tr>
<td>TICT</td>
<td>Twisted intramolecular charged separated state</td>
</tr>
<tr>
<td>WLL12</td>
<td>12 amino acid long helical peptide</td>
</tr>
</tbody>
</table>
1. Introduction

Proteins effectuate numerous tasks inside the cell. The three-dimensional structure of the protein is considered to be intimately connected with its function, as suggested by the central dogma of molecular biology. The central dogma of molecular biology\(^1\) says that the genetic code is the blueprint for the protein architecture. The protein structure is generally categorized in three hierarchic levels: \(^{i)}\) the primary structure, which is the amino acid sequence of the protein, is directly transferred from the genetic code. \(^{ii)}\) The secondary structure refers to organization of the polypeptide backbone of the protein. In its most developed forms, like the \(\alpha\)-helix or the \(\beta\)-sheet, the protein secondary structure is connected to the formation of internal hydrogen bonds between the carbonyl and amide-groups of the backbone. \(^{iii)}\) Different secondary structural elements can interact and form the tertiary structure. The tertiary level of the protein structure consists of a variety of motifs, e.g., the four-helix bundle \((\alpha_4)\). The four-helix bundle has \(\alpha\)-helical backbone structure where hydrophobic sidechains of the different helices interact and bring the four helices together via the hydrophobic core, which is the tertiary structural element.

How the tertiary structure is transcribed from the amino acid sequence is not fully understood, and is referred to as the protein-folding problem (Chapter 2). Contributions to the protein-folding mechanism are presented in this thesis, based on characterization of structural transitions in helical peptides.

The crystal structures of enzymes show a defined architecture on all structural levels, and defined structure appears to be significant for the enzyme’s function. Despite this, high enzymatic activity has been observed in a molten globule enzyme\(^2\), which challenges the structure–function paradigm.\(^3\) The molten globule is, as the name suggests, a globular protein structure, like a four-helix bundle, but with a less defined tertiary structure. The molten globule structure is proposed to be a folding intermediate with a rapidly fluctuating hydrophobic core.\(^4\); \(^5\) This finding proposes, in contradiction with the central dogma of molecular biology, that fluctuations in protein structure may be as important as the protein structure itself. The fluctuations of the protein structure, in its natural aqueous environment, is the main topic of this thesis, however, connections to the folding problem and disorganized proteins cannot be avoided, since they are so intimately linked to structural dynamics of polypeptides. What do these fluctuations look like? What parts of the protein structure fluctuate, and on what time-scale do these fluctuations
occur? These are the basic questions which this Thesis addresses. The hier-
archic architecture of the protein structure in general will be utilized as in-
gress towards the answers of these questions, and also influence the choice
of molecular systems (Chapter 4) and methodology (Chapter 3).

The central dogma of molecular biology encounters even more problems
in relation to so called disorganized proteins\textsuperscript{6-8}, which have important recog-
nition and signal-transduction functions in the cell.\textsuperscript{9, 10} What kind of struc-
tures are present in the intrinsically disordered proteins? Do they possess
tertiary structure, secondary structure or are they structure-less random
coils? Having characterized the structural content of a disordered polypep-
tide next question arises: how is the disordered protein structure functional in
molecular recognition events? Before attacking the mechanism of molecular
recognition and the description of fluctuations in the protein structure, it is
necessary to learn the lessons from the protein-folding field.
2. Structural dynamics of the protein

The goal of this thesis is to describe the structural fluctuations that occur in the protein structure. These fluctuations span multiple timescales. In order to succeed with such a description it is necessary, apart from choosing the appropriate methodology, to take a starting point in the protein folding problem for at least two reasons:

1. Conformational fluctuations or transitions in the protein structure may be equal to a certain folding/unfolding step.\textsuperscript{11}
2. A description of the conformational fluctuations occurring in the protein's folded and unfolded states may actually be needed to unmask the folding mechanism itself.

The Protein folding problem

How the protein finds its folded structure is still an unsolved problem.\textsuperscript{12} Understanding folding is crucial for understanding the function of proteins, and for finding treatment for diseases caused by misfolded proteins, like Alzheimer's and Parkinson's diseases. Anfinsen found that the amino acid sequence of the polypeptide chain ultimately determines the structure of the protein.\textsuperscript{13} The folded structure corresponds to the thermodynamically most stable conformation of the polypeptide. However, the conformational search of an unfolded protein for its lowest energy and folded conformation would take astronomical time if the conformational space of the protein is searched randomly, which is known as the Levinthal paradox.\textsuperscript{14} A random search of the conformational space can be likened to the search of a needle in a haystack.\textsuperscript{15} Since proteins \textit{do} fold on biological timescales there has to be a mechanism, which guides the unfolded protein into its folded form. How this mechanism works has been the subject for research for decades. Here follows a brief overview of the different views and research related to the protein folding problem.
The framework model – hierarchic folding via partly folded intermediates

Since the protein structure can be classified in a hierarchy of structures (Chapter 1) it is, perhaps, intuitive to think that the protein fold develops in a hierarchic manner as well. Hierarchic folding has been suggested as a possible folding mechanism that solves the paradox of Levinthal.16 The framework model17, which suggests hierarchic folding, predicts that the protein’s secondary structure forms before the tertiary structure. The secondary structure of myoglobin is α-helical and the tertiary structure is composed of interacting hydrophobic sidechains that make up the hydrophobic core of myoglobin.17 In the hierarchical model the secondary α-helical structure is formed as a first step, by local stabilization of the helical backbone in the helical regions of the chain. Folding of myoglobin then continues to form the globular shape (the tertiary structure) by seeking non-local interactions between hydrophobic sidechains on the different helices. The formation of the tertiary structure, however, does not occur in a single step, but through a folding intermediate. As the tertiary contacts start to develop, the partly folded chains with helical segments, collapse into a molten globule. This intermediate is characterized by a well developed secondary structure but an ill-defined tertiary structure.4; 5 The most stable pathway of a biased conformational search passes via partly folded intermediates.18; 19

Experimental studies of bovine α-interferon folding showed that the α-helical structure is present before formation of the tertiary structure, which is composed of five α-helical segments and two pairs of disulfide bonds in its folded state.20 Two forms of intermediate structures, with fully developed α-helical structure but different and non-native disulfide bond pairing, were both found to be predecessors of the native fold.

The accessibility of individual backbone amide protons to exchange with solvent protons has been used as a measure of preformed secondary structure during the refolding of ribonuclease A. Experiments showed reduced ability of proton exchange during the refolding reaction, which indicates that the secondary structure precedes the tertiary structure in the folding of ribonuclease A.21

Hydrogen exchange methods (HX), as mentioned above, measure the exchange rate of the amide protons in the backbone of the polypeptide with D₂O. The amide proton can be protected from exchange if it is a part of an internal hydrogen bond, like in an α-helix, leading to a slower exchange rate compared to unprotected protons. In combination with NMR analysis it is possible to follow hydrogen exchange with amino acid resolution during the refolding process. With this type of study it has been possible to extract structural information about the folding intermediates and in what sequence they occur during refolding. Instead of a folding intermediate, the foldon
concept is introduced as a cooperative folding unit. Folding then occurs in a sequence: the development of one foldon is the starting point for the generation of the next more folded foldon, and so on. Eventually, the native and fully folded protein structure is reached.

The hydrophobic collapse model

The presence of buried non-polar amino acids in the hydrophobic core of globular proteins has lead to the formulation of the hydrophobic collapse model for folding of proteins. In this model the unfolded protein chain is believed to collapse via non-local hydrophobic interactions into a collapsed globule. Folding progresses from this intermediate to form the tertiary structure, which at the same time dictates the secondary structure. This is in contrast to the framework model, which predicts that the secondary structure governs the tertiary.

The unfolded protein chain, a random coil, collapses because its hydrophobic parts are less soluble in the surrounding water, which induces a collapse of the unfolded chain. As a consequence the number of accessible conformations get reduced for the collapsed chain. From the collapsed globule state reconfiguration of the chain can proceed, avoiding a complete search of the random coil conformations, to form the native and globular fold with defined secondary and tertiary structure. The successive reduction of accessible conformations during the collapse and the reconfiguration of the collapsed globule increase the speed of the folding reaction that circumvents Levinthal's paradox.

The coil–globule transition is connected to a shortening of the chain and has been characterized experimentally with a number of techniques. Single molecule FRET measurements on protein L showed a continuous expansion of the protein's radius of gyration with increasing concentration of denaturant, implying that folding of protein L follows a collapse mechanism. Similar measurements on the cold shock protein revealed an analogous dependence between the protein's end-to-end distance and the concentration of the denaturant. The global reconfiguration time of the cold shock protein was measured to 50 ns.

Refolding experiments of globular proteins verifies sequential folding

The processes occurring at the early stages of protein folding are of special interest. These processes seem to occur on a timescale that is faster than the time-resolution of the instrumentation often used for measuring the refolding kinetics. In many refolding experiments a spectroscopic signal is followed during time. The signal, which for example can be the intrinsic tryptophan fluorescence from the protein, is then an overall measure of the degree of
the completed folding. What sometimes is observed in these types of experiments is that a significant part of the folding reaction is already completed when the instrumentation starts to monitor the signal.\textsuperscript{31, 32} This time-unresolved part of the folding process, the “burst phase”, most probably contains crucial information about the folding mechanism that occurs within milliseconds. The framework model suggests that the local secondary structure formation continued by tertiary contact formation are the early steps of folding, whereas the hydrophobic collapse model suggests non-local tertiary contact formation with proceeding generation of the secondary structure.

Using a time-resolution of 20 $\mu$s (mixing-time of the stopped-flow instrument) it was possible to resolve two different processes in this folding regime.\textsuperscript{33} The initial process was manifested as a spectral shift that terminated within the mixing time, and was attributed to a hydrophobic collapse of the chain. This interpretation is reasonable because the tryptophan emission spectrum will be blue-shifted as the tryptophan microenvironment becomes more hydrophobic due to the collapse of the chain.\textsuperscript{34} The second observed process was an intensity decay of the blue-shifted spectrum, and was attributed to the formation of early tertiary contacts. These folding reactions were found to be the same in lysozyme, apomyoglobin and cytochrome c\textsuperscript{33}, which are all helical and globular proteins.

In cytochrome c it was found earlier\textsuperscript{35} using time-resolved CD measurements, that the nascent phase corresponds to a hydrophobic collapse reaction, and not the formation of the helical backbone structure. This phase was terminated in 100 $\mu$s. From this first intermediate $\alpha$-helical structure starts to develop. This forms the second folding intermediate, \textit{i.e.}, the molten globule, which is characterized by a fully developed secondary structure but a fluctuating tertiary structure.\textsuperscript{4, 5}

**Folding through multiple pathways – the nucleation condensation model**

This folding model predicts that secondary and tertiary structure forms concomitantly in one cooperative step. Indeed many single domain proteins fold via a two state process on the millisecond timescale.\textsuperscript{36-38} The unfolded protein chain, however, explores a large number of conformations on a much faster timescale than that of folding. Some of these conformations are on the pathway towards the folded state but they are higher in energy, which determines the folding rate.\textsuperscript{39} This is in analogy with a transition state in a chemical reaction.

The ensembles of high-energy conformations have a characteristic folding nucleus, which has a compact globular structure similar to the folded state with some essential tertiary contacts. From this transition structure folding proceeds spontaneously by rapid condensation of the unfolded parts of the
chain around the preformed folding nucleus. In this way secondary and tertiary structure form concomitantly.

The structure of the folding nucleus can be investigated by so-called ϕ analysis, i.e., by mutating an amino acid, measuring the folding rate and comparing it to that of the native protein. From this comparison inferences can be drawn about the compactness of the transient structure. For example, changing a hydrophobic amino acid in the hydrophobic core to a hydrophilic may lead to a non-compact and more denatured-like transition structure. This in turn will slow down the rate of folding. In this way important tertiary contacts of the folding nucleus can be mapped. The group of amino acids in the transient structure are designated as a foldon. Some proteins have been found to have several foldons, which leads to folding through several pathways, which enhances the global cooperativity of folding.

Folding funnels

The folding process is frequently conceptualized with a funnel. The entrance to the funnel is broad, symbolizing a broad conformational space with high energy, whereas the exit further down is much more narrower, describing a largely reduced conformational space of the folded protein, which has lower energy than the unfolded conformations (Figure 1). The funnel concept circumvents Levinthal's paradox by implying that the lowest energy and folded conformation is found through multiple pathways down to the bottom of the funnel. In this way the lowest energy conformation of the protein is found at maximum speed.

Figure 1. Cross-sections of four different folding funnels: Funnel 1 illustrates an energy landscape that causes Levinthal's paradox. There is no energy bias leading to the well-defined low-energy fold in the bottom of the funnel. Funnel 2 shows an energy-biased barrier-less funnel. Funnel 3 and 4 illustrates one or several energy barriers in the funnels.
Cooperativity of folding and ruggedness of the folding funnel

Naturally, it is of great fundamental interest to describe the conformations in the funnel in order to understand the folding mechanism.\textsuperscript{46} For instance, the length of cytochrome c was measured with time-resolved FRET during re-folding. It was found that the length of the protein collapsed in a two-state mechanism.\textsuperscript{47, 48} This behavior was also found for the three-helix bundle protein A.\textsuperscript{49} A pure two state transition like this implies that the unfolded polypeptide chain folds in an all-or-none manner, \textit{i.e.}, the folding process is cooperative. This situation is illustrated with \textit{funnel 3} in \textbf{Figure 1}, where the folding reaction has to pass over the barrier before it enters the deep, narrow part of the funnel. In contrast to \textit{funnel 3}, \textit{funnel 2} represents a barrierless folding funnel (\textbf{Figure 1}).

However, other kinetic studies, such as that of the protein barstar show that refolding is a sequential process.\textsuperscript{50} During refolding from a urea-denatured state, the unfolded chain undergoes a collapse. The degree of collapse was found to be dependent on the concentration of urea. Time-resolved measurement of the tryptophan fluorescence showed that the intensity increased gradually, which implies that the unfolded chain collapses gradually and non-cooperatively. Increasing tryptophan fluorescence intensity is an indicative that the collapsed chain gets desolvated due to the collapse. This is at variance with the steady state picture of the tryptophan fluorescence that inferred a more cooperative chain collapse.\textsuperscript{50} A sequential folding process is illustrated by \textit{funnel 4} in \textbf{Figure 1}.

The early chain collapse of barstar poses questions about what kind of structures form in the early steps of folding. This essentially boils down to the views given by the hydrophobic collapse or the framework model (see above). A later study on the refolding of barstar revealed that the collapse consists of two phases\textsuperscript{51}, where the first one was assigned to a non-specific coil–globule transition, \textit{i.e.}, the transition is not assisted by the sidechains but induced by the solvent. The second process, which was specific, involved chain contraction with help of the sidechains. This process was believed to start from the collapsed globule, but it could also occur in parallel to the first process and lead to the formation of a partly folded structure with a native-like topology. A folding mechanism involving a sequence of partly folded intermediates, but with multiple pathways, can in principle be described by \textit{funnel 4}.\textsuperscript{15}

The role of water

The core of collapsed or folded protein structures are deficient of water, but can water assist folding? As mentioned in connection with the hydrophobic collapse model, hydrophobic sidechains tend to decrease their contact with
water, which causes the formation of hydrophobic structures between these sidechains. Seen from the polypeptide's point of view, the driving force for the formation of this hydrophobic structure is unfavorable because of the reduced chain entropy, i.e., the conformational space of the hydrophobic structure is reduced compared to the fully solvated chain. From the water's perspective, however, desolvation of the polypeptide chain is entropically favorable because the released water molecules gain translational entropy. This means in short, that globular protein structures could not form if the protein were not surrounded by water.

The driving force for the association of hydrophobic particles in water shows two types of favorable interactions. In the first type the hydrophobic particles, or sidechains, are separated by one water molecule, i.e., they are in a solvent separated state, which is different from the contact state where hydrophobic particles are interacting directly. These two states of the hydrophobic interaction suggest that the tertiary interactions between different hydrophobic sidechains can be diverse in the protein structure. The presence of a solvent separated state also suggests that the tertiary structure, like hydrophobic cores, can be partially hydrated. This has been observed in a folding simulation of SH3, which showed that the folded structure was preceded by a native-like, compact, but hydrated structure.

**Barriers in the folding landscape**

During the packing of hydrophobic sidechains during folding the protein encounters enthalpic barriers which, depending on the extent of favorable entropic compensation of the folding reaction, can lead to a net free-energy barrier in the folding landscape. Desolvation barriers have been suggested to explain the cooperative nature of folding, which can occur in one or several steps.

Desolvation barriers have been studied theoretically in α-helical systems. Hydrophobic association of two poly-alanine helices passes over enthalpic barriers during the transition from the solvent separated state to the contact state of the hydrophobic sidechains. However, the release of water from the solvent separated state makes this transition entropically favored and compensates almost completely for the enthalpic effect. This results in a small free-energy barrier for the association of the two poly-alanine helices.

**How does the protein structure fluctuate?**

Theories and experiments on protein folding imply, as it is briefly summarized above, that the protein structure can reorganize itself in many different ways on different timescales. In order to understand the dynamical phenom-
ena of the protein structure, the knowledge from the protein-folding field can be useful, because it points out the directions or pathways for how the protein structure is organized under non-equilibrium conditions. These pathways may be relevant for the equilibrium fluctuations as well, since they constitute the lowest energy or kinetically accessible pathways for reorganization of the protein structure. So the question is how do these pathways look, what structural features changes and on what timescale do the changes occur?

This perspective can also be reversed: the fluctuations in the protein structure will give information about how proteins fold, which is why the protein folding problem and dynamics of the protein structure are linked together.

Several refolding studies of single domain proteins report unresolved folding-events on the sub-millisecond or microsecond timescale (see above). The key questions of these studies concerns whether folding intermediates are formed\cite{31, 32} or not.\cite{30, 38, 59} Dynamical studies of single domain proteins on these and faster timescale should thus be capable to unravel these processes, which is the aim of the thesis.

Definition of the problem

Molecular dynamics simulations have been used to decipher the protein energy landscape, which has been found to have an hierarchical organization.\cite{60} Figure 1 in reference \cite{60} shows a conceptualized folding funnel of a protein. The funnel is in this context, however, used to display what kind of stable conformations there can be depending on the available energy. These conformations are characterized by the number of native contacts in each conformation described by the parameter $Q$. Inspired by this picture I have included a similar one (Figure 2) but with a set of additional of conformational descriptors, which I believe are useful for understanding structural fluctuations of the protein structure:

1. What is the diversity of the tertiary structure in the funnel?
2. What is the diversity of the secondary structure in the funnel?
3. How does the water content affect changes between conformations?
Figure 2. How do the structural fluctuations occurring in the protein affect the tertiary, the secondary and the water content of the protein structure?

Structural dynamics of the polypeptide chain

The properties of a polypeptide's denatured state are linked to protein folding. It is of paramount interest, in relation to the protein folding problem, to know whether the denatured state has some kind of structure, or if it is completely structureless. Traditionally the denatured polypeptide has been considered to be a structureless coil that can occupy a large number of conformations, i.e., a random coil. The switching between different conformations is believed to occur on a fast timescale. It is therefore also an objective of this thesis to describe the conformational dynamics occurring in denatured polypeptides. Here follows a brief summary of what is known about the structure of the denatured state.

Structure and dynamics of the denatured state

Many chemically unfolded proteins have been found to obey random coil statistics. Single molecule spectroscopy is a powerful technique to study...
conformational changes of only one polypeptide, in contrast to standard spectroscopic methods that monitor an ensemble-averaged signal.\textsuperscript{62} For instance, the end-to-end distance can be measured during time by following the energy transfer efficiency between two chromophores attached to the termini of the polypeptide. With this technique, it has been shown that the variations in end-to-end distances are distributed according to a Gaussian function\textsuperscript{63}, and that the conformational fluctuations of the denatured cold-shock protein occur on the nanosecond timescale.\textsuperscript{62}

What kind of transient structures exist in the denatured ensemble? Loop formation has been recognized as a fast conformational transition from the random coil state\textsuperscript{64; 65} and is believed to constitute a starting point for the folding process.\textsuperscript{66}

The denatured state may not be completely unfolded even under highly denaturing conditions. Residual structure, of unknown kind, is observed in cytochrome c at 6 M GdmHCl.\textsuperscript{67} The denatured protein chain cannot obey random-coil statistics if it contains residual structure of any kind. The residual structure may also be a starting point of the folding process since it prevents the chain from moving randomly.

The amount and type of residual structure may depend on the choice of chemical denaturant. Usually GdmHCl is considered to be a stronger denaturant than urea at equal concentrations. Denaturation with urea can in some cases maintain the native topology of the protein structure, even at 8 M.\textsuperscript{68} This property of urea has been related to the two states of the hydrophobic interaction (see above). Even though the native structure is affected by urea the native topology can still be maintained if the hydrophobic interactions are changed from the contact state to the solvent separated state.\textsuperscript{69} A urea-denatured structure like that does definitely not behave as a random coil.

The big question is thus if the denatured polypeptide can be considered to be non-structured or if it contains residual structure. Structural features of the denatured state imply that there is a bias towards the folded state, which will guide the unfolded structure to the folded structure. Residual structure will in this way be functional in order to circumvent Levintahl\'s paradox of protein folding. One key aspect related to this is if the denatured protein state has any kind of secondary structure, apart from the familiar $\alpha$-helix or the $\beta$-sheet, even at highly denaturing conditions.\textsuperscript{70; 71} Any ordering of the backbone of the denatured chain will limit the conformational space of the denatured chain and may constitute a predestined pathway towards the folded state. In fact, denatured CTPR-protein (consensus tetratricopeptide repeat protein) is reported to have a polyproline-II-helix-like backbone conformation, and the average length of the denatured CTPR was found to deviate from the expected random coil average because of the polyproline-helix like organization of the denatured backbone.\textsuperscript{72}

At the present time, there seem to be contradicting conclusions about the nature of the denatured state – if it is completely randomized or not. How-
ever, it is difficult to see how folding can occur on biological timescales if there is no structural bias from the unfolded to folded structure. If the denatured state has a structure similar to a polyproline helix, how can it find its way to a folded structure? The next part addresses this question that deals with fluctuations in the helical backbone.

**Structural Dynamics of the helical backbone**

The most frequent backbone structures found in globular proteins are the β-sheet and the α-helix. There are, however, more helical backbone conformations than the α-helix that must be considered in the exploration of protein dynamics. The helical structures are the π-helix, the α-helix, the 3_10-helix, the collagen helix and the polyproline-II-like helix. Equilibrium transitions between some of these states have been observed to occur as followed in the studies below.

**α-Helix-coil structural transitions**

Traditionally the formation of α-helices is considered to be a two-state process according to the Zimm–Bragg theory\(^7_3\), which predicts that α-helical structure develops cooperatively from a coiled-state. Experimentally, the transition from helix to coil can be followed by measuring the electronic CD spectrum as the helix protein is denatured. The α-helix has a characteristic negative CD at 222 nm that goes towards less negative values when the protein is denatured, which appears as a two state process. Accordingly, it is natural to anticipate that structural dynamics of α-helices can involve conformational switching between the helical and the coiled states. The structure of the coiled state is, however, not clear. Frequently, the lack of mean-residue ellipticity at 222 nm is interpreted as a transition to the random coil state during denaturation. However, it cannot be ruled out that the coil state corresponds to one or more of the other helical backbone structures mentioned above.

Helix–coil structural dynamics has been reported for polyalanine peptides using time-resolved spectroscopy\(^7_4\). However, it is unclear from the study what the coil state is.

**α-Helix to 3_10 helix structural transitions**

Conformational transitions of an α-helix do not necessarily mean a transition to an unstructured state like a coil or a random coil. In a perfect α-helix every φ and ψ dihedral angle of the amino acids is correlated to certain values (φ = –47°, ψ = –57°), which defines the α-helical secondary structure. The random-coil state has no correlation between its dihedral angles, which
makes it structureless. The α-helix–$3_{10}$-helix conformational equilibrium is known to exist in helices containing Aib-residues (aminoisobutyric acid). Structural transitions in this equilibrium cannot be characterized as a transition between a structured and structureless state. A transition between an ordered and somewhat less ordered state would probably be a more appropriate description. The $3_{10}$ helix has different but correlated values of the dihedral angles of the backbone ($\phi = -47^\circ$, $\psi = -26^\circ$), which form an internal hydrogen bond for every third residue between the amine and the carbonyl groups. This makes the $3_{10}$-helix longer than the α-helix, which has an internal hydrogen bond for every fourth amino acid. This type of conformational transitions may be relevant for non-Aib helices as well.

### α-Helix to polyproline-II structural transitions

What is the conformational space of a single α-helix? Do certain residues in the helix switch from a helical to a coiled state, or is a conformational transition of a helix consistent with a transition to another defined secondary structure like the $3_{10}$-helix? A population of a polyproline-II like conformation is reported for small non-proline-containing peptides, which also show an α-helical population as monitored by CD spectroscopy. The polyproline-II (PPII) helix ($\phi = -75^\circ$, $\psi = 145^\circ$) has a three-fold rotational symmetry, like the $3_{10}$-helix, but contains no internal hydrogen bonds, which makes the PPII helix longer and more exposed to the solvent.

Characterization of the backbone structure of polyalanine-based peptides has been accomplished with UV resonance Raman spectroscopy that measures the $\psi$ dihedral angle of the backbone. The measured distribution of the $\psi$ dihedral angle showed values corresponding to populations of α, $3_{10}$ and PPII conformations. The interpretation of the distribution was that the residues of the polyalanine peptide had different helical states. Some residues are in a α-state, other residues are in a $3_{10}$-state or a PPII-like state, which is in contrast to an equilibrium where all residues of the peptide are in the α state for one population, and that another population of peptides are in the $3_{10}$-state and so forth. Thermal unfolding experiments of the α-helical peptide showed that the central residues were in the α-state, which unfolded with a rate-constant of $200 \text{ ns}^{-1}$ when the temperature was raised from 5 °C to 30 °C. The exterior residues, which were in a $3_{10}$-state, unfolded with doubled rate to the PPII state. The data suggest, not surprisingly, that the $3_{10}$-state can be regarded as an intermediate helical structure that is on the pathway to the PPII-state, which is definitely not a random coil.
Dynamical properties of the protein-structure mediate recognition and binding to a target molecule

Many proteins do not possess a well-defined three-dimensional structure. These so-called intrinsically disordered proteins (IDP) are involved in many cellular processes. Recognition and binding of one protein to another is a common feature that often starts the cellular process. Many theories have been proposed to explain recognition between proteins, for example the key-lock, the induced-fit and the conformational-selection models. The key-lock model anticipates complementary shapes of the binding molecules, which requires a defined structure of both ligand and receptor. In the induced-fit model, the ligand induces a conformational change of the receptor upon binding. The conformational selection theory anticipates, on the other hand, that the ligand switches conformations on a fast timescale and that one conformation, which fits the structurally defined target, is selected. All these models for molecular recognition are more or less based upon the structure-function paradigm, and they do not explain how recognition is mediated.

The fly-casting model, however, suggests that an unstructured protein-ligand binds faster to its target. Indeed, unbound protein ligands have been shown to have a disorganized structure, which becomes organized upon binding to the receptor. The disorder-to-order transition is in that respect identical with folding of the ligand to a more structured state, which links the recognition and the binding reaction to the folding problem as it is described above. This connection between recognition, binding and folding has been expressed as a coupled binding and folding funnel. It is also a task of this thesis to propose a mechanism of molecular recognition. Molecular dynamics simulations suggest that molecular recognition is a hierarchic process.
3. Choice of methods for studying structural fluctuations in proteins

The protein structures in the Protein Data Bank (http://www.pdb.org) have all been determined by X-ray crystallography or NMR spectroscopy. X-ray crystallography gives the atomic coordinates of the protein, with a resolution down to 1.2 Å. With NMR spectroscopy the coordinates of the atoms can be determined when the protein is dissolved in water, which is not possible with X-ray crystallography. NMR spectroscopy can also resolve different conformations of a protein if the protein changes its structure, but the switching of conformations must not be faster than milliseconds in order to be resolved with NMR. This timescale is often too slow to resolve all types of structural fluctuations occurring in different structural elements of the protein, as evident from the results presented in this thesis.

If there was an experimental method, like X-ray crystallography, that could determine all the atomic coordinates of the protein in a continuous series of infinitesimal time points – say every femtosecond (10^{-15} s) – then the protein folding problem would have been solved (Chapter 2). We would know if proteins fold according to a hierarchic mechanism or if they fold in an all-or-none transition as proposed by the nucleation condensation mechanism. The dynamical properties of the denatured state would also be known, as well as how proteins interact in the molecular recognition process. With this time resolution we could follow how every atom vibrates around its equilibrium position and how the dihedral angles of the protein backbone change during the folding of the protein. We could also see how the surrounding water molecules change the solvation state of the protein as the protein structure changes. This is at present time a utopia, but knowledge about how the protein structure changes on a fast timescale is paramount for understanding the functions and properties of proteins.

There are methods that can resolve fast chemical and physical processes of small molecules in time, like time-resolved optical spectroscopy. If such methods are used they can give a time resolution down to sub-picoseconds. The measured quantities from time-resolved spectroscopy do not, however, contain any direct structural information. Small fluorescent parts are sometimes present as intrinsic parts of the protein, like the indole sidechain of tryptophan, which makes it possible to measure the tryptophan fluorescence intensity at small time steps. The fluorescence lifetime, which is determined
from such measurements, can be interpreted in structural terms, since the fluorescence lifetime depends on the composition of the molecular environment proximate to tryptophan. Using this strategy, it is possible to distinguish different protein conformations from each other, but the time-resolved structures from optical spectroscopy cannot, however, be as well described as the time-averaged structures from X-ray crystallography or NMR-spectroscopy.

Probing the local water content in protein conformations with Dansyl fluorescence

The fluorescence from dansyl is sensitive to the polarity of the molecular environment proximate to Dansyl. When dansyl reaches an electronically excited state by absorbing light of 335 nm, the excited state will not live forever because other processes will immediately start to de-excite it. Fluorescence is one process that depletes the excited state through emission of light. The emitted light does not, however, have the same wavelength as the absorbing light, because the excited state loses a part of its energy by relaxation that causes the emitted light, to be of longer wavelength, which is known as **Stokes shift**. The shift between absorption and emission maximum is illustrated in Figure 3. The absorption spectrum is plotted with a solid line and has an absorption maximum at 335 nm. When a solution of dansylglycine is excited at this wavelength the dansyl emission comes in the 450-700 nm region, which is a substantial red-shift. The size of the shift depends on the polarity of the solvent and the largest shift is found for water. This photophysical property of dansyl makes it suitable as a molecular device for probing the local polarity inside the protein structure, which is divided into hydrophobic and hydrophilic parts.
Figure 3. The absorption spectrum of dansylglycine (solid line). Emission spectra of dansylglycine, excited at 335 nm, in water (open circles), methanol (open boxes) and ethanol (filled boxes).

The correlation between the solvent polarity and the emission maximum is connected to the electronic property of one of dansyl's excited states. There are in fact two excited states: when dansyl absorbs light of 335 nm it reaches first a locally excited state, which is centered on the tertiary amine group (structure 1 in Figure 4). However, this structure is not stable and transforms to a charge-separated state as a consequence of an intra-molecular electron-transfer reaction, from the amine group to the sulfur group, which creates a large dipole in the excited molecule (structure 2 in Figure 4). The charge-separated state is stabilized by twisting of the amino group, which decreases its orbital overlap with the π-system of the arene. Because of the large dipole moment (32 D), the twisted intramolecular charge separated state (TICT) interacts strongly with polar solvents and causes a red-shift of the dansyl fluorescence spectrum compared to the absorption wavelength (335 nm). When the dipole is created in the dansyl molecule, the dipoles of the surrounding solvent molecules are forced to adapt their orientation to the much larger dipole of the TICT state. As a consequence, the solvent reorientation or relaxation lowers the energy of the TICT state, which can be experimentally determined from the emission maximum of the dansyl fluorescence spectrum or the center of gravity wavelength (Equation 1).
\[
\lambda_{\text{center of gravity}} = \frac{\sum \lambda_i \cdot I(\lambda_i)}{\sum I(\lambda_i)}
\]

Equation 1

**Figure 4.** Nuclear geometries of the dansyl fluorophore in the ground and the locally excited state are displayed with structure 1. The *twisted charge separated excited state* (2) is created by an intra-molecular electron-transfer reaction, from the amine-group to the sulfur group. This causes a large dipole over the dansyl fluorophore that interacts with polar solvents. The dipole is stabilized by the twisted amine group, which has no coupling with the \( \pi \)-system of arene.

The emission center-of-gravity for dansylglycine in water (\( \varepsilon_r=80 \)) is 613 nm, 573 nm in methanol (\( \varepsilon_r=32.7 \)) and 567 nm in ethanol (\( \varepsilon_r=24.3 \)). The measured emission center-of-gravity wavelengths in these solvents show that dansyl is sensitive to the polarity of the molecular microenvironment. Accordingly, hydrophobic or hydrophilic microenvironments in proteins can be probed by dansyl fluorescence. Non-radiative decay is also dependent on solvent polarity. As seen in the fluorescence spectra (Figure 3) the fluorescence intensity decreases with increasing solvent polarity.

The electron transfer reaction from the locally excited state to the TICT state occurs in \( \approx 10 \) ps for dansylglycine in buffer (100 mM phosphate buffer pH=7.2), and \( \approx 80 \) ps for dansyl in a polypeptide environment (unpublished data). Accordingly, virtually all dansyl emission that is observed in the steady state spectra (Figure 3) originates from the dansyl's twisted charge separated state.
Table 1. Photophysical data of dansylglycine

<table>
<thead>
<tr>
<th>Solvent</th>
<th>$\varepsilon_r$</th>
<th>$\lambda_{\text{center of gravity}}$ /nm</th>
<th>$\Phi$</th>
<th>$\tau_{\text{obs}}$ /ns</th>
<th>$\text{k}_r$ /s$^{-1}$</th>
<th>$\text{k}_{nr}$ /s$^{-1}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>H$_2$O</td>
<td>80</td>
<td>613</td>
<td>0.063</td>
<td>2.4</td>
<td>2.6·10$^7$</td>
<td>3.9·10$^8$</td>
</tr>
<tr>
<td>MeOH</td>
<td>32.7</td>
<td>573</td>
<td>0.41</td>
<td>11.3</td>
<td>3.6·10$^7$</td>
<td>5.2·10$^7$</td>
</tr>
<tr>
<td>EtOH</td>
<td>24.3</td>
<td>567</td>
<td>0.48</td>
<td>12.6</td>
<td>3.8·10$^7$</td>
<td>4.1·10$^7$</td>
</tr>
</tbody>
</table>

Time-resolved fluorescence spectroscopy

The lifetime of the dansyl excited state can be determined by *time-correlated single photon counting* (Figure 5). The observed fluorescence lifetime of the dansyl TICT-state ($\tau_{\text{obs}}$) is determined by the rate constants of the radiative ($\text{k}_r$) and the non-radiative ($\text{k}_{nr}$) de-excitation pathways. In water the observed fluorescence lifetime is 2.4 ns, in methanol 11.3 ns and in ethanol 12.6 ns (Table 1). Also the fluorescence lifetime of dansyl, like the emission maximum and intensity of fluorescence, depends on the polarity of the solvent. Consequently, the dansyl fluorescence lifetime is an observable that reflects the polarity of the dansyl microenvironment. Because of the relatively short lifetime of the dansyl fluorescence, this observable can be used for resolving fluctuations of the protein structure in time if the structure changes on a slower timescale than the fluorescence lifetime. One can imagine that a structural feature, like a hydrophobic core, can become solvated at a certain time point and de-solvated at another. If the conformational fluctuations of the protein involve a change of the hydration state, these structures can be probed by the dansyl fluorescence lifetime, manifested as two different lifetimes.

The quantum yield ($\Phi$) of the dansyl fluorescence was measured, using diphenylanthracene as a reference with a known quantum yield. From the quantum yields in Table 1 the radiative and non-radiative rate constant were calculated according to Equation 2 and Equation 3. Notably, the non-radiative rate constant is higher than the radiative for dansyl in water, but decreases as the solvent polarity decreases. The radiative rate constant, which is similar to the excited state lifetime in vacuum, is rather independent of the solvent polarity. These properties of the dansyl TICT state are appropriate for probing the micropolarity in different regions of the protein structure.

\[
\text{k}_r = \frac{\Phi}{\tau_{\text{obs}}}
\]

*Equation 2*
\[ k_{nr} = \frac{1 - \Phi}{\tau_{obs}} \]

*Equation 3*

Time-correlated single photon counting (TCSPC)

The fluorescence decay of a fluorophore can be measured with a *time-correlated single photon counting* setup (*Figure 5*). From the fluorescence decay the fluorescence lifetime (\(\tau_{obs}\)) can be calculated by fitting the data to a decay law. The functioning principle behind TCSPC is that the fluorophore is excited with a light pulse of short time-duration. The next step is to wait for the excited fluorophore to emit the light back (fluorescence). As evident from *Figure 5*, a TCSPC setup consists of two parts: one laser part that generates the short light pulses (FWHM\(\approx\)100 fs), and one part that registers the time points for excitation and emission of the fluorophore.

*Figure 5.* Time-correlated single photon counting setup together with the lasers for generating light pulses with a short time duration (300 fs) of 335 nm. The TCSPC system registers the time points when a light pulse excites the sample on the photodiode and when the first emitted photon is detected on the photomultiplier.

*Figure 5* displays schematically a TCSPC. Pulsed laser light (200 kHz repetition rate) at 800 nm is generated in the seed laser. This light has low intensity and is amplified in a chirped pulse regenerative amplifier. Both the seed laser and the amplifier are pumped by a CW pump laser at 527 nm. After generating the high intensity pulses (\(\approx5\ \mu\text{J}\)) of 800 nm this pulse needs to be
transformed to laser light of 335 nm in order to excite dansyl. This occurs in two steps: first, with help of the optical parametric amplifier (OPA) 800 nm is transformed to 670 nm. The second step is external frequency doubling on a bi-refringent BBO crystal in order to generate pulses at 355 nm (≈ 0.1 nJ), which are used for dansyl excitation.

The 335 nm pulse is split in two parts: a minor part is used as a start signal for the TCSPC system, while the major part continues towards the sample, which is almost instantly excited within the duration of the pulse (≈100 fs). The 200 kHz pulsed laser system delivers pulses every 5 μs, which means there will be a long period of darkness and a relatively long time for the excited fluorophores to emit before the next excitation pulse arrives. The time point for the first emitted photon is registered on the photomultiplier (PMT). The electronics of the TCSPC system calculate the elapsed time between excitation and emission. Filters are used in front of the photomultiplier in order to select specific wavelengths in the dansyl fluorescence spectrum (Figure 3) and to block scattered excitation-light. The emitted light is also polarized to magic angle conditions (θ=54.7° from the vertically polarized excitation pulse) in the fluorescence lifetime experiments, in order to avoid intensity decays due to rotational diffusion of the fluorophore, which will be manifested as an extra fluorescence lifetime (see fluorescence anisotropy).

The time-correlated single photon counting procedure is then repeated many times and results in the fluorescence decay, which is a histogram showing the number of emission events at different time points. From this decay the fluorescence lifetime is calculated.

Analysis of the fluorescence decay

The number of excited molecules, which is proportional to the intensity of the fluorescence, decays with an observed rate-constant equal to the sum of the radiative pathway and the non-radiative pathways (Equation 4).

$$\frac{d}{dt}I = -k_{obs}I$$

Equation 4

This differential equation has one solution according to Equation 5, which is the decay law for a fluorophore in one defined microenvironment.

$$I(t) = I_0 \exp\left(-t/\tau_{obs}\right)$$

Equation 5
If the fluorophore is present in several microenvironments, such as in a conformationally flexible protein, the decay law has to be expanded by extra terms, corresponding to the number of spectroscopically different fluorophores, indicated by index $i$ (Equation 6). Analysis becomes more difficult with more lifetimes, in particular if they are of the same order of magnitude.

$$I(t) = \sum_{i} a_i \exp(-t/\tau_i)$$

*Equation 6*

A single (Equation 5) or a multi-exponential decay law (Equation 6) can be used for analyzing a fluorescence decay by fitting the parameters of the chosen decay law, $a_i$ and $\tau_i$, to the measured decay. The measured decay does not, however, have the exact shape of a decreasing exponential function at early time points. In this region the fluorescence intensity is *convoluted* with the *instrumental response function*, which means that the calculated decay needs to be convoluted with the response function before the parameters are estimated according to a least-square algorithm. As a consequence of the finite response of the detection system one observes a small rising flank in the beginning of the decay. If the sample could be excited instantaneously, with a $\delta$ pulse and the emission detected with an infinitely fast detector, the measured decay would have a true exponentially decaying shape. The instrumental response function is usually recorded using a light-scattering medium such as Ludox silica sol.

How well the decay model describes the measured data can be determined by analyzing the residuals between the fitted decay model and the measured decay. This can be performed by plotting the residuals against time and to observe the distribution of them around zero. If the residuals are randomly distributed around zero this is a good indication that the decay model describes the photophysics of the system. The measured decay also contains experimental noise and it is convenient to calculate the goodness of the fit by relating the residual between model and data to the experimental noise ($\sigma$) of the measured data (Equation 7). An ideal $\chi^2=1$ indicates perfect agreement between the data and the decay model. There might, however, be several models that fulfill the $\chi^2$-criterion, which makes it difficult to judge which model is the true model if they have similar $\chi^2$. A $\chi^2>1.4$ indicates that the decay model does not describe the measured data accurately, because the systematic errors of the model exceed the experimental noise. $\chi^2=1.05$ is a typical value of the fitted decay models presented in this thesis.

$$\chi^2 = \sum_{i} \frac{1}{\sigma_i} (N_{i,\text{predicted}} - N_{i,\text{measured}})^2$$

*Equation 7*
Maximum entropy analysis of the fluorescence decay

As mentioned previously, with complex decays it is difficult to know if the applied, multi-exponential decay law is correct, and if the discrete time constants have to be connected to spectroscopic species in the sample. Because of this difficulty, it is of interest to do a model-free analysis, which can be accomplished by a maximum entropy analysis\(^\text{92}\) of the measured fluorescence decay. Unlike the multi-discrete model (Equation 6), where it is assumed that the decay consists of a number of time constants, the maximum entropy analysis tries to maximize the number of time-constants (more precisely the entropy) to use while minimizing \(\chi^2\) (Equation 7). Such a model-free analysis of the fluorescence decay is a powerful tool to extract distributions of fluorescence lifetimes in intrinsically heterogeneous systems. The width and number of the sub-distributions correspond to the heterogeneity of the subspecies in the sample.

\[
I(t) = \int_0^\infty p(\tau) \exp\left(-t/\tau\right) d\tau
\]

*Equation 8*

Measuring the rotational motion of the fluorophore

Measurement of the fluorescence lifetime is not the only way to probe the local composition of the protein structure next to a fluorophore. Time-resolved fluorescence spectroscopy can also be used to measure the rotational diffusion of a fluorescent molecule. If the fluorophore is a part of a folded protein the rotational motion of the fluorophore is often hindered, in contrast to if solvent molecules surround the fluorophore. Accordingly, the rotational correlation time \(\theta\), can be used as an experimental observable and a characteristic of protein structure.

The rotational diffusion of a fluorophore can be measured by polarizing the excitation light vertically and measuring the emitted light both in vertical (v) and horizontal (h) polarization. The fluorescence anisotropy is calculated from these decays (Equation 9):

\[
r(t) = \frac{I_{vv}(t) - I_{vh}(t)}{I_{vv}(t) + 2I_{vh}(t)}
\]

*Equation 9*

Since the excitation light is polarized, only fluorophores with the absorbing transition moment parallel to the electric field of the excitation light are ex-
cited. This is known as photoselection (Figure 6). The depolarization of the excited fluorophores can now be followed by monitoring the emission intensities in vertical and horizontal directions in the TCSPC setup (Figure 5). Note that the rotational depolarization is measured on the emitting transition dipole moment, and hence that in some fluorophores depolarization is caused by intrinsic photophysics rather than rotational diffusion.

**Figure 6.** Photoselection of an ensemble of excited fluorophores occurs at maximum probability if the fluorophores have their transition dipole moments aligned with the polarization of the excitation light. After the excitation moment the ensemble of excited fluorophores depolarizes due to rotational diffusion of the individual fluorophores.

The measured anisotropy decay is commonly analyzed according to a decay law that describes the rotational diffusion. The model, given by Equation 10, assumes that each fluorophore can be described as a rotating sphere. How fast an ensemble of emitting fluorophores rotates is given by the rotational correlation time ($\theta_i$).

$$r(t) = \sum r_i \exp(-t/\theta_i)$$

*Equation 10*

**Förster energy transfer – a ruler to measure distances in proteins**

With fluorescence spectroscopy and time-resolved fluorescence spectroscopy it is possible to measure distances inside proteins. An excited fluorophore can deliver, non-radiatively, its energy to an accepting chromophore, if the emission spectrum of the donor overlaps with the absorption spectrum of the acceptor.\(^\text{93}\) The overlap of the absorption and emission spectra deter-
mines the overlap integral $J(\lambda)$ and is a characteristic of the chosen donor–acceptor pair. The overlap integral determines the so-called Förster distance ($R_0$), at which the energy transfer efficiency is 50%, together with orientation factor ($\kappa$) and the quantum yield of the donor ($\Phi_D$) (Equation 11). The orientation factor $\kappa$ takes into account the orientation dependence of the dipole–dipole coupling and has a value of 2/3 when the chromophores can rotate freely.

$$R_0^6 = 8.79 \cdot 10^{-5} \kappa^2 \Phi_D J(\lambda)$$

Equation 11

The efficiency of the energy transfer depends on the distance $r$ between donor and acceptor (Equation 12). The efficiency ($E$) can be calculated by measuring the intensity of the fluorescence of the donor-only-labeled protein and the emission intensity of the donor–acceptor-labeled protein (Equation 13). The efficiency can also be determined, analogously, from the fluorescence lifetime of the donor alone and the lifetime of the donor in presence of the acceptor (Equation 14). Hence, intra-molecular distances can be estimated by measuring the degree of quenching of the fluorescence.

$$E = \frac{R_0^6}{R_0^6 + r^6}$$

Equation 12

$$E = 1 - \frac{I_{DA}}{I_D}$$

Equation 13

$$E = 1 - \frac{\tau_{DA}}{\tau_D}$$

Equation 14

If the fluorescence decay of the donor comprises several lifetimes (as in Equation 6), Equation 14 cannot be used, because it is generally difficult to match a lifetime in the donor decay to the correct lifetime in the donor–acceptor decay. It is in this case possible to calculate the average energy transfer efficiency by calculating the contribution of each component to the steady-state intensity (Equation 15). Then energy transfer efficiency is calculated with Equation 13.

$$E = \frac{R_0^6}{R_0^6 + r^6}$$
\[ I_{ss} \propto \sum_i a_i \tau_i \]

*Equation 15*

**Figure 7** shows two different FRET pairs that have been used to measure distances in peptides or distances between a peptide and a protein in a complex. The small helical peptide WLL12 was labeled with tryptophan at the N-terminus and nitrotyrosine at the C-terminus. This pair has a Förster distance of 26 Å. The nitrotyrosine acceptor does not emit light when it is excited, and energy transfer efficiency can only be measured through the quenching of the donor. The tryptophan–dansyl pair was used in molecular recognition experiments where a dansyl-labeled helix-loop-helix peptide, binds to carbonic anhydrase, which has six tryptophan residues. In contrast to nitrotyrosine, dansyl emits if it is sensitized by energy transfer from tryptophan. The Förster distance of the tryptophan–dansyl pair is 23 Å. Chapter 4 describes the details of the studied molecular systems.

**Figure 7.** FRET-pairs that have been used in this work.

**Resolving different backbone structures with Circular Dichroism**

Circular dichroism (CD) was used as a complement to the time-resolved fluorescence spectroscopy. UV CD spectra reveal the conformation of the peptide backbone of proteins. CD (\(\Delta A\)) is the difference between the absorption of left-handed and right-handed circularly polarized light, which is given by Equation 16 in analogy with the Beer–Lambert law, where \(\Delta \varepsilon\) is the difference in extinction coefficient between left and right-handed circularly polarized light for the molecular species. \(c\) is the concentration of the molecular species and \(l\) is the path length.
The CD spectrum of an \( \alpha \)-helix is characterized by negative differential absorption around 222 nm, which is the \( n \rightarrow \pi^* \) transition of the peptide bond. The \( \pi \rightarrow \pi^* \) transition around 208 nm also shows negative CD with similar amplitude as the \( n \rightarrow \pi^* \) transition in a \( \alpha \)-helical backbone structure.\(^95\) Negative CD around 200 nm is also a spectral feature of other helical backbone structures, like the polyproline II helix or the \( 3_{10} \)-helix. The CD at 222 nm (the \( n \rightarrow \pi^* \) transition) of the \( 3_{10} \)-helix is slightly negative\(^96\) and slightly positive for the polyproline-II-like helix\(^97\), which makes the negative CD-absorption at 222 nm a significant signature for the \( \alpha \)-helical content of the total protein structure. Usually, the \( \alpha \)-helical content of a protein is expressed as the mean residue ellipticity, where the CD has been normalized against the number of residues.

**Singular value decomposition of CD spectra**

CD spectroscopy can be used to resolve different backbone structures in a polypeptide system. It can, for instance, be the case that the backbone structure fluctuates between a \( \alpha \)-helical state and a polyproline-helix like state. The measured CD spectrum of the “conformational mixture” will then be a linear combination of the CD spectrum for the \( \alpha \)-helix and the spectrum of the polyproline-like helix.

Singular value decomposition can be used to deconvolute the number of spectral components that are needed to describe a series of measured spectra at varying conditions, \textit{e.g.}, different temperatures, which shifts the conformational equilibrium to one or the other state.

\[
CD = USV^T
\]

\textit{Equation 17}

The matrix \( CD \) contains all the measured spectra at different conditions. Matrix \( U \) contains a number of decomposed single-value components, which have the same dimension as the measured spectra. Matrix \( S \) is a diagonal matrix, which quantifies the significance of each single-value component in \( U \). The larger the singular value, the more important the component is to describe the measured data. Matrix \( V \) describes the contribution of the components in \( U \) at the various conditions.
4. Choice of polypeptide systems for studying structural dynamics

In order to understand the dynamics of protein structure it is necessary to choose appropriate model systems that can answer the questions posed in Chapter 2. Since in general protein structure is highly complex and categorized in several levels, the model systems should represent these structural levels as well. The four-helix bundle peptide (Figure 8) has all the structural levels, i.e., the helix bundle has α-helical regions (the secondary structure), which interact with each other to form a hydrophobic core (the tertiary structure). The four-helix bundle system is suitable for studying fluctuations related to the globular protein structure. The short 12 amino acid long helical peptide (Figure 8) was chosen for studying fluctuations of a defined backbone, without tertiary interactions. A helix-loop-helix motif and the carbonic anhydrase protein were chosen for studying molecular recognition and binding between polypeptides (Figure 8).

Figure 8. Polypeptide models for studying the dynamics of a four-helix bundle peptide (1), the dynamics of recognition between HCAII and a helix-loop-helix motif (2) and structural fluctuations of the helical segment (3).
The next question that emerges in relation to the “protein structural dynamics problem” (Chapter 2) is what can and what should be measured in the chosen systems. The two possibilities associated with time-resolved fluorescence spectroscopy, outlined in Chapter 3, are to probe the water-content locally inside the protein or to measure the length between two chromophores that are labeled to the protein chain. As might be expected, the precise location where the fluorescent label is attached to the model systems has to be chosen with care in order to collect experimental data that say anything about the structural changes occurring the model system. The advantage of small peptide systems is that the typical structural elements of the models are small and, for this reason, they can be probed accurately using a single fluorescent dye. Here follows a more detailed description of the model systems.

The four-helix bundle peptide

The four-helix bundle, (KE2D15)$_2$, is a homo-dimer that consists of two helix-loop-helix monomers (Figure 9a). The KE2 de-novo design$^{98}$ is based on a repeated heptad pattern where some amino acids in the heptad have a special function in connection with the formation of the four-helix bundle motif. Amino acids 1 and 4, in the heptad, build the hydrophobic core. Residues number 2 and 5 control dimerization and residues 3 and 7 are exposed to the solvent.$^{99}$ By labeling lysine-15 with dansyl the fluorescent probe can be introduced into the hydrophobic core of the four-helix bundle peptide. Dimerization of the KE2D15-monomers occurs with high affinity ($K_D<5 \mu M$) mainly through hydrophobic interactions. Because (KE2D15)$_2$ consists of two dansyl-labeled anti-parallel monomers the whole hydrophobic core region is probed by the two dansyl fluorophores. Consequently, this should be an appropriate system for probing structural changes on the tertiary level of the protein structure.

The KE2 design originates from the structurally well characterized SA42 four-helix bundle.$^{100}$ The four-helix bundle motif has been characterized by NMR as well as analytical ultra-centrifugation, which showed the presence of four-helix bundles but no further aggregation. Some residues are exchanged in the KE2 sequence but the hydrophobic core residues are maintained.$^{98, 99}$ The four-helix bundle structure of (KE2D15)$_2$ is manifested by the mean residue ellipticity of $-20$ 000 deg cm$^2$ dmol$^{-1}$ at 222 nm, which is typical for four-helix bundle dimers.$^{100, 101}$ Despite the developed $\alpha$-helical structure, broadened NMR-shifts of the hydrophobic core shows that these four-helix bundles switch conformations on a timescale too fast for resolving discrete conformations with NMR.$^{100}$
a) **KE2 monomer sequence**

```
1   15
HELIX

42  34
LOOP -G-P-V-D-
```

**Modifications: KE2D15 & KE2D15-8**

![Benzenesulfonamide on the C8-linker](image)

b) **WLL12**

```
(NH2-) W-L-L-K-K-L-L-E-E-L-K-(YNO2) (-CONH2)
```

**Figure 9.** a) The amino sequence of the KE2 de novo design, which can be modified with dansyl on lysine 15. This modification leads to the (KE2D15)2 construct, which was used for the dynamical studies of the four-helix bundle. The KE2D15 monomer can further be modified to KE2D15-8 with the benzene-sulfonamide ligand, with a 8 carbon linker, on lysine 34 that binds to active site of carbonic anhydrase. KE2D15-8 was used for molecular recognition studies. b) The amino acid sequence of the 12 amino acid long helical peptide WLL12.

**Recognition of the helix-loop-helix to carbonic anhydrase**

The KE2 monomer (Figure 9a) is also used as a receptor for binding human carbonic anhydrase II. A KE2D15 monomer does not bind to HCAII by itself. However, if it is modified by benzenesulfonamide inhibitor on an
alkyl spacer, attached to lysine 34 on the KE2D15 monomer (KE2D15-8, which has an eight-carbon linker), the helix–loop–helix monomer binds to the HCAII surface. The functionalized scaffold greatly increases the affinity of HCAII for the ligand. Pure benzenesulfonamide has a dissociation constant of 3000 nM, while the KE2D15-8/HCAII complex has a $K_D$ of only 4 nM. This large difference shows that the helix–loop–helix receptor has affinity for HCAII in the presence of the benzenesulfonamide/linker function. The linker is needed for the benzenesulfonamide ligand to reach the active site through the 15 Å long hydrophobic cleft (Figure 8). A binding event is reported as a change in the dansyl fluorescence intensity when dansyl is close to the HCAII surface.

Structural information from NMR shows perturbed amide shifts in the KE2D15-8/HCAII complex over a rather wide range of the HCAII surface102, and it is confirmed by NMR that KE2D15-8 binds as a monomer.101

The 12 amino acid long helical peptide – WLL12

WLL12 is another de novo designed peptide (Figure 9b), which has been found to have a $\alpha$-helical backbone structure in crystals.103 The amino acid sequence is, however, modified with tryptophan and nitrotyrosine as the first and the last residues, respectively. With this system, it is possible to measure variations in the peptide's length with Förster energy transfer. Changes of the backbone structure should be reflected in the peptide's length. The advantage of the tryptophan nitrotyrosine pair is that they are natural side-chains and should therefore not “unnaturally” disturb the conformational preferences of the original $\alpha_1$ conformations.103 The second advantage is that the chromophores are directly linked to the peptide, something that should give more precise length measurements than if they were attached through longer linkers. In the reference peptide nitrotyrosine is exchanged for alanine.
5. Results

This chapter presents a summary of the results obtained by studying the model systems described in Chapter 4. Paper I describes the structural dynamics occurring in a four-helix bundle and how they depend on the temperature. Paper II explains how the four-helix bundle conformational space changes during equilibrium unfolding with urea and guanidine hydrochloride (GdmHCl). Paper III investigates the conformational space of a helical segment of a short synthetic peptide. Paper IV describes the structural transitions occurring in a molecular recognition and binding process.

The four-helix bundle peptide (paper I)

The strategy chosen for probing structural fluctuations in the (KE2D15)₂ tertiary structure is to characterize the different conformations in terms of water-content using the polarity-dependent fluorescence of dansyl. Structural dynamics of another four-helix bundle protein\textsuperscript{104} and peptides\textsuperscript{105} are known to occur on the \(\mu s\) to ms timescale. The dansyl fluorescence lifetime is about 20 ns or less, and it should thus be possible to resolve different four-helix bundle conformations and categorize them as hydrated or non-hydrated.

Structural dynamics of the four-helix bundle

The fluorescence lifetime distribution of dansyl solvents

In order to verify the applicability of the maximum entropy fluorescence lifetime analysis to the peptide system, it is important to know what the apparent width of a single-exponential decay is. Inspection of the maximum-entropy fluorescence lifetime distribution of dansylglycine in common solvents reveals one sharp distribution (Figure 10b), the position of which is dependent on the dielectric constant \(\varepsilon_r\) of the solvent (Table 1). The maximum probable lifetime of these distributions coincides perfectly with the fluorescence lifetime of a single exponential term. Thus, the broadening, relative to the distributions obtained for dansylglycine, of the fluorescence lifetime distribution in the peptide, result from intrinsic heterogeneity of the peptide structures.
Figure 10 a) Fluorescence lifetime distributions of (KE2D15)$_2$ obtained from maximum entropy analysis, are shown at different temperatures. b) Fluorescence lifetime distributions of dansylglycine in three different solvents at room temperature. c) The fluorescence lifetime of dansylglycine is practically independent of temperature, which is why the changes of the distributions of (KE2D15)$_2$ are caused by a change of the four-helix bundle as temperature increases. [KE2D15]=100 μM in all measurements.
The fluorescence lifetime distribution of dansyl inside the four-helix bundle

The fluorescence lifetime distribution at room temperature (second from bottom in Figure 10a) is more complex than the fluorescence lifetime distribution of dansylglycine in solvents (Figure 10b). In fact, it is composed of three sub-distributions with peak-maxima at 4.2, 9.9 and 20.2 ns. The whole range of dansyl fluorescence lifetimes is present in this distribution – from 4.2 ns lifetime, similar to dansylglycine in water, to 20 ns, which is close to the natural lifetime of dansyl (≈26 ns). The maximum-entropy analysis of the (KE2D15)$_2$ fluorescence decay is evidence that the dansyl fluorophores experience different microenvironments inside the four-helix bundle. Judging from the peak fluorescence lifetimes these microenvironments can be categorized as dehydrated (τ=20ns) and partly hydrated (τ=9.9 ns) dansyl microenvironments. It is, perhaps, tempting to assign the 4.2 ns component to a maximally hydrated microenvironment. However, this component vanishes at 31°C, which is why this component likely is a signature of a well-folded four-helix bundle. The exact quenching mechanism is unknown, but is apparently related to the interaction of dansyl with the residues in the hydrophobic core. Consequently, what is seen from this distribution is that there are two major groups of four-helix bundle conformations, which differ in their water content. Since the group of conformations with the 20 ns lifetime is close to the radiative lifetime, it is evident that dansyl in this conformational subgroup is a part of the hydrophobic core of the four-helix bundle peptide. The four-helix bundle conformations with an average lifetime of 9.9 ns show that this type of conformations are exposed to, or contain, water to some extent. A fully hydrated peptide environment should, however, result in a dansyl fluorescence lifetime close to the 2.4 ns observed for dansylglycine in water. Indeed, this is observed at 80°C. The only way to explain the distribution of dansyl fluorescence lifetime of the four-helix bundle, at native conditions, is that the four-helix bundle makes conformational transitions that change the hydration state of the four-helix bundle's hydrophobic core on a timescale slower than the dansyl fluorescence. This is illustrated in Figure 11.
Figure 11. The structural fluctuations of the four-helix bundle are equivalent to fluctuations in the tertiary structure, which involves solvation and desolvation dynamics of the hydrophobic core.

**What structures fluctuate in the four-helix bundle?**

As mentioned previously, it would be impossible to observe a 20 ns fluorescence lifetime unless dansyl was protected from water inside the hydrophobic core. The 9.9 ns fluorescence lifetime corresponds to an average polarity close to that of methanol. There is a significant contribution from both lifetime groups, and it can be concluded that the structural dynamics are associated with fluctuations in the hydrophobic core. Judging from the CD spectrum of (KE2D15)$_2$ at room temperature (Figure 12a), one can see that the four-helix bundle has, on average, an $\alpha$-helical backbone structure. Consequently, from the CD data and the time-resolved fluorescence data it is possible to conclude that the parts that build the hydrophobic core, namely the hydrophobic sidechains from the different helices are fluctuating. Also, the presence of two fluorescence lifetime groups suggests that the hydrophobic interactions of the hydrophobic core are at some time points in the *solvent separated state* of the hydrophobic interaction. At other time points, the partly hydrated hydrophobic core dehydrates and develops hydrophobic interactions in the *contact state* in which the dansyl lifetime is long.$^{54, 55}$

These results and conclusions are in good agreement with the NMR studies of the KE2$_2$ parent four-helix bundle, which was characterized to be a *molten globule* four-helix bundle.$^{100}$ The molten globule is known as an unstable folding intermediate, with developed secondary structure but a fluctuating tertiary structure.$^4, ^5$ Does the molten globule concept, in some way, relate to the conformations observed in the time-resolved measurements above? The conformational group represented by the 9.9 ns fluorescence time is likely a group of molten globule conformations in which the hydrophobic core is partially hydrated, but the helical backbone structure is preserved. The conformational group represented by the 20 ns lifetime would then correspond to a group of fully folded four-helix bundle conformations with dehydrated hydrophobic cores, and the observed data is consistent with
the peptide undergoing transitions between the molten-globule state and the completely folded state of the four helix bundle.

![Figure 12](image.png)

**Figure 12.** a) CD-spectra at (from bottom) 20, 40, 60 and 80°C of (KE2D15)2. Fluorescence spectra (b) at (from top) 11, 16, 26, 32, 35, 39, 44, 49, 58, 68 and 81°C. The sharp peak at 670 nm is artifact from the double excitation wavelength.

**Structural transitions and their dependence on temperature**

Increased temperature changes the structure of the four-helix-bundle. The backbone becomes less α-helical, as is evident by the CD spectra in **Figure 12a**. Notably, the tertiary structure changes as well, as monitored by the increasing emission maximum wavelength and decreasing fluorescence intensity (**Figure 12b**). Both observables from the fluorescence steady-state spectra are in agreement with more water entering the hydrophobic core as temperature rises. The data from steady state fluorescence and CD are consistent with the picture that the four-helix bundle unfolds with increased temperature, and they also give an indication on how it unfolds:
Figure 13. Upper panel: Dansyl emission center of gravity wavelength (filled squares) and mean residue helicity (filled circles) of (KE2D15)$_2$ as a function of temperature. Lower panel: Rotational correlation time of dansyl inside (KE2D15)$_2$ as function of temperature. [KE2D15]=100 μM in all measurements.

Figure 13 displays the correlation between the mean residue ellipticity and the emission center of gravity, i.e., the graph shows how secondary and the tertiary structure are both gradually destroyed. These data may even tell how the unfolding starts: in the 10–30 °C range the emission maximum is nearly constant, but the helical content decreases. This indicates that the hydrophobic core is intact up to ≈30°C, but the helical backbone becomes disorganized. Above 30 °C, the hydrophobic core starts to disintegrate as well.

The time-resolved data provides additional details about the thermal equilibrium unfolding in terms of transient conformations and their tertiary-structure content. Figure 10a reveals at least three modes of structural transitions in the following temperature intervals: 9–30 °C, 31–60 °C and 61–80 °C.

Transitions between the molten globule and the native four-helix bundle state occur in the 9–23 °C temperature region, as is evident from the constant peak maxima in this region (Table 2). However, the population of the maximally dehydrated four-helix state is increased, along with a concomitant decrease of the molten globule population, when temperature is lowered to 9
°C (Table 2), and not surprisingly low temperature seems to favor the hydrophobic interactions of the four-helix fold. Accordingly, the pathway of thermal unfolding starts from the native to the molten globule state as judged by the conformational transitions in the 9–23 °C region.

Table 2. Areas and peak-maxima of the fluorescence lifetime groups in Figure 10

<table>
<thead>
<tr>
<th>Temperature/°C</th>
<th>A₁/%</th>
<th>A₂/%</th>
<th>A₃/%</th>
<th>τ₁/ns</th>
<th>τ₂/ns</th>
<th>τ₃/ns</th>
</tr>
</thead>
<tbody>
<tr>
<td>8.9</td>
<td>5</td>
<td>37</td>
<td>58</td>
<td>4.2</td>
<td>10.6</td>
<td>21.2</td>
</tr>
<tr>
<td>23</td>
<td>5</td>
<td>59</td>
<td>36</td>
<td>4.2</td>
<td>9.9</td>
<td>20.2</td>
</tr>
<tr>
<td>30.9</td>
<td>0</td>
<td>27</td>
<td>73</td>
<td>0</td>
<td>6.4</td>
<td>16.4</td>
</tr>
<tr>
<td>40.3</td>
<td>0</td>
<td>42</td>
<td>58</td>
<td>0</td>
<td>6.8</td>
<td>16.4</td>
</tr>
<tr>
<td>51.7</td>
<td>0</td>
<td>45</td>
<td>55</td>
<td>0</td>
<td>6.7</td>
<td>15.7</td>
</tr>
<tr>
<td>59.7</td>
<td>0</td>
<td>52</td>
<td>48</td>
<td>0</td>
<td>6.8</td>
<td>15</td>
</tr>
<tr>
<td>79.0</td>
<td>0</td>
<td>60</td>
<td>40</td>
<td>0</td>
<td>5.7</td>
<td>10.9</td>
</tr>
</tbody>
</table>

Approximate areas of the three sub-distributions (A₁, A₂, A₃) of fluorescence lifetimes as percentage of the total distribution, and maximal probable lifetime of each sub-distribution at different temperatures. The areas are calculated by integration from the minimum in the trough before the peak to the minimum after the peak. Uncertainties in areas are +/-3%. Peak lifetimes are +/-5%.

The four-helix bundle opens transiently above 30 °C

At 31 °C, the shape of the fluorescence lifetime distribution is dramatically changed. The bimodal form is conserved but the peak maxima of the fluorescence lifetime groups are shifted to lower values, indicating that the structures making up the conformational equilibrium become even more hydrated. This pattern persists up to 60 °C as shown in Table 2 and Figure 10. However, the widths of the fluorescence lifetime groups become narrower, which indicates that the rate of the structural changes increases with increasing temperature. In this region (31–60 °C), a new hydrated group appears, manifested by the 6 ns peak maximum, and the population of this component increases with increasing temperature (Table 2). This indicates that dansyl, at certain time points, experiences a very water-exposed microenvironment, which suggests that a portion of the four-helix bundle dimer (α₄) dissociates into the helix-loop-helix monomer (α₂). Along with the increased population of the monomer, the emission-center of gravity becomes more red-shifted and the α-helical backbone becomes more disorganized. Although the helical content of the different conformational subgroups cannot be estimated, the correlation between the population of the monomer and the overall disorganized backbone actually indicates that the monomer conformation is less α-helical than the dimeric form.

Above 30 °C, the peak maximum of the longest fluorescence lifetime group changes from 20 ns to 16 ns. The structural difference between the 20 ns and 16 ns lifetime groups is difficult to judge with the present methods. Nevertheless, judged by the slight increase (≈ 1 nm) in emission center of gravity from 23°C to 31°C, they should both be dimeric species. The differ-
ence in fluorescence lifetimes and the increased overall rotational motion suggest, though, that the hydrophobic core may not be as well packed at 31 °C compared to the packing at 9 °C, as is shown by the decrease in rotational correlation time for dansyl (Figure 13 lower panel). Transient openings of the dimer, undoubtedly, increase the overall rotational freedom of dansyl as well, which may not only be associated with the disruptions of the tertiary interactions, but also a rearrangement of the secondary structure of the monomer upon dissociation. Relating to Figure 2, the structural transitions occurring in this temperature interval involve changes of the tertiary and secondary structures, as well as a change of the hydration state of the differently folded species in the conformational equilibrium.

The four-helix bundle dimer dissociates above 60 °C

At 79 °C, the peak maximum of the longest fluorescence lifetime group has shortened from 15 ns to 11 ns (Table 2). This significant change of the longest fluorescence lifetime is an indication that all kinds of dimeric structures have dissociated. The shortest fluorescence lifetime (5.7 ns) is becoming close to that of dansylglycine in water (2.3 ns). Thus, at 79 °C we should be close to a four-helix bundle thermally denatured state. This denatured state contains residual structure, as is evident from the two fluorescence lifetime groups. The structures of the 10.9 ns group may involve tertiary interactions within the monomer. The structures of the 5.7 ns group may, accordingly, correspond to monomer with disrupted monomeric tertiary interactions, which maximally exposes dansyl to the surrounding solvent. The backbone organization is definitely different from an α-helix. The transition of the rotational correlation time, from 2 to 1 ns, also indicates disruption of tertiary interactions in the helix-loop-helix-monomer and a relatively unhindered motion of the dansyl.

Chemically denatured four-helix bundle conformations (paper II)

Chemical denaturants like urea and GdmHCl are often used to generate the denatured state of proteins for studying the mechanism of folding. In this paper these denaturants were used to see how they affect the fast conformational transitions in the four-helix bundle, described in paper I. A description of the structural dynamics at different denaturing conditions can in principle unravel the actions of the agents.
Equilibrium denaturation as followed by steady state spectroscopy shows sequential unfolding of (KE2D15)$_2$

Unfolding of proteins can be followed by measuring a spectroscopic signal as the denaturing conditions changes, e.g., by addition of a denaturant. Intrinsic tryptophan fluorescence has been used numerous times. Here the data from the equilibrium unfolding of (KE2D15)$_2$ are presented first, as followed by steady-state dansyl-fluorescence and CD-spectroscopy. For GdmHCl, the changes in the center of gravity wavelength of the dansyl fluorescence displays a steep transition with a transition-midpoint around 2.5 M (Figure 14a). This two-state behavior of the emission center of gravity could lead to the impression that (KE2D15)$_2$ folds in an all or none-transition, because the dansyl fluorescence can be considered to measure the degree of the overall four-helix bundle structure. It will be shown below that this interpretation is incorrect. However, a minor change in emission maximum between 0 and 2 M, seems to indicate the start of the unfolding process by increasing the micropolarity around dansyl. Exactly what the structural changes consist of cannot be told from the dansyl steady state fluorescence only, but it is known, from time-resolved fluorescence, that the four-helix bundle peptide undergoes conformational changes that weaken its tertiary structure already in its native state (Paper I).

![Figure 14](image-url)

**Figure 14.** Summary of equilibrium denaturation of (KE2D15)$_2$ with GdmHCl (a) and urea (b). Dansyl emission center of gravity (filled squares) is plotted on the left axis. Mean residue ellipticity at 222 nm (open squares) is plotted on the right axis. The emission center of gravity of dansylglycine is practically independent of the concentration of GdmHCl or urea. It changes from 613–608 nm and 613–609 nm for 0–6 M of GdmHCl and urea (paper II), respectively. Accordingly, the changes in the emission center of gravity reflect changes in the four-helix bundle structure due to denaturation.
The mean residue ellipticity at 222 nm shows that the secondary α-helical structure disappears in an all-or-none step, but with a transition midpoint around 5 M GdmHCl, which is different from the 2.5 M midpoint observed for the dansyl fluorescence (Figure 14a). This is a proof of that the tertiary and the secondary structure of the four-helix bundle peptide do not form concomitantly in an all-or-none transition.

The emission center-of-gravity for equilibrium denaturation of (KE2D15)_2 with urea shows a smoother transition compared to that observed with GdmHCl, with a transition mid-point around 5.5 M. The observed transition most probably displays the dissociation of the dimers, but before the dissociation a minor structural change of the four-helix bundle develops between 0 and 4 M.

The change in mean residue ellipticity for equilibrium denaturation with urea of (KE2D15)_2 shows, surprisingly, no transition to a non-α-helical state. Instead the mean residue ellipticity decreases, which likely corresponds to a strengthening of the α-helical backbone. This suggests that the (KE2D15)_2 urea-denatured state is not a structureless random coil.

The steady state measurements of the equilibrium denaturation show that the four-helix bundle motif unfolds in a sequence of steps. The time-resolved measurements reveal structural details of these steps.

Equilibrium denaturation as followed by time-resolved fluorescence spectroscopy reveals structural transitions corresponding to folding / unfolding steps of (KE2D15)_2

The conformational dynamics of the native four-helix peptide, described in paper I, involve structural changes of the tertiary structure of the four-helix bundle. The nearly bimodal shape of the fluorescence lifetime distribution, at native conditions, shows that the four-helix bundle peptides are divided into two structurally different groups, differing in the water content of the hydrophobic core. The group with the longest fluorescence lifetimes extends up to 30 ns unequivocally shows that dansyl is part of a hydrophobic and solvent protected structure. These structures are the fully folded four-helix bundle conformations with the most developed tertiary structure and, for this reason, they constitute the starting-point of the unfolding pathway. The 10 ns group includes the molten globule structures, which have a partly hydrated hydrophobic core, are equivalent to four-helix structures after the first unfolding step, which means a partial hydration of the hydrophobic core. The native conformational equilibrium corresponds to transitions between the unfolding starting-point and the first unfolding step (paper I). However, new conformational equilibria will take effect when chemical denaturants are added.

The first effect of GdmHCl or urea is a displacement of the conformational equilibrium, from the dehydrated state towards the molten globule
state. This is consistent with that the unfolding pathway goes through the molten globule state before the dimer dissociates. Moreover, the increased molten globule population is borne out in the slight increase in the dansyl emission center of gravity before the dissociation starts. Up to this point on the unfolding pathway, the actions of urea and GdmHCl appear to be similar.

GdmHCl induced unfolding

1 M of GdmHCl destroys hydrophobic structures, which is borne out in the loss of the tail of the 20 ns fluorescence lifetime group (Figure 15a). At the same time the population of the 4.2 ns group (Table 3) has increased from 4% to 13%, which means that this group also contains transiently opened dimers, apart from the quenched component that occurs in the native structures (Paper I). Notably, the mean residue ellipticity (Figure 14a) indicates a strengthened α-helical backbone, which suggests that a conformational equilibrium is maintained between all hydration states of the hydrophobic tertiary interaction, i.e., the helix-loop-helix monomer is either dissociated or has solvent-mediated tertiary interactions, or dehydrated tertiary interactions.

The dehydrated helix bundles (α₄), which are the starting-point structures in the denaturation process, vanish between 1 and 2 M. The lost population of dehydrated bundles are now located in the 4.2 ns fluorescence lifetime group, which probes the helix-loop-helix monomers (α₂). Up to 2 M, the population of the dehydrated four-helix bundles has shifted, via the molten globule state (partly hydrated), to the maximally hydrated monomer state (KE2D15).

At 3 M, the major part of the original four-helix bundle population is located in the broad 5.5 ns group (Table 3). This lifetime is quite close to that of dansylglycine in water, and this group likely contains helix-loop-helix monomers (α₂). These structures may contain tertiary monomeric helix-helix interactions, manifested in the width of the sub-distribution and the mean-residue ellipticity at 3 M (Figure 14a). Dissociation of the molten globule four-helix bundle is seen in the large population shift (≈46%) of the dansyl fluorophores in the 11.4 ns group to the 4.7 ns group, from 2 to 3 M. The dissociation of the molten globule is concomitant with the large transition in the dansyl emission center of gravity from 576 nm to 600 nm in the same concentration interval. The pathway for dissociation goes, therefore, via the four-helix bundle molten globule state. Dissociation of the molten globule four-helix bundle constitutes the second unfolding step.

4 M of GdmHCl destroys the monomeric tertiary structure, as seen from the narrow width of the 4.9 ns fluorescence lifetime group. The 9.7 ns group may originate in structures with water-mediated dimeric interactions. The α-helical backbone is preserved in all KE2D15 structural species, as seen from the mean residue ellipticity. Up to 4 M, monomeric and dimeric tertiary structures, in all forms, are disrupted by GdmHCl. 4 M of GdmHCl consti-
tutes the point in the protein structure hierarchy where the helix-loop-helix monomer is precisely on the secondary structure level. The 9.7 ns fluorescence lifetime group might consist of structures being close to developing tertiary structure, but they cannot be characterized further. Disruption of the tertiary structure of the monomer constitutes the third unfolding step. The next step on the GdmHCl unfolding pathway is the disruption of the secondary α-helical structure.

The unfolding of the α-helical structure secondary structure, which is the fourth unfolding step, starts at 4 M and ends at 6 M. During this transition the shapes of the fluorescence lifetime distributions do not change much even though the constitution of the backbone changes dramatically. The distributions are bimodal with a population of dansyl fluorescence lifetimes (≈70%), centered around 5 ns, together with the remaining population centered around 8 ns (Table 3). The dansyl microenvironment is equally water exposed during the transition to a less negative mean residue ellipticity.

The end-point for the GdmHCl unfolding pathway, which is close to 7 M GdmHCl, does not show a positive mean residue ellipticity (Figure 14a), something which is expected for a random coil. The sharp and bimodal shape of the fluorescence lifetime distribution at this concentration indicates residual structure in the KE2D15 monomer. This can be a sign of switching between different backbone structures, e.g., α→polyproline-II like helix, of which the latter is known to show positive mean-residue ellipticity at 222 nm.
Figure 15. Maximum-entropy fluorescence lifetime distributions of (KE2D15)2 at various denaturing conditions. The effect of urea and GdmHCl on dansylglycine itself is in principle negligible. The fluorescence lifetime of dansylglycine changes from 2.4–3.3 ns and 2.4–3.4 ns for 0–6 M urea and GdmHCl, respectively. Consequently, changes in the fluorescence lifetime distributions reflect changes of four-helix bundle structure due to denaturation.
Table 3. Areas and peak-maxima of the fluorescence lifetime groups in the fluorescence lifetime distributions of (KE2D15)$_2$ during denaturation with GdmHCl.

<table>
<thead>
<tr>
<th>[GdmHCl]/M</th>
<th>$A_1$/%</th>
<th>$A_2$/%</th>
<th>$A_3$/%</th>
<th>$\tau_1$/ns</th>
<th>$\tau_2$/ns</th>
<th>$\tau_3$/ns</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>3.9</td>
<td>61</td>
<td>35</td>
<td>4.2</td>
<td>10.6</td>
<td>21.2</td>
</tr>
<tr>
<td>1</td>
<td>13</td>
<td>58</td>
<td>29</td>
<td>4.5</td>
<td>10.2</td>
<td>19.3</td>
</tr>
<tr>
<td>2</td>
<td>34</td>
<td>66</td>
<td>29</td>
<td>4.7</td>
<td>11.4</td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>80</td>
<td>20</td>
<td>4.5</td>
<td>11.4</td>
<td></td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>80</td>
<td>20</td>
<td>4.9</td>
<td>9.7</td>
<td></td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>78</td>
<td>22</td>
<td>4.9</td>
<td>8.3</td>
<td></td>
<td></td>
</tr>
<tr>
<td>6</td>
<td>66</td>
<td>34</td>
<td>4.7</td>
<td>7.6</td>
<td></td>
<td></td>
</tr>
<tr>
<td>7</td>
<td>68</td>
<td>32</td>
<td>4.9</td>
<td>7.7</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Table 4. Areas and peak-maxima of the fluorescence lifetime groups in the fluorescence lifetime distributions of (KE2D15)$_2$ during denaturation with urea.

<table>
<thead>
<tr>
<th>[urea]/M</th>
<th>$A_1$/%</th>
<th>$A_2$/%</th>
<th>$A_3$/%</th>
<th>$\tau_1$/ns</th>
<th>$\tau_2$/ns</th>
<th>$\tau_3$/ns</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>3.9</td>
<td>61</td>
<td>35</td>
<td>4.2</td>
<td>10.6</td>
<td>21.2</td>
</tr>
<tr>
<td>1</td>
<td>2.8</td>
<td>48</td>
<td>49</td>
<td>3.6</td>
<td>8.4</td>
<td>19.3</td>
</tr>
<tr>
<td>2</td>
<td>6.1</td>
<td>50</td>
<td>44</td>
<td>4</td>
<td>8.8</td>
<td>19.8</td>
</tr>
<tr>
<td>3</td>
<td>6.6</td>
<td>48</td>
<td>45</td>
<td>4.4</td>
<td>8.4</td>
<td>17.6</td>
</tr>
<tr>
<td>4</td>
<td>7.4</td>
<td>54</td>
<td>39</td>
<td>4.1</td>
<td>7.3</td>
<td>16.4</td>
</tr>
<tr>
<td>5</td>
<td>11</td>
<td>58</td>
<td>31</td>
<td>4.1</td>
<td>7.5</td>
<td>16</td>
</tr>
<tr>
<td>6</td>
<td>23</td>
<td>57</td>
<td>20</td>
<td>4.4</td>
<td>7.5</td>
<td>15</td>
</tr>
<tr>
<td>7</td>
<td>77</td>
<td>23</td>
<td>23</td>
<td>5.7</td>
<td>11.9</td>
<td></td>
</tr>
</tbody>
</table>

Comments for tables: Approximate area of the three sub-distributions ($A_1$, $A_2$, $A_3$) of fluorescence lifetimes as percentage of the total distribution, and maximal probable lifetime of each sub-distribution at different concentrations of urea and GdmHCl. The areas are calculated by integration from the minimum in the trough before the peak to the minimum after the peak. Uncertainties in areas are +/-3%. Peak lifetimes are +/-5%.

Urea induced unfolding

The first effect of urea, like GdmHCl, is to affect fluorescence lifetimes above 25 ns, i.e., the fluorescence from helix bundles with completely dehydrated hydrophobic cores (Figure 15b). The second effect is a much sharper separation of the dehydrated and partially hydrated populations, which remain rather constant up to 4 M. The peak fluorescence lifetimes of the two groups decrease monotonously, however, to lower values, implying increased water content of the two dimeric states (Table 4). As mentioned in Chapter 2, a shift of the hydrophobic interactions from the contact state to the solvent-separated state is expected in presence of urea. The native structural dynamics can be described in these terms, i.e., the hydrophobic forces of the dimeric interaction oscillate between these two states. This dynamic pattern persists up to 4 M. It is first at 5 M that a significant part of the contact state population is shifted to the solvent-separated state, which constitutes the hydrophobic interactions of the four-helix bundle molten
globule state. Concomitantly, the population of maximally water exposed fluorescence lifetimes, in the 4.1 ns group, starts to increase (Table 4). The increased populations of the molten globules and monomers correlate with the transition in the emission center of gravity, starting at 4M (Figure 14b). Accordingly, dissociation of the four-helix bundle starts at 4M of urea.

Surprisingly, the mean residue ellipticity changes to more negative values indicating strengthened α-helical structure of all conformational species in the four-helix ensemble (Figure 14b). The strengthening of the α-helical structure continues to 7 M, where the helix-loop-helix monomer has α-helical backbone structure, which is completely hydrated, as indicated by the narrow 5.1 ns lifetime group in Figure 15b.

The denaturation pathways of GdmHCl and urea are summarized in Figure 16.
Figure 16. Stepwise unfolding of the four-helix bundle (KE2D15)_2 with urea and GdmHCl. This illustration is based on the time-resolved fluorescence data and the CD data, and displays the dominating peptide species of the conformational equilibria at different denaturing conditions. The unfolding pathways of urea and GdmHCl are the same concerning how the tertiary structure is denatured (green structures). However, GdmHCl unfolds the secondary structure as well (red structures), in contrast to urea.
The conformational backbone space of the α-helical peptide (Paper III)

What happens with the α-helical backbone of the four-helix bundle when temperature increases? The mean residue ellipticity at 222 nm increases but what kind of structural transitions, on the secondary level, occur under these conditions? Can the different conformations of the backbone be resolved within the time-resolution of time-resolved fluorescence spectroscopy (tens of nanoseconds)? There are many questions related to the conformational space of the α-helix, which is why the short α-helical segment in Figure 8 was chosen as a experimental model. The crystal structure\textsuperscript{103} shows, as mentioned, α-helical organization of the backbone. The α-helical conformation is, however, not the only the conformation in the solvated form. Paper III describes the conformational space of the helical segment using CD spectroscopy in combination with end-to-end measurements by FRET.

The peptide end-to-end distances reflects a correlated reorganization of the WLL12 backbone at different solvent conditions

The WLL12 peptide was labeled with tryptophan at the N-terminus and nitrotyrosine at the C-terminus. This donor–acceptor pair is referred to as DA in the text, and the corresponding reference peptide as D. From the tryptophan fluorescence lifetime the efficiency of energy transfer can be calculated (Equation 14), and from the efficiency the end-to-end distance can be calculated (Equation 12). Figure 17 shows the fluorescence decays of tryptophan from the TCSPC measurements of WLL12D at 20% TFE (trifluoroethanol) and WLL12DA at 0%, 10% and 20%TFE. TFE sometimes induces an α-helical backbone structure in polypeptides, something which is observed here. With increasing TFE concentration the end-to-end distance of WLL12 decreases. CD spectroscopy along with the end-to-end distance information was used to understand what the structural effects are.
Figure 17. a) Tryptophan fluorescence decays, from top, WLL12D at 20% TFE (green), WLL12DA at 0% (blue), WLL12DA at 10% (violet) and WLL12DA at 20% TFE (red). The fitted multiple discrete decay models are plotted upon the data points. The choice of a multiple-discrete decay law is supported by the maximum-entropy distributions shown in b.

Figure 18. CD spectra of WLL12 at (from top) 0% (blue), 5%, 10% (violet), 15%, 20% (red), 25% and 30% TFE.

The WLL12D and -DA decays were first analyzed with the non-biased maximum-entropy method, which systematically resulted in three sharp fluorescence lifetime groups for the D peptide and four narrow lifetime groups for the DA peptide (Figure 17b). Notably, all the DA decays are divided, as
already mentioned, into well-separated fluorescence lifetime groups, which indicates that the end-to-end distances are quite well defined. This is not expected for a random coil. The narrow distributions argue for that decay models with three time constants can be used for the donor decay and four time constants for the donor acceptor decay. The results from the multi-discrete analysis are shown in Table 5.

Table 5. Analysis of time-resolved fluorescence measurements of WLL12-D and WLL12-DA with multi discrete decay models at different concentrations of TFE.

<table>
<thead>
<tr>
<th>solvent</th>
<th>decay</th>
<th>a1</th>
<th>a2</th>
<th>a3</th>
<th>a4</th>
<th>τ1</th>
<th>τ2</th>
<th>τ3</th>
<th>τ4</th>
<th>I_{st}</th>
<th>&lt;E&gt;</th>
<th>&lt;r&gt;</th>
</tr>
</thead>
<tbody>
<tr>
<td>D</td>
<td>0.26</td>
<td>0.32</td>
<td>0.42</td>
<td></td>
<td>0.66</td>
<td>1.9</td>
<td>5.3</td>
<td></td>
<td>3.0</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>0.63</td>
<td>24</td>
<td></td>
</tr>
<tr>
<td>TFE</td>
<td>DA</td>
<td>0.15</td>
<td>0.36</td>
<td>0.35</td>
<td>0.14</td>
<td>0.08</td>
<td>0.49</td>
<td>1.4</td>
<td>3.1</td>
<td>1.1</td>
<td>±0.01</td>
<td>±0.1</td>
</tr>
<tr>
<td></td>
<td></td>
<td>±0.01</td>
<td>±0.01</td>
<td>±0.01</td>
<td>±0.02</td>
<td>±0.02</td>
<td>±0.03</td>
<td>±0.08</td>
<td>±0.08</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>D</td>
<td>0.21</td>
<td>0.31</td>
<td>0.48</td>
<td></td>
<td>0.65</td>
<td>1.9</td>
<td>4.6</td>
<td></td>
<td>3.0</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>10%</td>
<td></td>
<td>±0.01</td>
<td>±0.01</td>
<td>±0.01</td>
<td>±0.01</td>
<td>±0.02</td>
<td>±0.01</td>
<td></td>
<td></td>
<td>0.75</td>
<td>22</td>
<td></td>
</tr>
<tr>
<td>TFE</td>
<td>DA</td>
<td>0.21</td>
<td>0.41</td>
<td>0.29</td>
<td>0.09</td>
<td>0.07</td>
<td>0.36</td>
<td>1.1</td>
<td>2.7</td>
<td>0.7</td>
<td>±0.01</td>
<td>±0.1</td>
</tr>
<tr>
<td></td>
<td></td>
<td>±0.01</td>
<td>±0.01</td>
<td>±0.01</td>
<td>±0.01</td>
<td>±0.01</td>
<td>±0.02</td>
<td>±0.01</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>D</td>
<td>0.17</td>
<td>0.35</td>
<td>0.48</td>
<td></td>
<td>0.65</td>
<td>2.1</td>
<td>3.9</td>
<td></td>
<td>2.7</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>20%</td>
<td></td>
<td>±0.01</td>
<td>±0.01</td>
<td>±0.01</td>
<td>±0.01</td>
<td>±0.01</td>
<td>±0.01</td>
<td>±0.01</td>
<td></td>
<td>0.84</td>
<td>20</td>
<td></td>
</tr>
<tr>
<td>TFE</td>
<td>DA</td>
<td>0.22</td>
<td>0.46</td>
<td>0.26</td>
<td>0.06</td>
<td>0.06</td>
<td>0.28</td>
<td>0.75</td>
<td>1.8</td>
<td>0.4</td>
<td>±0.01</td>
<td>±0.1</td>
</tr>
<tr>
<td></td>
<td></td>
<td>±0.01</td>
<td>±0.02</td>
<td>±0.01</td>
<td>±0.01</td>
<td>±0.01</td>
<td>±0.02</td>
<td>±0.03</td>
<td>±0.05</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Tryptophan usually exhibits three fluorescence lifetimes in proteins. The lifetimes originate from different rotamer conformations of the indole sidechain in tryptophan. The existence of a fourth lifetime component in the DA decay indicates that there are at least two different Förster energy transfer processes occurring at the same time in the ensemble of WLL12 peptides. However, in order to determine the end-to-end distance (Equation 12) from energy transfer efficiencies the correct donor fluorescence lifetime has to be matched with the correct lifetime in the DA decay (Equation 14), which is difficult when the donor decay is complex. For this reason, as a start, the overall efficiency of energy transfer was calculated from the contribution of all the fluorescence lifetimes to the steady state intensity (Equation 15), which results in the time or ensemble averaged end-to-end distance of WLL12 at the different solvent conditions (Table 5). As observed, the end-to-end distance of WLL12 is shortened by TFE, indicating that TFE is inducing a significant structural rearrangement of the backbone of the WLL12.

CD spectra of WLL12 at different TFE/water mixtures are displayed in Figure 18, where the 0%, 10%, and 20% TFE concentration points have the
same color-coding as in Figure 17. The CD spectra confirm the conclusion from the time-resolved data that TFE causes a significant change of the WLL12 backbone, as seen by the abrupt change of the spectral shape when TFE concentration increases. This indicates a two-state reorganization of the backbone, from the unknown starting and native WLL12 state at 0% TFE, to an $\alpha$-helical state at 20% TFE, as judged by the shape of the CD-spectra at 20% and higher TFE concentrations.

Since all types of secondary structures of polypeptides are defined by the dihedral angles ($\phi$, $\psi$) of the backbone, the two-state reorganization implies that all amino acids of WLL12 changes from one set of initially defined values ($\phi_{\text{ini}}$, $\psi_{\text{ini}}$), to another set of defined values at the end of the transition ($\phi_{\text{end}}$, $\psi_{\text{end}}$), at least as seen from the time-averaged CD picture. The spectral shape at 20% TFE indicates, as mentioned before, $\alpha$-helical backbone, which implies $\alpha$-helical dihedral angles of WLL12 ($\phi_{\text{end}} = -47^\circ$, $\psi_{\text{end}} = -57^\circ$). Does the measured end-to-end distance of WLL12 at 20% TFE match the end-to-end distance of WLL12 in an $\alpha$-helical conformation? To answer this question WLL12 was modeled in an $\alpha$-helical conformation using the Molmol software. The $\alpha$-helical WLL12 model had an end-to-end distance equal to 18.7 Å, which is near the FRET-averaged end-to-end distance 20 Å (Table 5). This shows that the measured ensemble averaged end-to-end distance reflects the backbone structure of WLL12 in the same way as the CD spectra do.

**The native time-averaged WLL12 state is a $3_{10}$-helix**

Using the same method as above, the by FRET measured end-to-end distance of 24 Å at 0% TFE (Table 5) was tested by modeling the WLL12 backbone according to a $3_{10}$-helix. This is a reasonable approach, because $\alpha$-helical structures have been described to be in equilibrium with $3_{10}$-helices as described in Chapter 2. The end-to-end distance of the model $3_{10}$-helix was found to be 23.7 Å, which is very close to the observed 24 Å from FRET measurements. Accordingly, it can be deduced that the native solvated WLL12 peptide has a backbone conformation that corresponds to a $3_{10}$-helix. The CD spectrum of the $3_{10}$-helix has been characterized earlier, and is similar to that of WLL12-spectrum at 0% TFE. Consequently, it is clear that the solvated WLL12 does not have a completely disorganized backbone structure, even though the mean residue ellipticity at 222 nm is weak. Consequently, we now know that the backbone of WLL12 in water, at least on average, is slightly elongated, similar to a $3_{10}$-helix. However, this is, as already mentioned, the time averaged picture.
The FRET decay of WLL12 and SVD analysis of the CD spectra reveals four backbone conformations in equilibrium.

The fact that all WLL12D decays can be modeled with three time constants and that the WLL12DA decay has to be modeled with four implies that at least two distances are present in the WLL12 ensemble. This means that the conformational transitions of WLL12 equal contraction and elongation of the peptide’s end-to-end distance, due to reorganization of the peptide’s backbone. As mentioned before, if the correct pairs of $\tau^D$ and $\tau^{DA}$ can be found then lengths of the conformations can be determined. In paper III a strategy, based on the energy transfer rate, was tried to find the true distances in the ensemble of WLL12. All possible combinations of energy transfer rates between the tryptophan fluorescence lifetimes in the D and the DA decays were tried, under the constraint that $\tau^D$ must be larger than $\tau^{DA}$ to be a true donor-acceptor pair. Calculation of the energy transfer rate of all allowed pairs, at every concentration of TFE, resulted in four classes of Förster energy-transfer rates, which equals four end-to-end distances that are present at the same time in the WLL12 ensemble. This is consistent with that one WLL12 peptide explores four different backbone conformations. The four time-resolved WLL12 end-to-end distances are presented in Table 6.

### Table 6. Calculated time-resolved end-to-end distances of WLL12 from paper III

<table>
<thead>
<tr>
<th>TFE%</th>
<th>$r_1$/Å</th>
<th>$r_2$/Å</th>
<th>$r_3$/Å</th>
<th>$r_4$/Å</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>27</td>
<td>22</td>
<td>18</td>
<td>13</td>
</tr>
<tr>
<td>10</td>
<td>27</td>
<td>21</td>
<td>17</td>
<td>13</td>
</tr>
<tr>
<td>20</td>
<td>25</td>
<td>20</td>
<td>17</td>
<td>13</td>
</tr>
</tbody>
</table>

The distances in Table 6 suggest that four conformations, with different end-to-end distances, are in equilibrium. The 13 Å distance obtained from time-resolved data matches perfectly the modeled end-to-end distance of WLL12 in a $\pi$-helix conformation (12.6 Å). 18 Å matches the distance of an $\alpha$-helix (18.7 Å). 22 and 27 Å are distributed around the end-to-end distance of WLL12 in a $3_{10}$-helix conformation. The different helices are illustrated in Figure 20.

Increased concentration of TFE does not change the end-to-end distances much. However, the populations of $\tau_1$ and $\tau_2$ increase in the DA decay as the populations of $\tau_3$ and $\tau_4$ decrease. This is consistent with that the population of the short helices increases. The effect of TFE is, consequently, a desolvation of the peptide backbone, something which is observed both from the time-averaged and the time-resolved aspect.
**SVD analysis of CD spectra shows the presence of different conformations of the WLL12 backbone**

Singular value decomposition of all CD-spectra in Figure 18, 0–30% TFE, were performed according to Equation 17, in order to determine the independent number of spectral components that exist in the dataset, which is already known to contain different conformations (from FRET). Three spectral components (the three first components in U) are necessary to reproduce the data properly. In particular, the spectra between 0% and 10% TFE, need all three components in order to be satisfactorily reproduced. The validity of the SVD model is demonstrated in Figure 19. No spectra could be properly reproduced with the first single value component SVD1. The spectra between 15% and 30% TFE are reproduced with SVD1 and SVD2, and all spectra are perfectly reproduced with SVD1, SVD2 and SVD3. The singular value components are plotted on top in Figure 19. SVD1 is reminiscent of the CD spectrum of an α-helix. This component dominates the spectra at 15–30% TFE, along with SVD2. Below 15% TFE, the SVD2 and SVD3 components are more important.

The SVD analysis confirms the conformational equilibrium of WLL12 peptides, with different backbone structures, by finding three independent singular value components.
Figure 19. a) Reproduced CD spectra (grey line) with SVD1 (boxes, on top) do not match the measured spectra (from Figure 18, black line). b) Reproduced CD-spectra with SVD1 and SVD2 (circles, on top) reproduces WLL12 spectra 15–30% TFE. c) Reproduced CD spectra with SVD1, SVD2 and SVD3 reproduce at all spectra. d) The significance of each component at the different concentrations of TFE.
Elongation and contraction of the WLL12 backbone involves rearrangements of the internal hydrogen bonds

The FRET analyses suggest the following pathway of the conformational changes upon addition of TFE: \( \text{3}_{10} \rightarrow \alpha \rightarrow \pi \). This pathway of structural change seems be connected to the internal hydrogen-bond pattern of WLL12 and its coupling to the surrounding water. The \( \text{3}_{10} \)-helix is the first helix with internal hydrogen bonds, with every third residue, in the family of helical backbone structures (Figure 20). A contraction to the \( \alpha \)-helical state, which has internal hydrogen bonds between every fourth residue, can be regarded as a desolvation of the peptide. The end-to-end distance is reduced by 5 Å in the \( \text{3}_{10} \rightarrow \alpha \) transition, something which reduces the solvent accessible area as well. The subsequent \( \alpha \rightarrow \pi \) transition corresponds to a contraction of 6 Å. During the contractions, the longer conformation has to be desolvated at the same time as the internal hydrogen bonds are reorganized in the new helix geometry. Because of the desolvation and the breaking of internal hydrogen bonds the contraction might be energetically activated, which makes the contraction cooperative. This refers to the framework and the hydrophobic collapse theories in the Chapter 2. Based on the conformational transitions of WLL12, the early steps of protein folding consist of a stepwise desolvation, which can be considered as a stepwise hydrophobic effect, where two processes are correlated: rearrangements of the internal hydrogen bonds and desolvation of the polypeptide chain.
Figure 20. The conformational backbone space of the WLL12 peptide in water.
A coupled folding and binding funnel leads to recognition between the helix-loop-helix and human carbonic anhydrase (Paper IV)

This paper presents the results of the experiments with human carbonic anhydrase II (HCAII) (Figure 8) and the helix-loop-helix peptide (KE2D15-8), modified with dansyl and benzenesulfonamide (Figure 9). In paper I and II it was demonstrated how the dansyl fluorescence works as a probe of the tertiary structure. The dansyl fluorescence lifetime is used in this paper as well to characterize different structures that develop as binding occurs between KE2D15-8 and HCAII. Aside from the dansyl fluorescence lifetime, the rotational diffusion of the dansyl fluorophore is another central observable that provides additional information about the nature of the fast fluctuating structures, which appears when the polypeptides bind to each other. This information can be extracted by time-resolved measurements of the dansyl fluorescence. Furthermore, HCAII can also give spectroscopic information, since it contains six tryptophans, which can be used in energy transfer to dansyl on the helix-loop-helix peptide.

Dansyl probes new and folded tertiary structure as a result of binding

The dissociation constant ($K_D$) in Equation 18 has been determined earlier to be 4 nM by measuring the change in steady state fluorescence intensity when HCAII was titrated into KE2D15-8. The dissociation constant of benzenesulfonamide decreased three orders of magnitude when it was coupled to the helix-loop-helix peptide, which gives extra affinity to HCAII by binding to it. It is the peptide/protein interaction that is the subject of investigation for this work – not the benzenesulfonamide/HCAII interaction.

$$K_D = \frac{[HCAII]_{free}[KE2D15-8]_{free}}{[KE2D15-8:HCAII]}$$

*Equation 18*

Fluorescence spectra from a similar experiment are shown in Figure 21. From the fluorescence intensity at 550 nm the concentration of formed KE2D15-8:HCAII complexes was estimated and fitted to a 1:1 binding-model (Equation 18), which resulted in a dissociation constant of 3 nM. The
dansyl emission center of gravity was blue-shifted during the titration, which indicates that dansyl is located in a hydrophobic environment in the complex.

**Figure 21.** a) Fluorescence spectra of KE2D15-8 (1 μM) during titration with HCAII. The thick spectrum corresponds to pure spectrum of KE2D15-8 without HCAII. b) The concentration of formed complex was estimated from the fluorescence intensity at 550 nm (boxes), which was fitted to 1:1 binding model (solid line). Dansyl emission center of gravity is blue-shifted as the concentration of complex increases, indicating that a new molecular environment has been formed between the ligands.

**Time-resolved dansyl fluorescence indicates three interaction states**

The fluorescence lifetime distributions of uncomplexed and complexed KE2D15-8 with HCAII are shown in **Figure 22**. The broad uncomplexed KE2D15-8 dansyl fluorescence lifetime distribution, stretching up to 17 ns, indicates that the tertiary structure of 1 μM KE2D15-8 is very disorganized. This makes for an interesting comparison with the dimeric, much more well-structured four-helix bundle structure (**Figure 10**). However, the shape of the dansyl fluorescence lifetime distribution changes dramatically when KE2D15-8 binds to HCAII. In the bound state, three defined groups appear. Judging from their peak fluorescence lifetimes, these groups correspond to hydrated, partly hydrated and dehydrated conformations of the bound peptide. The dansyl fluorescence lifetime distributions reveal that the binding and recognition of the two polypeptides corresponds to a disordered-to-ordered transition of the interfacial structure between the two polypeptide-ligands. This is seen in that the 5.9 ns peak-maximum of the uncomplexed distribution that shifts to 17.4 ns when the helix-loop-helix is complexed. The shift corresponds to a transition of a completely hydrated and disorganized structure, to a dehydrated tertiary structure between the ligands. This
transition does not, however, occur in an all-or-none step. The bimolecular structure folds via at least one or perhaps two intermediate bimolecular structures, given by the hydrated and certainly the partly hydrated population (Figure 22 upper part). The time-resolved measurements reveal that these states are in equilibrium with each other, and interchange conformations on a timescale that is slower than the lifetime of the dansyl fluorescence. The pathway through the coupled folding and binding funnel is thus associated with desolvation of their initially disorganized parts, which occurs via a molten globule like state that is manifested by the 7.6 ns group.

![Figure 22](image)

**Figure 22.** Lower left: Dansyl fluorescence (excitation at 335 nm) lifetime distribution from maximum-entropy analysis shows a very disorganized tertiary structure of the uncomplexed helix-loop-helix monomer. Upper left: The dansyl fluorescence lifetime distribution of the KE2D15-8:HCAII complex indicates three interactions, i.e., hydrated, partly hydrated and dehydrated states between the helix-loop-helix and HCAII. Right: Sensitized dansyl fluorescence, by energy transfer from tryptophans in HCAII (excitation at 290 nm), reaches only the partly hydrated and the dehydrated interaction conformations between the helix-loop-helix and HCAII.

**Table 7.** Summary of the dansyl fluorescence lifetime distributions in Figure 22 (left part) for KE2D15-8 and KE2D15-8 complexed with HCAII. Excitation at 335 nm.

<table>
<thead>
<tr>
<th>Fluorescence lifetimes of dansyl</th>
<th>Uncomplexed KE2D15-8</th>
<th>Complexed KE2D158:CA</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\tau_{\text{max} 1}$/ ns</td>
<td>0.2 (0.6%)</td>
<td>0.2 (0.5%)</td>
</tr>
<tr>
<td>$\tau_{\text{max} 2}$/ ns</td>
<td>1.1 (5.2%)</td>
<td>0.9 (3.1%)</td>
</tr>
<tr>
<td>$\tau_{\text{max} 3}$/ ns</td>
<td>3.3 (9.5%)</td>
<td>3.1 (14.4%)</td>
</tr>
<tr>
<td>$\tau_{\text{max} 4}$/ ns</td>
<td>5.9 (64%)</td>
<td>7.6 (46%)</td>
</tr>
<tr>
<td>$\tau_{\text{max} 5}$/ ns</td>
<td>n/a</td>
<td>17.4 (37%)</td>
</tr>
</tbody>
</table>

Peak maxima and fractions of each sub-distribution, compared with the total distributions of fluorescence lifetimes, is given in the parenthesis. Uncertainties in the data is less than ±5%.
Three dansyl rotational correlation times in the KE2D15-8:HCAII complex also indicate three conformations of interactions

Fluorescence anisotropy was measured on the uncomplexed and complexed KE2D15-8 with HCAII. The fluorescence depolarization decays are displayed in Figure 23. A bi-exponential decay model (Equation 10) for the uncomplexed KE2D15-8 describes the measured decay well. Two rotational components of 0.21 ns and 2.1 ns were extracted from the decay. This confirms the previous interpretation of the fluorescence lifetime distribution (Figure 22) that the tertiary structure of the free helix-loop-helix monomer is disorganized. For instance, a hindered dansyl rotational motion of 2.1 ns can arise because dansyl is a part of the tertiary structure of the helix-loop-helix monomer. The 2.1 ns component should be compared to the rotational motion of 2 ns of the four-helix bundle peptide at 60 °C, at which temperature it is monomeric (Figure 13).

Figure 23. Fluorescence anisotropy (depolarization) decays of uncomplexed (a) and complexed (b) KE2D15-8:HCAII (1 μM : 2 μM). The fitted decay models are plotted with a solid line.

The anisotropy decay of the complex can be described by three rotational correlation times: 0.18 ns (22%), 1.2 ns (44%) and 19 ns (37%). This result was achieved when all six decay-model parameters were allowed to vary freely. Notably, the different populations of the rotational correlation times correlates rather well with the different fluorescence lifetime populations of the KE2D15-8:HCAII complex in Table 7. This correlation suggests that the less hindered rotational correlation time is the most water exposed, which results in a fluorescence lifetime of 3.1 ns. Analogously, the rotational motion of the dansyl fluorophore is maximally hindered if it is a part of a bi-molecular dehydrated tertiary structure, which results in fluorescence life-
time of the fluorophore of 17.4 ns. The same reasoning holds for the inter-
mediately hindered rotation and the intermediate fluorescence lifetime

group. A perfect correlation between the anisotropy and fluorescence life-
time data of the KE2D15-8:HCAII was found when the population of the
0.18 ns rotational component was locked to equal the population of the 3.1
ns lifetime group (14%). The final results of the fluorescence anisotropy
analysis of are shown in Table 8.

The 23 ns rotational-correlation time is typical for globular proteins of the
same size as HCAII, which is why this population of dansyl fluorophores
follows the rotational motion of HCAII when the helix-loop-helix is strongly
bound to the HCAII surface. For comparison, the rotational correlation time
of dansylglycine in water is around 60 ps.

The dansyl fluorescence anisotropy data supports, strongly, the coupled
folding and binding process, including the partly folded intermediates. This
is illustrated with Figure 24.

Figure 24. Illustration of the two most important conformations of interaction be-
tween the helix-loop-helix peptide and human carbonic anhydrase II. Recognition of
the helix-loop-helix to HCAII is consistent with a coupled folding and binding proc-
ess, which involves the formation of solvent assisted hydrophobic contacts in a mol-
ten globule like bimolecular structure (1). The folding and binding of the final com-
plex involves a dehydration step, which creates contact state hydrophobic interac-
tions (2). Note that the crystal structure of HCAII displays a disordered region where
the helix-loop-helix is believed to bind.

Table 8. Dansyl fluorescence anisotropy data of uncomplexed and complexed
KE2D15-8 with HCAII

<table>
<thead>
<tr>
<th>Anisotropy of dansyl</th>
<th>$\phi_1$/ns</th>
<th>$\phi_2$/ns</th>
<th>$\phi_3$/ns</th>
<th>$r_0$</th>
<th>$\chi^2$</th>
</tr>
</thead>
<tbody>
<tr>
<td>KE2D15-8</td>
<td>0.21±0.04 (41±7%)</td>
<td>2.1±0.1 (59±2%)</td>
<td>n/a</td>
<td>0.30</td>
<td>1.16</td>
</tr>
<tr>
<td>KE2D15-8:HCAII</td>
<td>0.18±0.1 (14 ± 6%)</td>
<td>1.2±0.2 (47±4%)</td>
<td>23±2 (39±2%)</td>
<td>0.27</td>
<td>1.08</td>
</tr>
</tbody>
</table>
The most important interaction conformations of the helix-loop-helix can be seen from the HCAII surface through energy transfer

Because of the six tryptophans close to the surface of HCAII there is a possibility for having energy transfer from tryptophan to dansyl. The tryptophan fluorescence of HCAII overlaps with the dansyl absorption of KE2D15-8 (Figure 25.a). Only the tryptophans close to the dansyl will be able to undergo energy transfer. The Förster distance for the Trp:dansyl pair is 21.5 Å. Figure 25b shows that the tryptophan emission in HCAII decreases with increasing concentration of KE2D15-8. The energy transfer from tryptophan or several tryptophans in HCAII results in sensitized emission of dansyl. The increasing dansyl fluorescence and tryptophan quenching, due to energy-transfer between the two chromophores, levels out at 1:1 stoichiometry of HCAII and KE2D15-8. The tryptophan intensity is at this point quenched to approximately 70% of the intensity in absence of donor. This results in an apparent energy transfer efficiency of 34% (Equation 13), which corresponds to an apparent distance of 24 Å (Equation 12). However, the true distance can be shorter if not all tryptophan donators participate in the energy-transfer to dansyl. It is a task for time-resolved experiments to resolve the rate constant for energy-transfer, from which the real distance between the chromophores can be found. Sensitized time-resolved dansyl fluorescence can be used to characterize the nature of the interacting helix-loop-helix conformations.

Figure 25. a) Dansyl absorption of KE2D15-8 (squares) and tryptophan emission of HCAII (circles). b) Energy-transfer between one or several tryptophans in HCAII to dansyl on KE2D15-8. Excitation at 290 nm.
Where does KE2D15-8 bind to HCAII?

The apparent energy-transfer efficiency was also calculated from tryptophan fluorescence lifetime measurements of uncomplexed and complexed HCAII with KE2D15-8 (1:1). The fluorescence decay is complex and composed of five different tryptophan fluorescence lifetime groups for both the complexed and uncomplexed HCAII. From the fluorescence lifetime analysis, the apparent energy transfer efficiency was calculated to 37%, using Equation 15 and Equation 12. This is in good agreement with the 34% obtained from the steady-state measurement. The energy transfer data can be used to understand where the helix-loop-helix peptide binds.

The crystal structure of HCAII shows that five of the six tryptophans are located in the unstructured N-terminus region (Figure 26). An energy-transfer efficiency of 34% corresponds to full quenching of two tryptophans out of the six available (assuming equal quantum yields of the fluorescence), which can only be accomplished if KE2D15-8 binds to the five-tryptophan side of HCAII. Consequently, there are at least two donors involved in the energy transfer to dansyl.

Figure 26. Locations of the six tryptophans in HCAII. The lowest tryptophan in the middle is located on the opposite side of the globule relative to the five tryptophans that are closest to the reader in the disorganized N-terminus.

Global analysis of the tryptophan and the dansyl fluorescence decays reveals the energy-transfer rate-constant and the nature of the conformations of interaction that are close to HCAII surface

The decay rate of the tryptophan fluorescence should be equal to the rate with which the dansyl fluorescence intensity increases, when energy is transferred from tryptophan to dansyl. Therefore the fluorescence decay models of tryptophan and dansyl were coupled, assuming that one time constant is common, but with opposite signs of the pre-exponential factor, for both decays. The decaying component of the tryptophan fluorescence and the in-
growing component of the dansyl fluorescence was found to be 140 ps, through global analysis. The following model for the time-dependent dansyl fluorescence was established: one in-growing component 140 ps (−12%), and two decaying components of 8.2 ns (54%) and 18.5 ns (32%) (χ²=1.09). The inverse of the 140 ps in-growing time constant equals the energy transfer rate, which corresponds to an apparent distance of 12 Å between at least two tryptophans in HCAII and the dansyl on KE2D15-8.

If the two interaction conformations in Figure 24 have different distances between donor and acceptor, because of a water layer, it implies that two different in-growing components would exist in the dansyl decay. This is not case, which is explained partly by that the second energy transfer is too fast to be resolved on the TCSPC setup, partly by direct excitation of dansyl at 290 nm. However, the decaying components show two different dansyl microenvironments, i.e., one partly hydrated and one dehydrated, but the decay model does not contain the most hydrated dansyl conformation, which results in a fluorescence lifetime around 3 ns. The hydrated conformation is too far away from the HCAII surface to be excited by energy transfer. The other two conformations are already known from direct dansyl excitation and the rotational motions of dansyl. Accordingly, the hydrated and the strongly interacting conformation, together with the intermediate molten globule like structure, are closest to the HCAII surface. A maximum entropy fit of the decaying part of the sensitized dansyl fluorescence is shown in Figure 22, which indicates defined interaction structures.
6. Conclusions

The findings in thesis demonstrate that structural fluctuations of the protein occur on all levels of the protein architecture.

Pathways of structural changes on the tertiary level

Different four-helix bundle conformations, which are probed with nanosecond time resolution, show differing degrees of hydration in the hydrophobic core. The time-resolved fluorescence spectroscopy experiments reveal that some conformations have a completely dehydrated hydrophobic core whereas the other group is partly hydrated. With respect to the different water content in this tertiary structural element the partly hydrated form must be considered to be less folded and more disorganized than the dehydrated version. Consequently, the observed structural dynamics consist of structural transitions between different states on the four-helix-bundle’s folding pathway. It has been shown that the partly hydrated conformation corresponds to a molten globule, which consists of solvent separated tertiary hydrophobic interactions. The switching between the dehydrated four-helix bundles and the molten globules is a cooperative process, which is manifested by the clear separation of the two conformational groups, as they are resolved on the nanosecond timescale. This is especially clear at lower temperatures or at moderate concentrations of urea, which freeze out the populations of the dehydrated and the molten globule four-helix bundle even further. The dehydrated four-helix bundle and molten globule structural transitions are compatible with folding/unfolding fluctuations in the four-helix bundle’s tertiary structure.

Tertiary structural transitions may involve structural changes on the secondary level as well, depending on the solvent conditions

If the tertiary fluctuations involve structural rearrangements on the secondary level as well depends on the solvent conditions, *i.e.*, the presence of chemical denaturant, and the temperature. It is clear from the temperature studies that the tertiary structure persists when temperature increases, despite the overall decreasing $\alpha$-helical content. This is consistent with that the in-
creased thermal energy affects the motions of the backbone structure. In contrast to increased temperature, urea or GdmHCl initially influence the side chains instead of the backbone structure. Denaturation of the four-helix bundle with these chemical agents strengthens (0–4 M GdmHCl), surprisingly, the α-helical backbone structure. This is consistent with that the structural transitions only concern the tertiary level, which involves hydration and dehydration of its hydrophobic parts due to increased motions of the hydrophobic sidechains. The structural differences between thermal and chemical denaturation show that the pathways of the structural transitions are different, which has implications for protein folding experiments that aim to elucidate the folding mechanism of proteins. How the denatured ensemble is induced will both affect the structure of it and the refolding pathway. The three ways of inducing the four-helix bundle unfolded state passes, however, always over the molten globule state before the tertiary structure of the four-helix bundle is disrupted.

GdmHCl, in contrast to urea, disrupts the α-helical structure after the denaturation of the tertiary structure.

Pathways of structural changes on the secondary level

The 12 amino acid long helical peptide (WLL12) provides a detailed picture of the structural transitions occurring in a helical segment. This structure was found to expand and contract, which contrasts looping and helix–coil transitions reported in other studies. The elongation and contraction motion is compatible with rearrangement of the backbone structure. Under native conditions, i.e., solvated in water at neutral pH at room temperature, the time-averaged end-to-end distance and CD-spectrum suggest a 3_10-helix, which is in contrast to the α-helical structure that is reported from the crystal structure. The time-resolved picture shows, however, that the helical peptide switches between three different helical states, namely: the 3_10-helix state, the α-helix state and the π-helix state. Changing the solvent conditions, by adding trifluoroethanol (20%) as a cosolvent, has a desolvating effect on the peptide, something which is monitored as an increased population of α-helices and π-helices.

The experiments with WLL12 do not show how these different, but all helical states, are related to the formation of the tertiary structure, e.g., a four-helix bundle. However, it is a speculation of mine that the formation of tertiary structure implies a desolvation of the helical backbone (via the molten globule state), which is to some degree mimicked by TFE. If this is true, structural transitions of a four-helix bundle, or any other globular protein, involves concomitant rearrangements on the tertiary and the secondary levels.
The pathways of conformational transitions of WLL12, starting from the most elongated and solvated state $3_{10}$-helix state, a conformational transition from here to the $\alpha$-helix state is equal to a contraction of the end-to-end distance. This pattern of structural transitions can be regarded as a sequential hydrophobic effect. However, it is not driven by non-local hydrophobic interactions as suggested in the hydrophobic collapse model (Chapter 2), rather it is the contraction of the end-to-end distance and the formation of internal hydrogen bonds that add up to the total hydrophobic effect. The framework model (Chapter 2) is more compatible with this sequence of hydrophobic steps, if more helical backbone structures are incorporated in the framework model, like the $3_{10}$-helix. The experiments with WLL12 show that other helical backbone structures than the $\alpha$-helix state have to be considered when following protein unfolding by circular dichroism. The reorganization of the helical backbone occurs on the microsecond timescale, which is poorly resolved in many kinetic folding experiments (Chapter 2).

Which folding model is supported by the structural dynamics observed in the helical peptides?

The large-scale equilibrium fluctuations of the four-helix bundle, between the dehydrated and the molten globule state, is evidence of that the four-helix bundle folds in a stepwise manner. Folding via the molten globule intermediate is originally proposed in the framework model. This theory was formulated on the basis of self-organization of myoglobin, which contains nine helical regions. The framework model predicts that the secondary $\alpha$-helical backbone structure precedes the globular tertiary structure, which forms via the molten globule intermediate. The time-resolved fluorescence study of (KE2D15)$_2$ is to my knowledge unique in the sense that the fully folded four-helix bundle and the molten globule version are resolved as separate structures, somewhere on the microsecond timescale. This type of conformational search is entirely compatible with the framework model concerning the formation of the tertiary structure of the four-helix bundle motif.

Folding intermediates, like the molten globule, are not well characterized because of the difficulty to trap this state where it can be studied with ordinary structural methods. The time-resolved fluorescence shows that the molten globule is on the unfolding/folding pathway, no matter of chosen method for inducing the unfolded state. Kinetic intermediates have been observed in refolding experiments of globular and helix proteins, which confirm a sequential folding mechanism. A transient structure has been observed in the refolding of $\alpha$-lactalbumin and for a three-helix bundle. Equilibrium unfolding as well as kinetic refolding studies of cytochrome c (four-helix bundle) shows intermediate structures. In fact two sequential
intermediates are observed in refolding of the bundle protein cytochrome c\textsuperscript{33,35}, apomyoglobin\textsuperscript{33} and lysozyme.\textsuperscript{33} Accordingly, it is from these studies clear that four-helix bundles and other helical proteins folds via intermediates. The time-resolved fluorescence study and the CD measurements of (KE2D15)\textsubscript{2} strongly indicate that the intermediate is a molten globule, with a partly solvated hydrophobic core with $\alpha$-helical structure.

The small scale motions observed in the 12 amino acid long helical segment are entirely compatible with reorganization of the backbone of the peptide, which occurs on the same timescale as the large scale motions observed in the tertiary structure. This might indicate the backbone reorganization is coupled with the formation of the tertiary structure. The nascent kinetic phases observed in cytochrome c, apomyoglobin and lysozyme were attributed to the coil–globule transition\textsuperscript{33,35}, however, it cannot be excluded that the nascent kinetic phase is compatible with a reorganization of the backbone, e.g., a $3_{10}$–$\alpha$-helix transition. The denatured cytochrome c state has in fact a $3_{10}$-helix reminiscent shape, which turns into an $\alpha$-helical shape within 2.5 ms.\textsuperscript{35} Stepwise contraction of the unfolded chain through several helical backbone states is, as already mentioned, compatible with the framework model if more helical states than the $\alpha$-helix are incorporated in this model. The hydrophobic effect causes contraction of the backbone into a shorter helix state, through desolvation, which might cause energetic barriers in the folding funnel.\textsuperscript{55-57}

The function of the disorganized helix-loop-helix structure in connection with molecular recognition

The time-resolved structural studies of the four-helix peptide and the WLL12 peptide have shown that the tertiary as well as the secondary structures are disorganized, meaning that these polypeptides make conformational transitions corresponding to folding and unfolding steps of the polypeptides’ structural motifs. Moreover, water is involved in these transitions, which can be regarded as sequential solvation and desolvation steps. Accordingly, the disorganized protein structures fluctuate between more folded and less folded states. Concerning the helix-loop-helix monomer it is now known, from the studies of the four-helix bundle at native and denaturing conditions, that it has a disorganized tertiary structure.

When the helix-loop-helix monomer recognizes and binds to carbonic anhydrase, new bimolecular tertiary structures are developed between the binding polypeptides. Moreover, the formation of the bimolecular tertiary structure occurs via a bimolecular molten globule like intermediate. The formation of new bimolecular tertiary structure shows that the helix-loop-helix monomer and the carbonic anhydrase has a coupled binding and folding
funnel. The disorganized tertiary helix-loop-helix structure is, in this respect, functional in the recognition process where the recognition and binding corresponds to a disordered-to-ordered transition, which occur via partly folded intermediates.
I would like to thank my advisor Hans-Christian Becker for sharing his knowledge, advice and thorough scrutiny. I also thank my co advisor Leif Hammarström for good advice and growing support, and all people at the Fotomol and Physical Chemistry departments.
8. Summary in Swedish

Proteiner är en grupp av biomolekyler som har flera viktiga funktioner i cellen. Ett proteins struktur är nära sammankopplad med dess funktion, och man har ansett att funktionen är entydig med proteinets tredimensionella struktur. Nya rön visar emellertid att fluktuationer i proteinstrukturen också är betydelsefulla för att ett enzym skall fungera. I samband med funktionen och strukturella fluktuationer uppstår frågan på vilka sätt fluktuerar proteinstrukturen, och hur snabbt detta går.

Utöver detta finns en stor grupp av så kallade oorganiserade proteiner, som inte har någon entydig struktur, men som ändå utför viktiga cellulära funktioner, i t.ex. överföring av information genom bindning och igenkänning till ett annat protein. Frågan är då hur den molekylära igenkänningsprocessen går till.


Utöver tertiärstruktureren, vilken är en kombination av olika sekundärstrukturer, så har en proteinmodell med bara sekundärstruktur valts för att studera vilka fluktuationer som är relaterade till densamma. På en peptid, med känd α-helikal struktur mättes variationerna i peptidens längd. Längdvariationerna motsvaras av att den undersöka peptiden byter konformation mellan tre olika helixstrukturer, vilket troligtvis sker på kort mikrosekundstidsskala eller långsammare.

Avhandlingen beskriver inte bara hur de valda peptidmodellerna fluktuerar utan också hur nya och mer organiserade proteinstrukturer uppstår när två proteiner känner igen och binder till varandra. Med tidsupplöst laserspektro-
skopi kan man identifiera hur igenkännningen och bindningen sker stegvis av den ena liganden, som har ett "helix-loop-helix" motiv. Stegen i igenkännningsprocessen motsvaras av att den bimolekylära strukturen blir mer ordnad, mer väldefinierad och partiellt dehydratiserad innan den fullt utvecklade och dehydratiserade strukturen i proteinet komplexet uppstår. Den observerade igenkänningsreaktionen motsvaras av en kombinerad bindnings och veckningsprocess, som föreslagits som en mekanism för igenkännningsprocessen i den vetenskapliga litteraturen.
9. References


86


Acta Universitatis Upsaliensis

Digital Comprehensive Summaries of Uppsala Dissertations from the Faculty of Science and Technology 683

Editor: The Dean of the Faculty of Science and Technology

A doctoral dissertation from the Faculty of Science and Technology, Uppsala University, is usually a summary of a number of papers. A few copies of the complete dissertation are kept at major Swedish research libraries, while the summary alone is distributed internationally through the series Digital Comprehensive Summaries of Uppsala Dissertations from the Faculty of Science and Technology.

(Prior to January, 2005, the series was published under the title “Comprehensive Summaries of Uppsala Dissertations from the Faculty of Science and Technology”.)