Unfolded, Misfolded, and Self-Organized Short Alanine-Rich Peptides: Implications for Fundamental Science, Human Disease, and Biotechnology

A Dissertation
Submitted to the faculty of Drexel University
In partial fulfillment of the
requirements for the degree of
Doctor of Philosophy
In
The Department of Chemistry

By
Thomas J. Measey
B.S., Drexel University, 2005
M.S., Drexel University, 2008
June 2010

Dedications

This thesis is dedicated to my parents, for their unconditional love and support, their insistence on my well-being and happiness, and their continued willingness to make selfless sacrifices for their children. I would also like to dedicate this thesis to my saintly grandmother, Alice Urbaniec (Babci), and to the memory of my grandfather, Eugene Urbaniec (Dzaidzi).

Acknowledgements

I would like to first acknowledge my adviser, Prof. Reinhard Schweitzer-Stenner. To say the least, I don't think I could have found a better adviser/mentor. Reinhard has been pivotal to my success as both a graduate student and researcher, and to my development of skills and sentiment necessary for a continued career in scientific research. In addition to being a mentor and supervisor, Reinhard also become a good friend, always eager to help and offer genuine support, regardless of the situation or problem. His intellectual curiosity and enthusiasm certainly made the journey through graduate school delightful.

Moreover, various people have contributed to the research and work presented herein. In addition to my adviser and various members of the Schweitzer-Stenner lab, these include Dr. Fatma Eker and Prof. Kai Griebenow during their time at the University of Puerto Rico (Río Piedras), for their assistance during my initial learning/training process; Prof. Claudia Forte and Dr. Silvia Pizzanelli during their time at the Istituto per I Processi Chimico-Fisici (Pisa, Italy), Prof. Frank Jordan and Lazaros Kakalis during their time at Rutgers University (Newark), Prof. Sean Decatur and Kathryn Smith during their time at Mt. Holyoke College, and Prof. Konstantin Kornev and Vijoya Sa during their time at Clemson University.

I would like to acknowledge those individuals whom I have had the pleasure of working with during various periods of my graduate career. These include Dr. Andrew Hagarman, Isabelle Dragomir, Ronak Shah, Dan Verbaro, and Jon Soffer, for making my experience at Drexel University fun, productive, and memorable. In addition, I would like to acknowledge those individuals who had no direct connection to my research or this thesis, yet I owe a great deal of gratitude. These include various family members and friends. In particular my parents, for all of their sacrifices, my siblings, Michael, Kristin, Mariah, and Mikayla; Babci and Dzaidzi,

Ms. Brandee Lozito, various Aunts and Uncles, cousins and kin. Also, my long-time friends, Joe Ponisciak, Gary Martin, and George Chamberlin.

Finally, part of my research was funded by the National Science Foundation (NSF) from a grant to Reinhard Schweitzer-Stenner (CHEM - 0804492).

Table of Contents

LIST OF FIGURES

LIST OF TABLES

LIST OF ABBREVIATIONS

A:	Alanine
Aβ:	Amyloid β
AFM:	Atomic Force Microscopy
ATR:	Attenuated Total Reflectance
C:	Cysteine
CCD:	Charge Coupled Device
COSY:	Correlation Spectroscopy
Cytc:	Cytochrome c
D:	Aspartic Acid
DLS:	Dynamic Light Scattering
DNA:	Deoxyribonucleic Acid
DQF-COSY:	Double-Quantum Filtered Correlation Spectroscopy
E:	Glutamic Acid
EM:	Electron Microscopy
ER:	Endoplasmic Reticulum
F:	Phenylalanine
FCS:	Fluorescence Correlation Spectroscopy
FRET:	Förster Resonance Energy Transfer
FTIR:	Fourier Transform Infrared Spectroscopy
FWHM:	Full-Width at Half-Maximum
G:	Glycine
GndCl:	Guanidinium Chloride

H:	Histidine
HSR:	Heat Shock Response
I:	Isoleucine
IDP:	Intrinsically Disordered Protein
IPH:	Isolated Pair Hypothesis
K:	Lysine
L:	Leucine
M:	Methionine
MD:	Molecular Dynamics
N:	Asparagine
NMR:	Nuclear Magnetic Resonance
NOE:	Nuclear Overhauser Enhancement
NOESY:	Nuclear Overhauser Enhancement Spectroscopy
P:	Proline
PEM:	Photoelastic Modulator
PGA:	Poly-L-Glutamic Acid
PLK:	Poly-L-Lysine
PLP:	Poly-L-Proline
PLQ:	Poly-L-Glutamine
PolyQ:	Poly-L-Glutamine
PPII:	Polyproline II
Q:	Glutamine
R:	Arginine
REMD:	Replica Exchange Molecular Dynamics
RNA:	Ribonucleic Acid

ROA:	Raman Optical Activity
S:	Serine
SAXS:	Small Angle X-Ray Scattering
SRE:	Self Recognition Elements
T:	Threonine
TDC:	Transition Dipole Coupling
TK Hypothesis:	Tiffany/Krimm Hypothesis
UPR:	Unfolded Protein Response
UV-CD:	Ultra-Violet Circular Dichroism
UV/vis:	Ultra-violet/Visible
V:	Valine
VCD:	Vibrational Circular Dichroism
W:	Tryptophan
Y:	Tyrosine

ABSTRACT

Protein folding is the reversible transition by which an unordered polypeptide chain attains its functional 3-D native structure. A detailed understanding of the principles which govern the protein folding process, such as how sequence codes for structure, remains elusive. Achieving a complete picture of the folding process requires information regarding structural preferences in the unfolded state. Moreover, understanding the principles which govern protein aggregation is of significant biomedical and biotechnological importance. Herein, short alanine-based peptides are used as model systems for studying both the structural preferences in the unfolded state as well as protein aggregation in relation to human disease, and exploitation of the self-assembly process for various biotechnological applications.

It is now a central dogma of protein science that the unfolded state is not conformationally random, as was originally believed, but that, instead, residual structure exists. Here, we elucidate the conformational propensities of alanine in the unfolded state using short alanine-rich peptides as model systems. The intrinsic conformational propensities of alanine, as well as nearest neighbor effects are illuminated using various vibrational spectroscopic methods, combined with NMR results.

Protein and peptide aggregation is affiliated with various seemingly unrelated diseases, including several neurodegenerative diseases and the systemic amyloidoses. It is of current belief that aggregation is a general feature of the protein energy landscape, suggesting that the various unrelated human pathologies linked to protein aggregation are linked by common principles. Herein, fibril formation of a short alanine-based peptide

with no known disease affiliation is probed by vibrational circular dichroism (VCD) spectroscopy. In particular, it is demonstrated that peptide fibrils give rise to VCD intensity enhancement, illustrating the use of the technique as a novel means to probe aggregation kinetics.

In addition to the biomedical relevance, protein and peptide self-assembly can be exploited as a means of constructing biomaterials with inherent biofunctionality. In this regard, oligopeptide-based hydrogels have shown potential as drug delivery systems and tissue engineering scaffolds. Herein, the unique properties of a novel class of self-assembling alanine-rich oligoopeptides are presented. In particular, it is demonstrated that conformational instability can be exploited to tune the physicochemical properties of hydrogels formed by such systems, for the potential use in various biotechnological applications.

CHAPTER 1

Intrinsic Conformational Propensities in the Unfolded State

1.1 Synopsis

A complete understanding of the unfolded state of proteins and polypeptides requires the systematic investigation of the conformational propensities of amino acids in unfolded chains. Herein, the combined use of FTIR, VCD, Raman, and NMR spectroscopic methods is employed to determine the intrinsic conformational propensities of alanine in short alanine-based peptides. To this end, an excitonic coupling model is employed to simulate the amide I' band profile of various short alanine-rich peptides. These simulations are constrained using $^{3}J_{C\alpha HNH}$ coupling constants obtained from NMR experiments.

Results from this study affirm that alanine has a clear propensity to adopt polyproline II (PPII)-like conformations in the unfolded state, at odds with the random coil model commonly used to describe unfolded proteins. Lysine, on the other hand, was found to behave like a statistical coil in an alanine context, in contrast to polylysine peptides of varying lengths which are known to adopt local segments of PPII-like conformations. Moreover, the PPII propensity of alanine is found to be modified by neighboring charged residues, e.g. lysine, which results in a decreased sampling of PPII

for alanine. This result emphasizes that nearest neighbor interactions play a crucial role in individual amino acid propensities.

1.2 Background Information

1.2.1 Protein Folding

Protein folding is the reversible transition by which an unordered polypeptide chain attains its fully functioning 3-D native structure, as depicted in **Figure 1**. Many questions regarding the principles which govern the folding process, including a detailed understanding of how the amino acid sequence codes for the 3-D native structure, remain elusive. From a thermodynamic viewpoint, protein folding is the simple two-state reaction between the unfolded (U) and folded (F) states of a polypeptide chain, e.g. U ↔ F. Owing to the size of typical proteins, however, the folding reaction tends to be much more complex. Ways in which the protein folding field is emerging is comparable to the progress made on small molecule chemical reactions, where knowledge of basic potential energy surfaces between reactants (unfolded peptides) and products (folded protein) is a priority (1). Owing to the interplay between the large entropic and enthalpic contributions to the free energy of the protein system during folding, the free energy of the native state tends to be only slightly lower than that of the unfolded state, so that the protein folding reactions are much more complex than typical chemical reactions of small molecules (*1*).

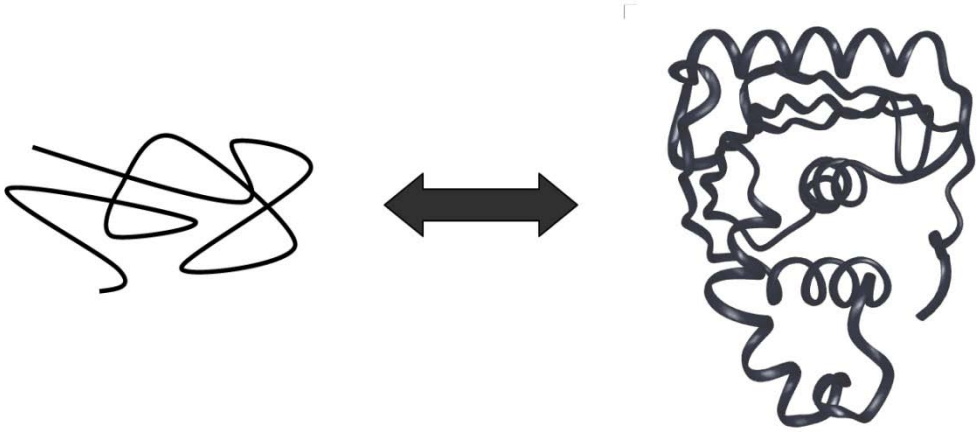


Figure 1: Typical schematic of protein folding depicting the reversible transition between the unfolded (left) and folded state (right).

In vivo, the primary amino acid sequence of any protein is encoded by DNA in the cell nucleus (2). This genetic information is transcribed from DNA to RNA, where the information is then translated into a polypeptide chain from the corresponding nucleotide sequence. Some proteins begin to fold while still attached to the ribosome, i.e. co-translationally, while others fold only after they have left the ribosome. *In vivo*, proteins fold with the aid of numerous chaperones, some of which interact with the nascent polypeptide chain as it emerges from the ribosome, while others guide the newly synthesized chains to other locations for folding at later stages (3). Thus, in the same way that biochemical reactions are catalyzed by enzymes, protein folding is 'catalyzed' by molecular chaperones *in vivo* (4).

1.2.2 Thermodynamic Hypothesis of Protein Folding

In vitro protein folding studies can yield fundamental mechanistic insight into the folding process without needing to consider contributing factors from the often highly crowded biological environment. The history of protein folding studies *in vitro* dates back to the early 1960's, when Charles Anfinsen and associates first demonstrated that chemically denatured ribonuclease can restore itself back to its native fold when re-exposed to conditions which favor the native state (5). Anfinsen reasoned that proteins fold in such a way so as to ultimately yield a structure with a minimum Gibbs energy (ΔG). Thus, in Anfinsen's "thermodynamic hypothesis" of protein folding, the native fold is that which attains a minimum free energy. From these early experiments, Anfinsen concluded that the native conformation is encoded by the protein's amino acid sequence. Although methods of secondary structure prediction have been around since the early 1970's, the most famous being the Chou and Fasman method (6-7), a complete and detailed understanding of how a protein's native fold is encoded in the primary amino acid sequence remains a mystery of the biological sciences (8).

1.2.3 Levinthal Paradox

A contradiction to the thermodynamic hypothesis of protein folding was put forth by Cyrus Levinthal in 1968. Levinthal reasoned that even if each residue in a sequence containing 100 amino acid residues can sample only two distinct conformations, so that the total # of conformations is $2^{100} = 10^{30}$, and that if it is assumed that each conformation

can be sampled in 1 picosecond (a somewhat conservative assumption), then it would take > 10^{10} years to fold, that is, close to the lifetime of the universe (9). In contrast to Levinthal's estimate, proteins typically fold in the microsecond to second regime. In particular, α-helical segments and β-turns can form in μs or less (10-11), and simple proteins have even been shown to fold from their unfolded state in these short times (12-13). Based on his thought experiment, Levinthal suggested the possibility of protein folding pathways which dominate the "classical view" of protein folding (14). In this view, proteins fold kinetically, that is via a series of sequential steps, to ultimately attain their native 3-D structure with a minimum free energy, ΔG.

1.2.4 The 'New View' of Protein Folding: Energy Landscapes and Folding Funnels

Recently, the concepts of energy landscapes and folding funnels have been introduced to help visualize the protein folding process. In this 'new view' of protein folding, a typical protein energy landscape consists of a multidimensional funnel in which various minima exist along this energy funnel, with the native state residing at the global energy minimum. In particular, the unfolded state and folded states reside at the top and bottom of an energy funnel, respectively (15-17).

During folding, an unfolded polypeptide chain undergoes a stochastic search of the many accessible conformations as it advances down the energy funnel. Since native-contacts are more energetically favorable and thus more persistent than non-native ones, the conformational search is greatly reduced, and the polypeptide ultimately attains its

global free energy minimum, i.e. the native state, (See **Figure 1**) (1, 3-4, 16). While two state transitions have been documented between the unfolded (U) and folded (F) states, e.g. U↔ F, the folding process is much more complex, especially for larger systems (more than 100 amino acid residues), where one or more kinetic intermediates are often populated (3, 18). Various techniques are available which can provide information about residues involved in folding intermediates and transition states (3, 19) and their respective stabilities. The more well-known of these techniques include pulsed hydrogen exchange NMR (18) and the protein engineering method (20), where the latter attempts to identify specific amino acids involved in native-like contacts in the transition or intermediate states by introducing point mutations and probing the subsequent effects on the folding rate (19).

The conversion of natively folded proteins into well-ordered aggregates is a common phenomenon of the protein folding process, and is affiliated with various debilitating human pathologies. Recently, the funnel concept of protein folding has been modified to include protein aggregation, which is now considered a generic feature of all proteins and peptides. This extension of the protein energy landscape to include the aggregated state will be addressed in **Chapter 2**.

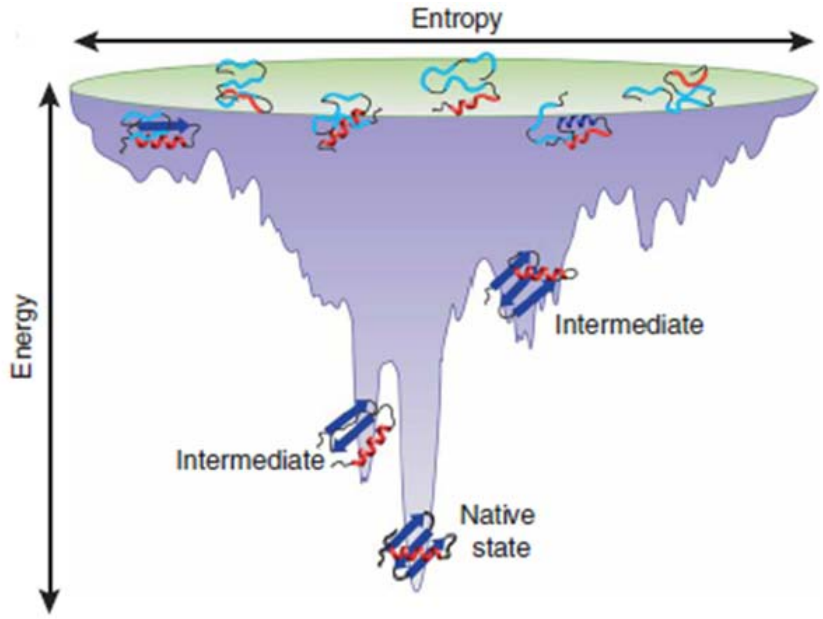


Figure 2: **Energy Landscape of a typical protein, where the native state resides at the free energy minimum of the funnel. This figure was adapted from ref(21).**

1.2.5 Protein secondary structure

The polypeptide backbone consists of three degrees of freedom resulting from the three dihedral angles, ω, ϕ, and ψ, defined by rotation about the C_α-C'-N-C_α, C'-N-C_α-C', and N-C_α-C'-N backbone atoms, respectively (2). Owing to its partial double bond character, the peptide group (H-N-C'-O) tends to be very rigid, making the four atoms nearly co-planar. Aside from proline, polypeptides formed from the other naturally occurring L-amino acids prefer the formation of peptide groups with a *trans*- conformation, so that ω is constrained to 180°. As such, the secondary structure of the polypeptide chain is usually governed only by the dihedral angles ϕ and ψ, as depicted in **Figure 3**.

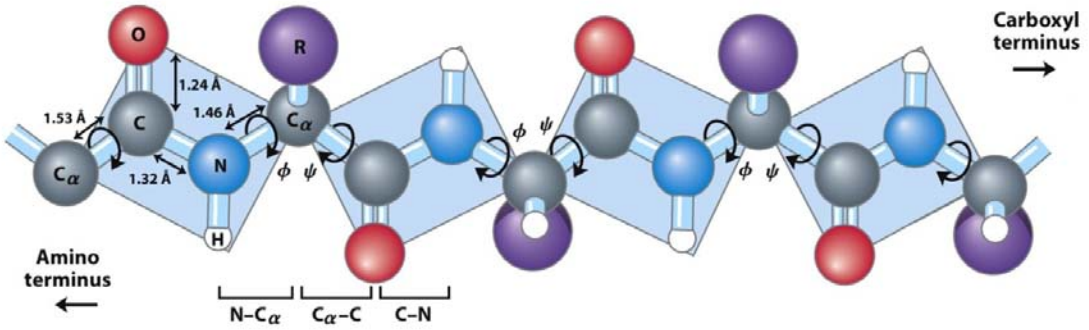


Figure 3: Polypeptide backbone, showing the dihedral angles (ϕ, ψ). Figure was adapted from ref(2).

Due to steric restrictions, however, not all possible combinations of ϕ and ψ are allowed. A convenient way to view the allowed conformations of a protein is via a Ramachandran map, developed by G.N. Ramachandran in the early 1960's (22). **Figure 4** shows the Ramachandran map for proteins, illustrating the allowed regions, as well as pointing out the locations of various common elements of secondary structure.

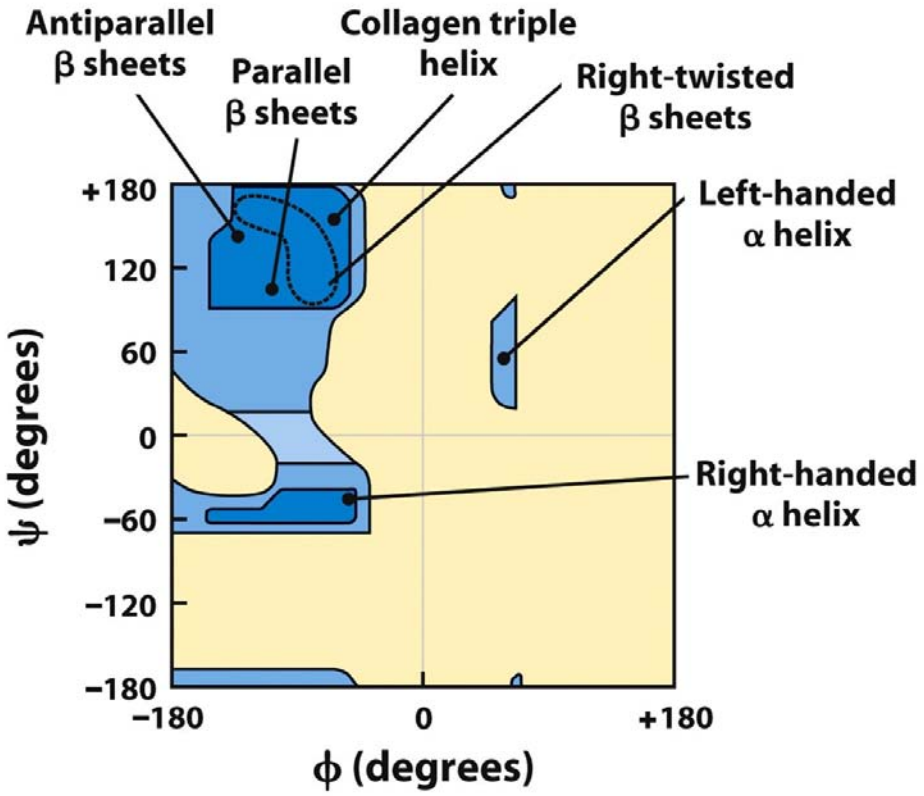


Figure 4: Ramachandran map showing the sterically allowed regions (blue). The locations of various secondary structures are also indicated. This figure was adapted from ref(2).

The simplest method of qualitatively deducing the elements of secondary structure present in a particular polypeptide is ultraviolet circular dichroism (UV-CD) spectroscopy (23-26). UV-CD measures the differential absorption of right and left-circularly polarized light, in the regime of electronic transitions associated with the peptide group. The distinct basins of the Ramachandran map are indicative of well-defined polypeptide secondary structures, such as α-helices and β-sheets. Each of these

major secondary structures exhibits a distinct UV-CD spectrum, as depicted in **Figure 5** for the α-helix, β-sheet, β-turn, and PPII conformations.

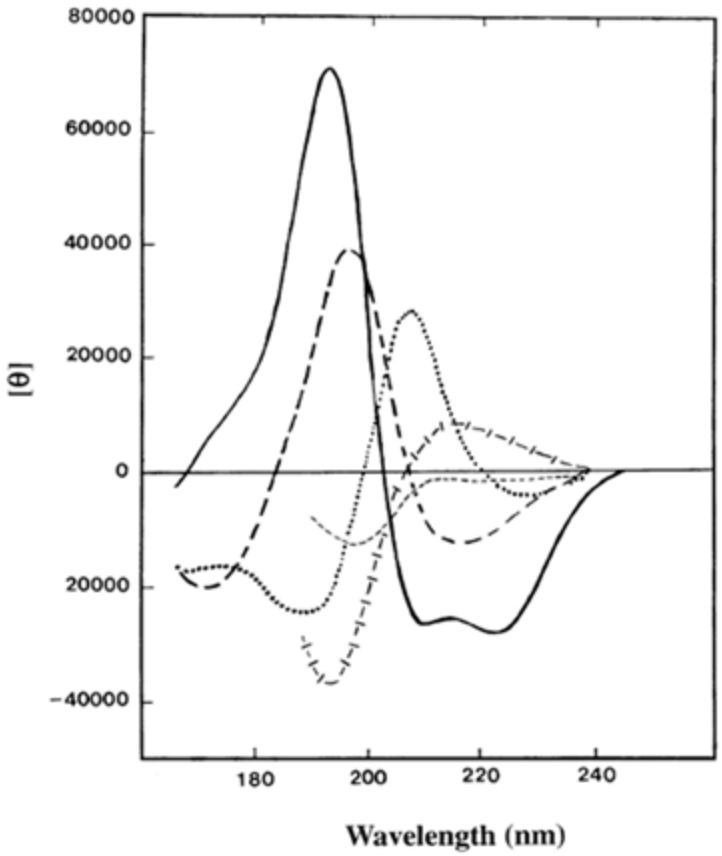


Figure 5: **Far-UV circular dichroism (CD) spectra of various secondary structures. Taken from ref(23) and modified.**

1.2.6 The Unfolded State

Determining the structure of folded proteins is somewhat of a routine endeavor, at least for non-membrane proteins. As crystallization is often required for structure determination, membrane proteins present a challenge owing to the requirement of surfactants and/or denaturants required for their solubility, which can interfere with crystallization (*27*). High-resolution techniques such as X-ray crystallography and nuclear magnetic resonance (NMR) spectroscopy can provide atomic resolution of even large proteins. Approximately 62,000 protein crystal structures have been solved to date (*28*). Due to these high resolution methods a great deal is known about the structure and dynamics of the folded state. The unfolded state of proteins, on the other hand, has eluded such a detailed structural and energetic characterization. This is due to the inherent heterogeneity of the unfolded state and complex dynamics (8).

In spite of the elusive nature of the unfolded species, a complete understanding of the mechanistic principles which govern the protein folding process requires a detailed thermodynamic characterization of the unfolded state, including both structure and dynamics (8). Classic textbook wisdom dictates that the unfolded state is purely random, such that all amino acid residues, with the exceptions of glycine and proline, can sample the entire sterically allowed Ramachandran space, thus suggesting that structural studies of the unfolded state are futile. The presence of residual structure in the unfolded state, however, has significant implications for protein folding theories, as the number of conformations needed to be sampled by a polypeptide chain would be greatly reduced (*29*).

1.2.7 Unfolded Proteins and the Random-Coil Model

The structure of unfolded proteins was first pioneered by Tanford using chemical denaturants to unfold native proteins (*30*). Tanford observed that denatured proteins behave like polymers in a good solvent (*30-32*), in contrast to folded proteins globules which behave like polymer chains in poor solvents (*30*). This pioneering work led the foundation for treating unfolded proteins with a statistical model from polymer theory, now known as the random- or statistical-coil model, developed by Paul J. Flory (*33*). Due to implicit steric restrictions of the peptide backbone, the term statistical coil is preferred to random coil (*34-35*).

In the statistical-coil model, the length of an unfolded polypeptide chain length and radius of gyration are related by the following power-law (*36-37*):

$$R_g = R_o N^{\nu} \tag{1}$$

where N is the number of amino acid residues, R_o is measure of the excluded volume of the polymer chain (which allows for favorable solvent interactions), and ν is an exponential scaling factor which depends on the quality of the solvent, where $\nu = 0.33$ for a poor solvent, $\nu = 0.5$ for an ideal solvent, and $\nu = 0.6$ for a good solvent (*37-38*). Typically, $\nu = 0.6$ and R_o = 1.93 ± .001 are used in the treatment of unfolded and denatured polypeptides (*36*). Such a random-coil scaling of the radius of gyration

suggests that the unfolded polypeptide does not depend on the amino acid composition of the peptide chain.

The statistical-coil theory of unfolded polypeptides is based on the assumption that each pair of (ϕ,ψ) dihedral angles is independent of neighboring pairs. This is known as the isolated pair hypothesis (IPH). The IPH hypothesis suggests that the total number of possible conformations which can be adopted by a polypeptide can be determined as the product of the number of conformations sampled by each amino acid residue. The isolated pair hypothesis may be formulated mathematically as (*39*):

$$S_n \leq nS_{n,i} \tag{2}$$

where S_n is the entropy of a polypeptide with n residues, and $S_{n,i}$ is the entropy of residue i. The isolated pair hypothesis is valid when the equality holds.

1.2.8 The Tiffany and Krimm Hypothesis (TK Hypothesis)

Homopolypeptides of charged amino acid residues, e.g. lysine, glutamic acid, etc., adopt different secondary structures depending on the protonation state of the side chains. In particular, poly-L-lysine (PLK) can adopt an α-helical structure in water at pH values above the pK_a of the lysine side chain. Below the pK_a, however, PLK was believed to adopt a more 'random-coil' structure due to electrostatic repulsion between the positively

charged lysine residues. In 1968, Tiffany and Krimm showed that the UV-CD spectra of poly-L-glutamic acid (PGA) and PLK, at pH values where the side chains were charged, resemble the UV-CD spectrum of *trans*-poly-L-proline (PLP), which was known to adopt an extended left-handed helix, known as poly-L-proline II (PPII) (*40-41*). Thus, the Tiffany and Krimm hypothesis (TK hypothesis) dictates that unfolded peptides actually exhibit a significant amount of PPII-like structure. Tiffany and Krimm suggested that segments of local order exist in polypeptides which were thought to adopt the so-called 'random-coil' structure, with admixtures from conformational sampling of the sterically allowed regions of the Ramachandran map obtained for the alanine dipeptide (*22, 42*).

1.2.9 Model Systems for Studying the Unfolded State

Following the findings of Tiffany and Krimm in the late 1960's and early 1970's, little attention was drawn to their conclusions. In recent years, however, the TK hypothesis has been revived based on the results of numerous experimental studies of short peptides (*43-58*). This rekindling of the TK hypothesis is due to the vibrational circular dichroism (VCD) studies of Dukor and Keiderling on proline oligomers in the early 1990's (*59*). Since then, various experimental studies on short peptides suggest that they significantly sample the PPII conformation in water (*43-58*). Such peptides, which are too short to adopt any stable elements of secondary structure such as α-helices, β-sheets, and β-strands, are continually being used as models for probing the unfolded state of peptides and proteins.

In addition to short peptides, various other model systems have been employed over the years to better characterize the unfolded state of peptides and proteins. These include the charged homopolypeptides such as PLK and PGA studied by Tiffany and Krimm in their seminal work (*40-42*), as well as denatured proteins, natively unfolded proteins, and coil libraries.

Studies of denatured proteins are usually carried out using chemical denaturants, such as guanadinium chloride (GndCl) or Urea (8), but such chemically denatured species might not reflect the true unfolded state of polypeptides. In this regard, natively unfolded proteins, such as synuclein, which is linked with Parkinson's disease, are better suited for studies of the unfolded state of longer polypeptides (*60*). A coil library is essentially a catalog of (ϕ, ψ) dihedral angles sampled by a particular amino acid obtained by analyzing the various protein structures archived within the Protein Data Bank (*28, 61*). It is thought that all contextual dependencies will be eliminated, such that coil libraries are representative of the intrinsic conformational propensities of a particular amino acid. Recently, Serrano (*61*) used a less-restricted coil library, where those pairs of dihedral angles which correspond to elements of major secondary structure, e.g. α-helix, β-strand, β-sheet, etc., were removed, to yield new conformational distributions. Such an approach is more representative of intrinsic conformational propensities (*62*).

1.2.10 Polyproline II (PPII) Conformation

The conformation adopted by *trans*-PLP obtained from the crystal structure of Cowan et al. (*63*) is found to be an extended left-handed helix, known poly-L-proline II (PPII), as shown in **Figure 1**(a.). The canonical dihedral angles of PPII are (ϕ, ψ) = (-78°, 145°), so that it is located in the upper-left quadrant of the Ramachandran map (see **Section 1.2.5**). PPII is a 3_1, helix, containing three amino acid residues per helical turn, and therefore exhibits three-fold rotational symmetry, as shown by the axial view in **Figure 1**(b.).

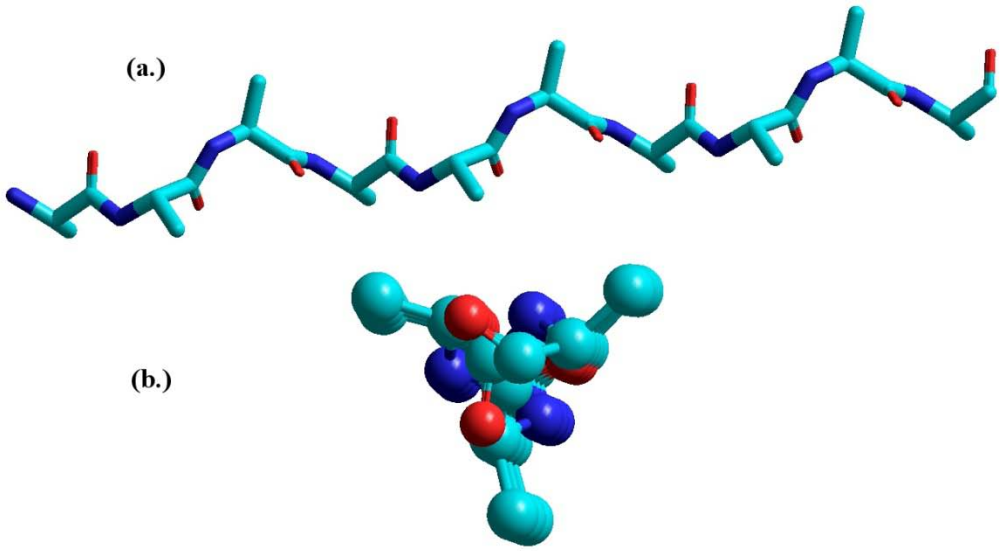


Figure 6: Illustration of deca-alanine (A_{10}) in the polyproline II (PPII) conformation. (a.) Side view showing the extended left-handed helical conformation and (b.) a view down the axis of the polypeptide chain, illustrating the 3-fold rotational symmetry of PPII.

1.2.11 Alanine-Based Peptides

Short alanine-Based peptides have been the subject of intense research activities since the late 1980's. This interest stemmed from the discovery by Marqusee et al. that short alanine-based peptides with at least 12 amino acid residues could adopt stable α-helices in aqueous solution if doped with a limited number of charged residues (*64-65*). The addition of charged residues increases the solubility of the peptide in aqueous solution. Such short alanine-rich sequences also serve as convenient model systems for studying helix ↔ coil transitions of proteins (*66-67*). Short alanine-based peptides have been used extensively in recent years as convenient model systems to study the unfolded state of peptides and proteins (*45*, *68*). This mostly stems from the fact that they are too short to form any stable elements of secondary structure. They are, in essence, intrinsically unfolded, so that they can be used to study the unfolded state under physiological conditions. Moreover, recent experimental and theoretical studies suggest that the unfolded state of short alanine-based peptides cannot be described by Tanford (*30*) and Flory's (*33*) statistical coil model (*36*, *43*, *48*, *69*).

1.2.12 The Reconciliation Problem

Despite numerous theoretical and experimental studies which suggest that conformational preferences exist in the unfolded state, such unfolded polypeptides typically exhibit radii of gyration that scale with random-coil behavior (see **Section 1.2.7**). The reconciliation

problem concerns the question of how one can reconcile the obvious local structure observed for unfolded and denatured peptides and proteins, with the fact that their radii of gyration appear independent of the identity of the amino acid sequence, and suggest rather that they behave more like random or statistical coils (*32, 36-37*).

For example, Tran et al. (*36*) simulated ensembles of conformations which are accessible to peptide and proteins, and showed that, in the limit of excluded volume, a tug-o-war exists between steric overlap, which promotes chain stretching, and conformational entropy, which opposes chain stretching. Minimization of steric overlap in a polypeptide chain results in the formation of local structure, consistent with experimental studies, which indicate a PPII-preference in the unfolded state. Conformational entropy, on the other hand, opposes the formation of long, locally ordered segments, resulting in short persistence lengths and random-coil scaling behavior.

Fitzkee and Rose (*38*) introduced a counterexample to show that polypeptides with rigid segments of native protein structure exhibit random coil statistics when they are interconnected by various flexible hinge residues. These studies clearly indicate that polypeptides with local order can still exhibit random-coil characteristics, and that other criteria are necessary to suggest complete randomness of a polypeptide chain, as was originally pointed out by Tanford (*30*).

1.2.13 Violation of the Isolated Pair Hypothesis

The isolated pair hypothesis (IPH) is a pillar of protein folding (*70*) and helix-coil theories (*71-73*). Recent studies, however, have provided results which clearly violate the IPH (*74-77*). These include studies from the Rose and Sosnick groups. Pappu et al., for example, divided the Ramachandran map into so-called mesostates, and used an all-atom representation to enumerate all of the possible sterically allowed conformations for short alanine-rich peptides (*77*). These authors showed that steric effects can extend beyond nearest-neighbor residues, at least for residues which adopt conformations residing in the lower left quadrant of the Ramachandran map, such as α-helices. Pappu et al. demonstrated that while the IPH holds for residues in the upper left quadrant of the Ramachandran map (PPII and β-strand), that mixtures of residues adopting PPII and/or β-conformations, with residues adopting α-helical conformations, also violates the IPH. Based on their results, these authors suggested that the entropy of the unfolded state is much less than expected based on a random- or statistical-coil model. A more recent study by this group provides further support for this notion (*78*).

The results of Pappu et al., however, were challenged by Ohkubo et al., who argued that the non-nearest-neighbor clashes that occur in the short peptides studied by Pappu et al. are due to end effects, and are expected to be absent in longer peptides (*39*). In particular, Ohkubo et al. demonstrated that the conformational entropy of polyalanines containing 3 – 20 amino acid residues is proportional to the chain length, and thus support the validity of the IPH (*39*).

1.3 Intrinsic Conformational Propensity of Alanine in the Unfolded State

1.3.1 Introduction

The unfolded state of peptides and proteins has been the subject of an increasing number of experimental and theoretical studies (*8, 36, 43, 48, 51, 79-82*). This is due to the discovery of naturally disordered, yet biologically functioning proteins and peptides (*82*), as well as the general relevance for a thorough understanding of the protein folding process. In this context, the possible existence of local residue structure would certainly affect the initial phase of the folding process (*83*). The existence of such locally ordered segments was first proposed by Tiffany and Krimm (see **Section 1.2.8**) based on UV-circular dichroism (UV-CD) measurements on PLP, PLK, and PGA (*41*). The authors concluded that charged polypeptides assume, at least locally, a rather ordered PPII conformation. This notion was later confirmed by vibrational circular dichroism (VCD) studies on a variety of unfolded polypeptides and proteins (*47, 59*).

More recently, the interest of the protein/peptide folding community has focused on the unfolded state of alanine based peptides, after theoretical and experimental results suggested that it cannot be described by the classical statistical coil model of Flory (*33*). In particular, tri- (*43, 49*) and tetra-alanine (*69*) have recently been suggested to adopt predominant fractions of PPII-like conformations in aqueous solution. Thus, alanine has emerged as a paradigm for the breakdown of the statistical coil model.

Moreover, the solution structure of the hepta-alanine polypeptide Ac-$X_2A_7O_2$-NH_2 (XAO, X and O denote diaminobutyric acid and ornithine, respectively) has been a matter of controversy in the current literature. Shi et al., for instance, used NMR and

UV-CD measurements to study the structure of XAO, and interpreted their results as indicating that the individual residues predominantly adopt a PPII conformation at room temperature, with a < 20% population of β conformations at room temperature (*45*). The results of Shi et al. have been corroborated by numerous experimental and theoretical studies on short peptides, all of which revealed a substantial PPII propensity for alanine (*36, 43, 48, 51, 79, 81*). Here the word propensity is used to indicate a quantified mole fraction or probability to adopt a specific conformation. This notion is also in perfect agreement with distributions which Serrano (*61*) and Jha et al. (*80*) inferred from coil libraries, but at variance with a less restricted library investigated by Dobson and associates (*84*). The results of MD simulations are force field dependent, and generally do not reproduce a PPII propensity of alanine without force-field modifications (*85*).

Zagrovic et al. investigated XAO by small angle x-ray scattering (SAXS) experiments and obtained a radius of gyration of 7.4 Å (*86*). If one assumes the 'random walk scaling' between radius of gyration and end to end distance, this value corresponds to an average end-to-end distance of 18.1 Å, which is significantly shorter than what one would expect for a pure PPII structure (radius of gyration = 13.1 Å, end to end distance = 32.04 Å). Zagrovic et al. also performed several molecular dynamics simulations with six variants of the Amber and Gromos force fields (*86*). The simulations reproduced neither the extended, PPII dominated structure, nor the very short radius of gyration obtained from the SAXS data. Makowska et al. combined UV-CD and NMR measurements of XAO (involving the measurement of $^{3}J_{C\alpha HNH}$ coupling constants, NOE's, and chemical shifts) with MD simulations performed with the AMBER 99 force field (*87*). The results lead the authors to conclude that PPII is only one of many

conformations sampled by alanine, along with multiple β-turn structures. Additionally, the residues of XAO were found to sample multiple β-turn structures, without leading to the formation of stable structures with intra-peptide hydrogen bonding. This is a surprising result since alanine is not known for its propensity for turn structures (7). Altogether, the results of Makowska et al. reemphasize the validity of the classical statistical coil model, though their Ramachandran plot looks somewhat different from the random coil distribution of alanine (*84, 88-89*).

Molecular Dynamics (MD) simulations of the alanine dipeptide and polyalanines of varying lengths have increased the confusion in that different force fields yield rather different Ramachandran plots for the very same peptide. Some of these simulations support the high PPII propensity of alanine (*90-91*), while others do not even yield a substantial population of this conformation (*85, 92*). Distributions derived from coil libraries seem to support a PPII preference of alanine, but this depends on how the data set is chosen (*61*). Even those simulations supporting the PPII propensity of alanine indicate an additional sampling of helix-like conformations, which is generally not considered in the analysis of experimental data (*48, 69*).

Recently, Jun et al. used electron spin resonance distance measurements to probe the distance of two spin labels attached to an alanine based peptide in the folded and (thermally) unfolded state, and found the respective value for the latter to be too short for a pure PPII state of the peptide (*93*). Analogously, Tucker et al. used an intrinsic fluorescence resonance energy transfer (FRET) pair to deduce that the supposedly

disordered mastoparan X peptide actually exists in aqueous solution as an ensemble of compact conformations (*94*).

One might try to resolve the above discussed conflicting results by arguing that the PPII propensity of alanine solely applies to very short peptides, i.e. tri- and tetra-peptides, whereas it behaves like a statistical coil in longer ones (*48, 69, 95*), which would violate the IPH. A clarification of this issue has a high bearing on the understanding of both the unfolded state of peptides and proteins, as well as the protein unfolding process, in view of the abundance of alanine in nature, and its high propensity for helix formation (*65*).

In what follows, we combine the use of FTIR, polarized Raman, VCD (vibrational circular dichroism), UV-CD and NMR data to explore the conformations sampled by short alanine-based peptides, in particular the unblocked cationic tetrapeptide H-AAKA-OH (AAKA), and the controversial XAO peptide, containing a hepta-alanine segment. An excitonic coupling model is employed to simulate the amide I' band, where the backbone ϕ angle for each residue is constrained using the $^{3}J_{\mathrm{C\alpha HNH}}$ NMR coupling constants reported in the literature along with a modified Karplus equation.

AAKA was chosen for two reasons. First, obtaining the conformational propensity of lysine as a guest residue in an alanine host system is necessary to reconcile the contextual dependence reported for lysine, since polylysine peptides of varying lengths have been reported to prefer PPII conformations (*41, 50*), while the PPII propensity of the central lysine residue in the tripeptide AKA was found to depend on the C-terminal protonation state (*48*). These contrasting results raise the question of whether

lysine itself has a high PPII propensity, which is of relevance for the understanding of naturally unfolded proteins in which lysine residues occur much more frequently than in folded proteins (*96*). Also, lysine is one of the residues used in helix forming alanine based peptides, because it increases the overall helix stability, most likely by dehydrating the adjacent part of the backbone (*97*).

We performed a structure analysis of AAKA and XAO by exploiting the excitonic coupling between the amide I' modes of the peptide groups, which produces conformationally sensitive IR, Raman and VCD band profiles. For the simulation of the profiles, we used a more extended approach than that employed for tri- and tetrapeptides in previous studies by explicitly considering coexisting conformations for the individual residues (*48, 69*). The statistical modeling was constrained by the $^{3}J_{C\alpha HNH}$ coupling constants of the residues obtained from ^{1}H NMR experiments. For XAO, which remains at the center of the debate, results from SAXS experiments were also used as constraints.

Subsequently, we utilized the same approach to reanalyze earlier published amide I' band profiles of tetra- and trialanine, in conjunction with corresponding NMR data. For short alanine peptides (AAA, AAAA, and AAKA), the results led us to conclude a modified view about the conformational propensity of alanine, which is in excellent agreement with MD simulations of Gnanakaran and Garcia (*90*). With respect to lysine, our results show that it exhibits a reduced PPII propensity in an alanine context, e.g. AAKA.

The conformational manifold of XAO on the other hand was found to be quite heterogeneous. The best reproduction of the experimental data could only be achieved by

assuming an ensemble of conformations, which contains various β-turn conformations (~26%), in addition to β-strand (~23%) and PPII (~50%) conformations. PPII is the dominant conformation in segments not involved in turn formations. Most of the residues were found to sample the bridge region connecting the PPII and right-handed helix troughs in the Ramachandran plot, which is part of the very heterogeneous ensemble of conformations generally termed as type IV β-turn. Technically, the current study shows that the combination of NMR and vibrational spectroscopy provides a powerful tool to explore the conformational landscape of peptides.

1.3.2 Vibrational Spectroscopy as a Tool to Probe Residual Structure in the Unfolded State

Protein structure determination remains as one of the most important endeavors in molecular biology and biochemistry. In this regard, x-ray crystallography and nuclear magnetic resonance (NMR) spectroscopy are the most common techniques used to determine the structure of a protein, representing ~87% and ~ 12% of the ~62,000 solved protein crystals, respectively (*28*). Intrinsically disordered/unfolded peptides and proteins typically adopt an ensemble of conformations in solution, so that even if crystallization can be achieved on such a dynamic system, the resulting structure may only represent one of many conformations occurring in the solution ensemble (*98*). Owing to the ensemble of conformations sampled, complementary techniques are often required to gain a clear picture of the structural properties of intrinsically disordered and unfolded polypetides in solution. In this regard, vibrational spectroscopic techniques

such as Fourier transform infrared (FTIR), vibrational circular dichroism (VCD), and Raman spectroscopies can be coupled to quantify distinct elements of secondary structure that contribute to a protein's overall structure (*99-101*). Owing to the larger time windows ($> 10^{-12}$ s) made available by Raman and IR spectroscopies, such methods are particularly well-suited for exploring fast dynamic processes such as conformational fluctuations (*102*).

1.3.3 Amide I Normal Mode

The normal modes of polypeptides have been well-documented (*99, 103*). The most structurally sensitive vibrational mode is the amide I mode, which results predominantly from backbone carbonyl stretching (CO s), with some contributions from N-H in plane bending (NH ip b), CN stretching (CN s), and CH bending (CH b) (*104-107*). The structural sensitivity of amide I arises from the ability of the backbone carbonyl oxygen to participate in hydrogen bonding with other peptide groups as well as the solvent (e.g. water). In deuterated solvents, the amide I downshifts due to a result of decoupling of the NH s, and is denoted as amide I' (*102*). The position and intensity of the amide I' have been shown to be dependent on the type of amino acid (*108*).

In addition to the sensitivity of amide I due to hydrogen bonding and amino acid type, the major structural sensitivity of the amide I mode arises from delocalization along the polypeptide backbone, which results predominately from transition dipole coupling (TDC) and through-bond coupling (*109-110*). These delocalized excited vibrational

states are termed excitonic states, and can be described using a coupled oscillator model. The following section will introduce the excitonic coupling model, which can be used to simulate the vibrational spectra (e.g. infrared (IR), Raman, and vibrational circular dichroism (VCD)) of short polypeptides.

1.3.4 Theory of Excitonic Coupling Model to Describe Delocalization of Amide I

Excitonic coupling of the amide I mode has previously been successfully described by using a coupled quantum mechanical oscillator model (*69*, *79*, *111-113*). Each vibrationally excited amide I oscillator in a polypeptide containing *n* amide I oscillators gives rise to a manifold of *n* excited vibrational states, termed excitonic states. These excitonic states can be written as a linear combination of local excited amide I oscillators, such as:

$$\chi'_i = \sum_{i=1}^{n} a_{ij}\chi_j \qquad (3)$$

where χ'_i is wavefunction of the *i*th excitonic state, χ_j is the local uncoupled excited state wavefunction of the *j*th oscillator, and a_{ij} are the mixing coefficients, which reflect the contribution of the local excited state wavefunctions to the *i*th excitonic state.

The Schrödinger equation for a peptide with *n* residues can be written as:

$$\left(\hat{H}_o + \hat{H}_{ex}\right)|\chi'\rangle = E|\chi'\rangle \tag{4}$$

where $\hat{H}_o$ and $\hat{H}_{ex}$ are the Hamiltonian operators for the unperturbed local amide I oscillators and excitonic coupling, respectively, E is the set of *n* eigenenergies, and $|\chi'\rangle$ is a state vector comprising the *n* excitonic amide I' states (*114*).

The total Hamiltonian, $\hat{H}_T = \hat{H}_o + \hat{H}_{ex}$, therefore contains the interactions between the oscillator of the *j*th residue and its neighboring residues, i.e. $j\pm1$, $j\pm2$, $j\pm3$, and $j\pm4$, namely:

$$\hat{H}_T = \begin{bmatrix} \tilde{\nu}_1 & \Delta_{12} & \delta_{13} & \delta_{14} & \delta_{15} & . & . & . & . & 0 \\ \Delta_{21} & \tilde{\nu}_2 & \Delta_{23} & \delta_{24} & \delta_{25} & \delta_{26} & . & . & . & 0 \\ \delta_{31} & \Delta_{32} & \tilde{\nu}_3 & \Delta_{34} & \delta_{35} & \delta_{36} & \delta_{37} & . & . & 0 \\ \delta_{41} & \delta_{42} & \Delta_{43} & \tilde{\nu}_4 & \Delta_{45} & \delta_{46} & \delta_{47} & \delta_{48} & . & 0 \\ \delta_{51} & \delta_{52} & \delta_{53} & \Delta_{54} & . & . & . & . & . & . \\ . & \delta_{62} & \delta_{63} & \delta_{64} & . & . & . & . & . & . \\ . & . & . & . & . & . & . & . & . & . \\ . & . & . & . & . & . & . & . & . & \delta_{n-2,n} \\ 0 & . & . & . & . & . & . & . & \tilde{\nu}_{n-1} & \Delta_{n-1,n} \\ 0 & 0 & 0 & 0 & . & . & . & \delta_{n,n-2} & \Delta_{n,n-1} & \tilde{\nu}_n \end{bmatrix} \tag{5}$$

Here, the term 'residue' is defined as a peptide group and the amino acid side chain. For amide I oscillators which are more than 4 residues apart, the coupling between them can be considered as negligible except in the case of intermolecular coupling, e.g. in β-sheets. In the matrix representation of the total Hamiltonian above, $\tilde{\nu}_j$ is the energy of the unperturbed local oscillators, *j*, $\Delta_{j,j\pm1}$ represents nearest-neighbor interactions, while $\delta_{j,j'}(j' = j \pm 2, j \pm 3, j \pm 4)$ accounts for the coupling between non-nearest-neighbors. The latter is assumed to result solely from through-space interactions, i.e. transition

dipole coupling (TDC), while the former involves both through-bond and through-space coupling (*109*). TDC can be calculated by utilizing the following equation (*103*):

$$\delta_{jj'} = \xi \left(\frac{\vec{\mu}_j \cdot \vec{\mu}_{j'}}{\left|\vec{T}_{jj'}\right|^3} - 3 \frac{\left(\vec{\mu}_j \cdot \vec{T}_{jj'}\right)\left(\vec{\mu}_{j'} \cdot \vec{T}_{jj'}\right)}{\left|\vec{T}_{jj'}\right|^5} \right) \quad (6)$$

where $\vec{\mu}_j$ and $\vec{\mu}_{j'}$ are the vibrational transisiton dipole moments of the jth and j'th residues, respectively, and $\vec{T}_{jj'}$ is the distance vector connecting the centers of the jth and j'th dipole moments. ξ is a constant which depends on the units used to express the transition dipole moments and distance vector (*103*). In electrostatic units (esu), it is equal to 9.047 x 10^{15} esu^{-1}(*115*).

Diagonalization of the above Hamiltonian yields eigenfunctions and thus the mixing coefficients, a_{ij}, which can be used to calculate the Raman tensor and the transition dipole moment of the ith excitonic state:

$$\hat{\alpha}'_i = \sum_{j=1}^{n} a_{ij} \hat{\alpha}_j \quad (7)$$

$$\vec{\mu}'_i = \sum_{j=1}^{n} a_{ij} \vec{\mu}_j \quad (8)$$

Here, the Raman tensors and transition dipole moments for the local modes, j, are written as (*116*):

$$\hat{\alpha}_j = \hat{\alpha} = \begin{pmatrix} a^* & c^* & 0 \\ c^* & 1 & 0 \\ 0 & 0 & d \end{pmatrix} \quad (9)$$

$$\vec{\mu}_j = \vec{\mu} = \begin{pmatrix} \mu_o \cdot cos\vartheta \\ \mu_o \cdot sin\vartheta \\ 0 \end{pmatrix} \quad (10)$$

Where μ_o is the magnitude of the transition dipole moment of the amide I mode, ϑ is its angle with respect to the *x*-axis of the coordinate system. The tensor elements and dipole strengths of the local amide I' oscillators were obtained from a previous study on AX and XA dipeptides (X signifies one of the 20 naturally occurring amino acid residues) (*108*). The elements, a^*, c^*, and d, are expressed in units of α_{yy}, which is set to 1. **Eq. 7** describes the Raman tensor of the molecular frame depicted in **Figure 7** (*117*). In this frame, the respective nitrogen atom is at the origin of the coordinate system, and the x-axis is chosen to be co-linear with the NC_α bond. The rotational axis of x is therefore the rotational axis associated with the dihedral angle ϕ. The y-axis is (nearly) coplanar with the peptide group, and the z is the out-of-plane coordinate. The small deviation from coplanarity (~1.1°) results from the fact that the NC_α bond is not exactly coplanar with the peptide group (*117*).

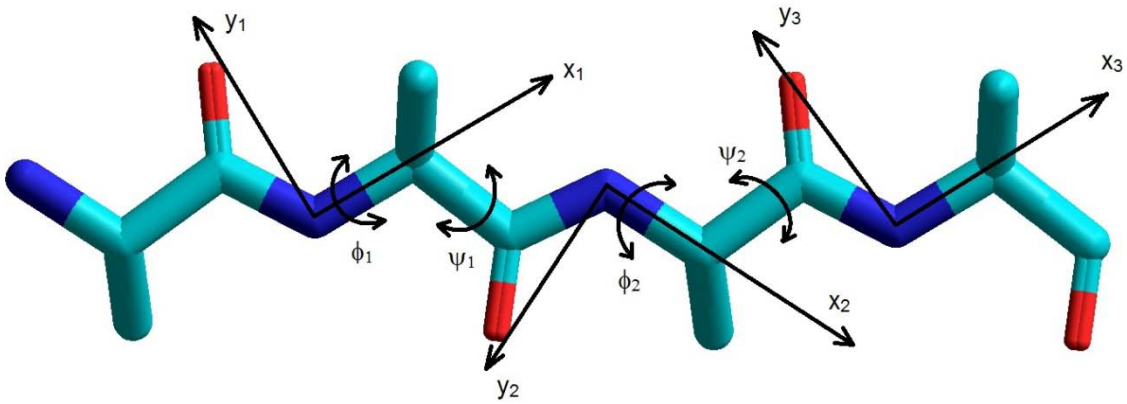


Figure 7: Local coordinate systems for expressing the transition dipole, Raman tensor, and distance vectors shown for the model system tetraalanine. The peptide is depicted as adopting a completely extended β-sheet structure with (ϕ, ψ) = (180°, 180°). For clarity, the z-axes are not shown.

To utilize **eqs. 5 - 8** for determination of the Raman tensors and transition dipole moments of the excitonic states, all local Raman tensors and dipole moments must first be transformed into a common reference system. We arbitrarily choose the coordinate system of the C-terminal residue as the reference system. Namely, Raman tensors and transition dipole moment vectors in the *j*th coordinate system have to be rotated $(n - 1) - j$ times (for an unblocked peptide) in order to have their elements expressed in the basis of the chosen C-terminal reference system (*116*). As an example, the respective transformation of the Raman tensor of the 1st residue in coordinate system $S_1(x_1, y_1, z_1)$ to the coordinate system $S_2(x_2, y_2, z_2)$ of the 2nd residue, can be written as follows (*116*):

$$\hat{\alpha}_1(S_2) = R^T(\omega)\left(R^T(\psi)\left[\left(R^T(\xi)[R^T(\phi')\hat{\alpha}_1(S_1)R(\phi')]R(\xi)\right)\right]R(\psi)\right)R(\omega) \qquad (11)$$

In particular, S_1 is first rotated by an angle $\phi' = \phi - \pi$, followed by a rotation by an angle ξ in the xy plane, to result in the y coordinate coinciding with the $C_\alpha C$ bond, which is then the rotational axis for y. Here, ξ is the angle formed by the y_1 axis and the $C_\alpha C$ bond. Next, the system is rotated by the dihedral angle ψ. The fourth step involves the rotation by an angle ω formed by the $C_\alpha C$ bond and the y_2 axis. This causes the x-axis to coincide with the NC_α bond. The angles ω and ξ have been obtained from textbooks on peptide structure, and are given as 96° and 20°, respectively (2).

The Raman tensor, $\hat{\alpha}_i$, of the *i*th excitonic state is thus calculated by first rotating all of the local amide I' Raman tensors into the coordinate system of the C-terminal residue, in combination with **eq. 5**. It can then be used to calculate the isotropic and anisotropic Raman intensity via the respective tensor invariants, $\beta'^2_{s,i}$ and $\gamma'^2_{aniso,i}$, via the following equations (*118*):

$$\beta'^2_{s,i} = \frac{1}{9}(Tr\ \hat{\alpha}'_i)^2 \tag{12}$$

$$\gamma'^2_{aniso,i} = \frac{1}{2}\left[\left(\hat{\alpha}'_{xx,i} - \hat{\alpha}'_{yy,i}\right)^2 + \left(\hat{\alpha}'_{yy,i} - \hat{\alpha}'_{zz,i}\right)^2 + \left(\hat{\alpha}'_{zz,i} - \hat{\alpha}'_{xx,i}\right)^2\right] + \frac{3}{4}\left[\left(\hat{\alpha}'_{xy,i} - \hat{\alpha}'_{yx,i}\right)^2 + \left(\hat{\alpha}'_{yz,i} - \hat{\alpha}'_{zy,i}\right)^2 + \left(\hat{\alpha}'_{zx,i} - \hat{\alpha}'_{xz,i}\right)^2\right] \tag{13}$$

IR absorption intensity is proportional to the square of the transition dipole moment associated with the respective vibrational transition, so that the IR intensity of the *i*th excitonic mode is calculated as:

$$I_i^{IR} = \tilde{\nu}_o \frac{\vec{\mu_i} \cdot \vec{\mu_i}}{\zeta} \quad (14)$$

where ζ is a constant that depends on the units of $\vec{\mu}_i$ and $\tilde{\nu}_o$, which is the first moment of the entire amide I band profile In electrostatic units, ζ is equal to 9.2 x 10^{-39} esu (*119*).

The IR and Raman amide I profiles can be calculated as a superposition of Gaussian bands representing the respective bands of the excitonic modes.

$$I_{iso}(\Omega) = \sum_{i=1}^{n-1}\left[45 \cdot \beta_{s,i}^2 \cdot f_i(\Omega, \Omega_o)\right] \quad (15)$$

$$I_{aniso}(\Omega) = \sum_{i=1}^{n-1}[7 \cdot \gamma_i^2 \cdot f_i(\Omega, \Omega_o)] \quad (16)$$

$$I_{IR}(\Omega) = \sum_{i=1}^{n-1}[I_i^{IR} \cdot f_i(\Omega, \Omega_o)] \quad (17)$$

where:

$$f_i(\Omega, \Omega_o) = \frac{1}{\sigma_i\sqrt{2\pi}} exp\left(-\frac{(\Omega-\Omega_i)^2}{2\sigma_i^2}\right) \quad (18)$$

Here, Ω_i denotes the eigenenergy of the *i*th excitonic mode in units of cm^{-1} and σ_i is full width at half maximum (FWHM) of the respective band. It should be mentioned here that the choice of Gaussian profiles for the individual excitonic transitions is one of

convenience, and in accord with common practice. The real bands, in fact, exhibit a Voigtian profile, with a Lorentzian half-width of ~ 11 cm^{-1} (*104, 120*).

If multiple representative conformations are considered, each configuration of a peptide gives rise to an individual band profile. The total amide I intensity of a peptide ensemble thus calculated as a superposition of intensity profiles of co-existing conformations, which are weighted with their respective Boltzmann terms according to the following equation:

$$I(\Omega) = \frac{\sum_{i=1}^{n} \sum_{j_1, j_2 \ldots j_n = 1}^{n_{c_1}, n_{c_2} \ldots n_{c_n}} \left(I_{i j_1 j_2 \ldots j_n}(\Omega) e^{-\left(\sum_{k=1}^{n} G_{jk}\right)/RT} \right)}{Z} \tag{19}$$

where $I_{i j_1 j_2 \ldots j_n}(\Omega)$ is the respective IR or Raman intensity profile of the configuration $\{j_1, j_2, \ldots, j_n\}$ calculated from **eqs. 15 - 17**, G_{jk} is the Gibbs energy of residue *k* in the *j*th conformation, and *i* denotes the respective excitonic state. R is the gas constant, T the absolute temperature, Z is the partition sum of the ensemble, and n_{c_k} are the number of conformations considered for residue *k*. The conformation of the C-terminal residue does not affect the amide I' band profiles. Hence, *n* excitonic states depend on the conformation of *n - 1* residues.

Vibrational circular dichroism (VCD) intensity is proportional to the rotational strength of the *i*th excitonic state which is the imaginary part of the scalar product between the transition dipole and transition magnetic moments of the respective excitonic state (*121-122*):

$$R_i = Im[\vec{\mu}_i \cdot \vec{m}_i] \quad (20)$$

The generalized equation for the rotational strength of the *i*th excitonic state for a polypeptide containing *n* – 1 residues is given as (*114*):

$$R_i = Im\left[\sum_{j=1}^{n-1} a_{ij}\vec{\mu}_j \left(\sum_{j=1}^{n-1} a_{ij}\vec{m}_j - \frac{i\pi}{2}\left(\sum_{l=1}^{n-2}\sum_{m=1}^{n-1} \tilde{\nu}_{lm}\vec{T}_{lm} \times (a_{il}\vec{\mu}_l - a_{im}\vec{\mu}_m)\right)\right)\right] \quad (21)$$

where *j*, *l*, and *m* indicate local amide I oscillators. The terms within the third set of parentheses reflect the chirality brought about by the magnetic moment at residue *l* induced by the transition dipole moment of residue *m*. As an example, the rotational strength of a tetrapeptide is given as (*114*):

$$R_i = Im\left[\sum_{j=1}^{3} a_{ij}\vec{\mu}_j \left(\sum_{j=1}^{3} a_{ij}\vec{m}_j - \frac{i\pi}{2}\left(\tilde{\nu}_{12}\vec{T}_{12} \times (a_{i1}\vec{\mu}_1 - a_{i2}\vec{\mu}_2) + \tilde{\nu}_{13}\vec{T}_{13} \times (a_{i1}\vec{\mu}_1 - a_{i3}\vec{\mu}_3) + \tilde{\nu}_{23}\vec{T}_{23} \times (a_{i2}\vec{\mu}_2 - a_{i3}\vec{\mu}_3)\right)\right)\right] \quad (22)$$

Here, $\vec{m}_j$ and $\vec{\mu}_j$ are the magnetic and electric transition dipole moments associated with the local amide I mode of the *j*th oscillator. The first term in the second set of parenthesis, namely $\tilde{\nu}_{12}\vec{T}_{12} \times (a_{i1}\vec{\mu}_1 - a_{i2}\vec{\mu}_2)$ describes the chirality brought about by the magnetic moment at residue 1 induced by the transition dipole moment of residue 2.

The terms $\tilde{\nu}_{13}\vec{T}_{13} \times (a_{i1}\vec{\mu}_1 - a_{i3}\vec{\mu}_3)$ and $\tilde{\nu}_{23}\vec{T}_{23} \times (a_{i2}\vec{\mu}_2 - a_{i3}\vec{\mu}_3)$ account for the corresponding interaction between the first and the third residue and the second and and third residue, respectively. The VCD profile of the amide I can be calculated as a superposition of Gaussians using the expression:

$$\Delta\varepsilon(\Omega) = \frac{\tilde{\nu}_o}{2.3\times10^{-39}}\sum_{i=1}^{n-1}[R_i \cdot f_i(\Omega, \Omega_o)] \quad (23)$$

where $\tilde{\nu}_o$ is the first moment of the entire amide I band profile.

1.3.5 AAKA

The amide I' band profiles in the FTIR, Polarized Raman, and VCD spectra of AAKA in D_2O were measured at acidic pD and room temperature, and are shown in **Figure 8**. Acidic conditions were chosen to allow for maximal spectral resolution and to avoid the instabilities (e.g. precipitation) sometimes encountered at neutral pD. A comparison of the band profiles reveals a significant split between the peak positions of the isotropic Raman and IR band profile. The anisotropic band has its peak between those of the IR and isotropic Raman band, but it is slightly closer to the IR band position. This is typical for a peptide with a predominant PPII fraction (*114*, *123-124*). The respective VCD spectrum displays an asymmetric negative couplet with a large negative signal at the position of the lowest wavenumber amide I' sub-band, which is predominantly assignable to the C-terminal peptide group (*43*). This negative signal reflects, to some extent, an intrinsic magnetic moment of the C-terminal amide I' oscillator, which gives rise to a negative Cotton effect even in the spectrum of a dipeptide. This is illustrated by the amide I' VCD spectrum of L-lysyl-L-alanine (KA) in **Figure 9**, which was also measured at acidic pD. The $^{3}J_{C\alpha HNH}$ coupling constants and the respective chemical shifts of residues 2 – 4 of AAKA are listed in **Table 1**.

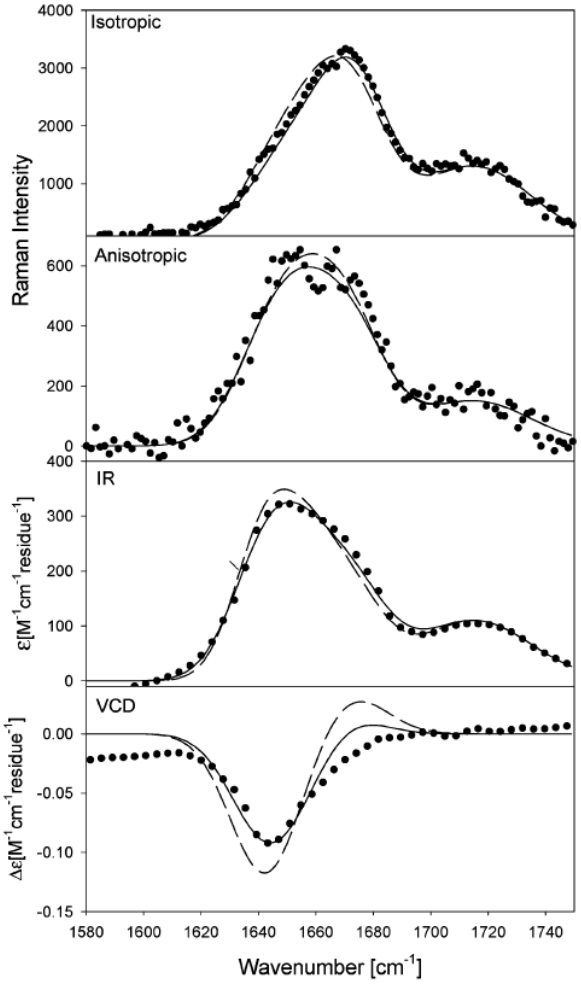


Figure 8: Amide I' band profile (dotted) of the isotropic Raman, anisotropic Raman, IR, and VCD spectra of AAKA measured at pD = 1.5. The solid line results from a simulation based on a three-state (per residue) model, encompassing PPII, β, and helix-like conformations derived from the coil library of Avbelj and Baldwin (*125*) (PPII and β) and the MD simulations of Duan et al. (*91*) (helical). The dashed line was calculated with a two-state model (PPII and β). The band of the C-terminal carbonyl stretching mode has been modeled for the sake of completeness by fitting a Gaussian profile to the experimental data with the wavenumber and half-width as free parameters.

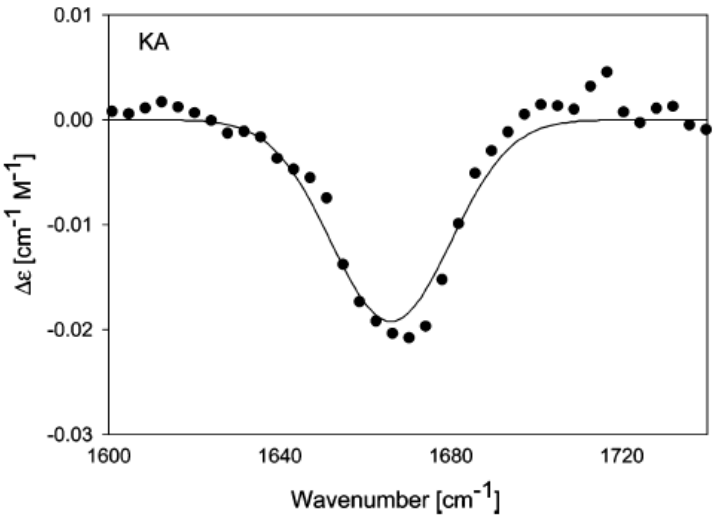


Figure 9: VCD spectrum of KA (pD = 1.5). The solid line results from a fit described in the text.

Residue No.	$^3J_{C\alpha HNH}$ (Hz)
2 (Ala)	5.18
3 (Lys)	6.46
4 (Ala)	6.72

Table 1: $^3J_{C\alpha HNH}$ NMR coupling constants of the 2nd, 3rd, and 4th amino acid residues of AAKA.

First, the VCD spectrum of KA was analyzed by using:

$$\Delta\varepsilon(\Omega) = \frac{\Omega_o Im(\vec{\mu}\cdot\vec{m})}{2.3\times10^{-39} M\cdot esu^2\cdot cm^2} f(\Omega,\Omega_o) \quad (24)$$

where Ω_o is the peak wavenumber of the amide I' and $\vec{\mu}$ the electronic and $\vec{m}$ the magnetic transition dipole moment. $f(\Omega,\Omega_o)$ denotes the Gaussian band profile. The dipole moments reported for cationic KA were utilized (*108*). A magnetic dipole moment of 2.3×10^{-23} esu cm and an orientational angle of 10° for $\vec{m}$ with respect to the NC axis of the C-terminal residue (*69*). The wavenumber and half-width were obtained from the respective IR band profile (data not shown). Thus, the fit of the VCD signal visualized by a solid line in **Figure 9** was obtained.

The excitonic coupling model used to describe the delocalization of the amide I mode is described in **Section 1.3.4**. In particular, it describes the amide I band profiles of IR, isotropic Raman, anisotropic Raman, and VCD in terms of nearest neighbor and non-nearest neighbor excitonic coupling between the excited states of the local oscillators (*111*) and the dihedral angles ϕ and ψ of the residues between the interacting amide I modes. The first central residue (starting at the N-terminal) is counted as I and the subsequent one as 2. It should be noted that the two terminal residues are not probed by the amide I' band profile.

The total intensity of a peptide ensemble is calculated according to eq. 19 (Section 1.3.4). In our analysis, three representative conformations, *j*, were considered, namely PPII (j = 1), β (j = 2), and helical (j = 3). The Gibbs energy of PPII was set to zero. In a first step, we based our simulation on $(\phi, \psi)_1$ = (-65°, 150°) and $(\phi, \psi)_2$ = (-125°, 115°) for alanine, which represent their respective distributions in the coil library of Avbelj and Baldwin for PPII and β, respectively (*125*), and correspond to the $^3J_{C\alpha HNH}$ coupling constants reported by Shi et al. (*81*). For lysine, these authors reported a slightly different $^3J_{C\alpha HNH}$ value for PPII, which corresponds to ϕ = -68°. Coil libraries as well as molecular dynamics simulations for alanine dipeptides generally exhibit helical conformations (*91*). These coordinates are closer to a 3_{10} than to a canonical α-helical conformation (*126*). In the following, we refer to this structural model as the "coil library model."

The $^3J_{C\alpha HNH}$ constants of these conformations were utilized, and our experimentally determined coupling constants to determine the mole fraction of three considered conformations as follows:

$$\chi_{1,k} = \frac{(J_k - J_{\beta,k}) - \chi_{3,k}(J_\alpha - J_{\beta,k})}{J_{p,k} - J_{\beta,k}} \quad (25a)$$

$$\chi_{2,k} = 1 - \chi_{1,k} - \chi_{3,k} \quad (25b)$$

Where $\chi_{j,k}$ is the mole fraction of the *k*th residue in the *j*th conformation, J_k is the measured coupling constant of the *k*th residue, and $J_{p,k}$ and $J_{\beta,k}$ are the coil library

coupling constants of the *k*th residue for the PPII and β as reported by Shi et al. (*81*). $J_\alpha = 4.1$ Hz is the representative coupling constant of the helical conformation mentioned above. The mole fraction, $\chi_{3,k}$, of this conformation has been used as an adjustable parameter in the simulation. The mole fractions were used to calculate the relative Gibbs energies of the residues:

$$G_{j_k} = RTln\left(\frac{\chi_{j,k}}{\chi_{1,k}}\right) \tag{26}$$

which were used in **eq. 21** to calculate the intensity profiles.

The amide I' band profiles of AAKA were simulated assuming solely a coexistence of PPII and β for each residue, i.e. $\chi_{3,k} = 0$. Cooperativity was not considered, in accordance with the isolated pair hypothesis (*33*) and the temperature dependence of the UV-CD spectra reported below. The nearest neighbor coupling constants (5.0 for PPII and 2.5 for β) were taken from the *ab initio* study on a glycine dipeptide by Torii and Tasumi (*109*). For PPII, this is in the interval of values observed earlier from isotropic Raman and femtosecond IR data of tripeptides (*43*, *79*). The corresponding β-strand coupling constant is slightly lower than that obtained for trivaline (*79*). The coupling between the N- and C-terminal amide I' modes was calculated using the transition dipole coupling (TDC) formalism (see **Section 1.3.4**) (*103*). The program MULTIFIT (*127*) was used to self-consistently decompose the IR and Raman band

profiles in **Figure 8** into three bands as described in an earlier study (*69*) and used the respective peak positions as a basis for guessing the wavenumber positions for the unperturbed amide I' modes. The Gaussian bandwidths for the individual bands were also obtained from this analysis. Moreover, we used the reported electronic transition dipole strengths of AA and KA (*108*). This yielded band profiles which reproduced the IR and Raman data reasonably well (dashed line in **Figure 8**), though some small deviations between simulated and experimental IR band profiles are notable. The VCD simulation, however, totally underestimates the large negative signal, which is much more pronounced than in the spectrum of KA. Additionally, it overestimates the positive signal at higher wavenumbers. A satisfactory simulation required a much larger magnetic transition dipole moment for the C-terminal amide I' mode, i.e. $1.3 \cdot 10^{-22}$ esu cm (dashed line in the lower panel of **Figure 8**). In a third step, we checked for the possibility that a small fraction of helical conformations are present in the sample. Indeed, we obtained a better agreement for the IR band profile, the highly asymmetric shape of which was nicely reproduced with a mole fraction of 0.3 for both residues (solid line in **Figure 8**). Larger values for $\chi_{3,k}$ yield substantial deviations from the experimental data in that too much intensity is accumulated at the highest wavenumber position, as expected for a predominantly helical band profile (*128*). We also checked the possibility that only one of the central residues adopts helical conformations, but that did not improve the profiles compared with the two-state (PPII/β) simulation. Thus, our final result yields PPII fractions of $\chi_{1,1} = 0.63$ for alanine and $\chi_{1,2} = 0.36$ for lysine, the β-fractions are $\chi_{2,1} = 0.07$ for alanine and $\chi_{2,2} = 0.34$ for lysine. Hence, alanine has a clear

PPII propensity, but the values for lysine suggest a statistical coil behavior, in contrast to what has been observed for polylysine peptides (*41, 50, 95*).

Additionally, we considered slightly different representative structures for our three basic conformations. The band profiles of a simulation based on the PPII, β and helix conformations in **Figure 10** emerged from the MD simulation of Garcia, namely ($(\phi,\psi)_1 = (-60^o, 120^o)$, $(\phi,\psi)_2 = (-150^o, 120^o)$, and $(\phi,\psi)_3 = (-60^o, 60^o)$), respectively (*129*). The respective $^3J_{C\alpha HNH}$ constants were adjusted using the modified Karplus equation of Vuister and Bax (*130*). The representative PPII conformation resembles more the structure adopted by the first residue in a type II β-turn. The best reproduction of the experimental profiles was obtained with $\chi_{3,k} = 0.20$. The agreement between simulated and experimental profiles is not bad, but small systematic differences are discernable for IR, isotropic and anisotropic Raman. The respective PPII fractions are $\chi_{1,1} = 0.60$ for alanine and $\chi_{1,2} = 0.45$ for lysine, while the β-fractions are $\chi_{2,1} = 0.19$ and $\chi_{2,2} = 0.35$ for alanine and lysine, respectively. Thus, alanine has still a clear PPII propensity. The respective value for lysine is slightly larger than that obtained from the 'coil library model.'

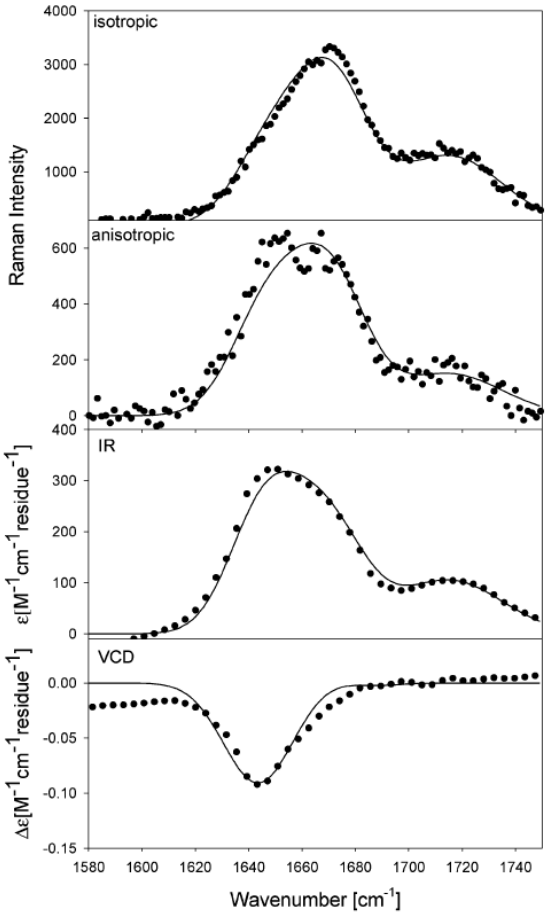


Figure 10: Amide I' band profile (dotted) of the isotropic Raman, anisotropic Raman, IR, and VCD spectrum of AAKA measured at acidic pD 1. The solid line results from a simulation based on the three-state (per residue) model, encompassing PPII, β, and helix-like conformations derived from MD simulations of Gnanakaran and Garcia (*90*). The band of the C-terminal carbonyl stretching modes has been modeled for the sake of completeness by fitting a Gaussian profile to the experimental data with the wavenumber and half-width as free parameters.

The temperature-dependent UV-CD spectra of AAKA plotted in units of M^{-1} cm^{-1} $residue^{-1}$ is depicted in **Figure 11**. In contrast to what has been obtained for the heavily debated XAO peptide, the spectra exhibit the characteristic minimum (195 nm) and maximum (215 nm), which is diagnostic of a substantial PPII contribution (*131*). Compared with tetraalanine, however, both the maximum and minimum are reduced by a

factor of 2 and 3, respectively (*69*). This is qualitatively consistent with the much larger PPII fraction reported earlier for tetraalanine.

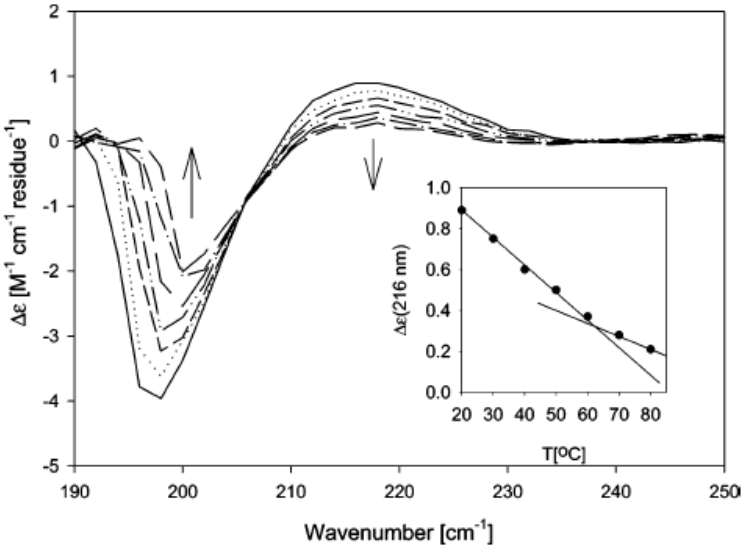


Figure 11: Temperature-dependent UV-CD spectra of AAKA at pD = 1, measured from 20 to 80°C at 10° intervals.

It should be emphasized that the two sets of representative structures considered above are most likely part of a distribution of conformations assignable to different basins in the Ramachandran plot. More realistic models require the consideration of distributions of conformations centered about the basins of the Ramachandran map, whose maxima have been considered here as representative conformations. More recent studies in our laboratory rely on an extension of the currently employed model which utilizes two-dimensional Gaussian distributions of conformations from different regions of Ramachandran space, instead of representative sets of dihedral angles (*120, 132*).

Different coil libraries and results from MD simulations are far away from agreeing with each other with respect to the corresponding mole fractions and distribution widths of the sampled conformations. A classical random coil distribution of e.g. alanine, as it emerged from the work of Ramachandran et al. (*88*) and Brant and Flory (*89*) can definitely be ruled out from our experimental data. We also considered the possibility of various turn conformations to contribute to the AAKA ensemble. The agreement with the experimental profiles was poor at best even if only small fractions of e.g. β-turns were considered.

1.3.6 Tetra- (A_4) and Tri-alanine (A_3)

In the past, Schweitzer-Stenner and associates have used a somewhat less advanced model to determine the average structure of tri- and tetrapeptides, in that they identified the dihedral angles which reproduce the intensity ratios of the individual amide I' sub-bands derived from spectral decomposition (*44, 69, 79*). The result was subsequently interpreted in terms of a two state model, which considered solely PPII and an extended β-strand conformation, which is different from the more parallel β-sheet like dihedral angles used in the present study. Therefore, we felt a necessity to re-analyze both tetra- and trialanine by means of our statistical model. The earlier reported band profiles of (cationic) tetraalanine and (zwitterionic) trialanine are depicted in **Figure 12** and **Figure 13**, respectively. Additionally, we measured the $^{3}J_{C\alpha HNH}$ coupling constants for tetraalanine; these values are listed in **Table 2**, while the coupling constant for the central alanine residue in AAA (i.e 5.6 Hz) was taken from a study by Graf et al. (*133*). Again,

we used these constants to restrict the simulations of the band profiles by means of **eqs. 21 – 23**. In a first step, we considered the representative conformations of the 'coil library model' and obtained insufficient agreement with the experimental data, irrespective of our choice of the helical fraction. That particularly concerned the IR and anisotropic Raman band profiles, for which the relatively large intensity of the lowest wavenumber band could not be reproduced. For tetraalanine, however, we could obtain quite good agreement with the experimental profiles by using the representative conformation obtained in our earlier study (with ϕ angles of -70^o and -80°), but this is not consistent with the observed $^3J_{C\alpha HNH}$ coupling constants listed in **Table 2**. Inclusion of any of the different types of turns in the simulation led to poorer agreement with the experimental data. Subsequently, we decided to modify the coordinates of the representative PPII conformations and finally obtained the simulation in **Figure 12** for tetraalanine, by assuming that the difference between the $^3J_{C\alpha HNH}$ coupling constants reflects different PPII coordinates rather than different conformational mixtures. The agreement with the experimental data can be considered as satisfactory. The respective PPII conformations are $(\phi,\psi)_{11}$ = (-65^o, 150^o) and $(\phi,\psi)_{12}$ = (-70^o, 140^o). Helical contributions had to be considered, and we obtained different fractions for the two residues, namely $\chi_{3,1}$ = 0.20 for the first and $\chi_{3,2}$ = 0.10 for the second of the non-terminal alanines (counted from the N-terminal). The PPII mole fractions are $\chi_{1,1}$ = 0.66 and $\chi_{1,2}$ = 0.73, while the β-fractions are $\chi_{2,1}$ = 0.14 and $\chi_{2,2}$ = 0.17. With respect to PPII, this is less than what was estimated in our earlier study, but the result still corroborates the notion that alanine has a high PPII propensity, contrary to recent claims (*87*). Additionally, we

simulated the profiles with the coordinates representing the distributions of Gnanakaran and Garcia (*90*), but this did not reproduce any of the experimental profiles.

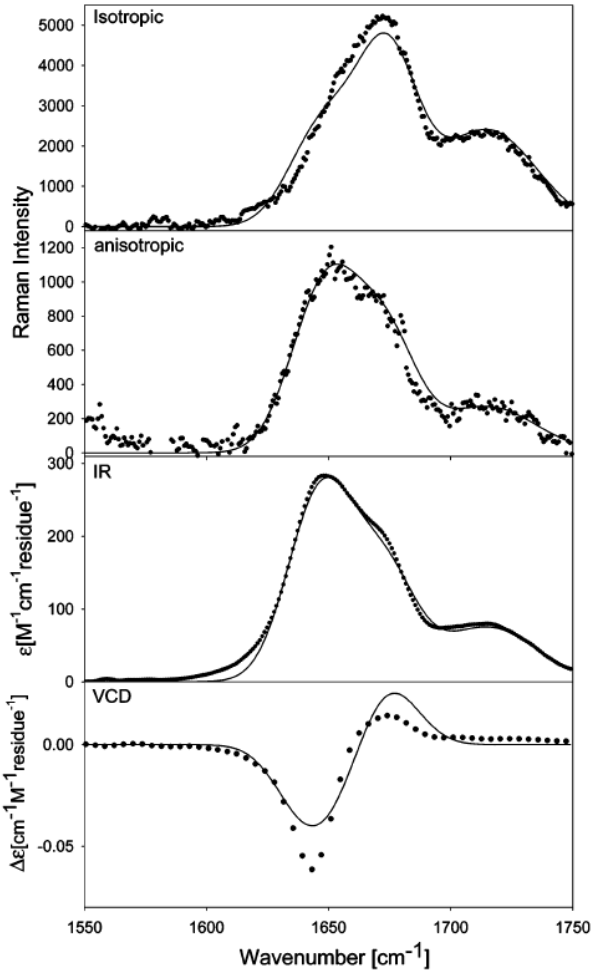


Figure 12: Amide I' band profile of the isotropic Raman, anisotropic Raman, IR, and VCD spectrum of tetra-alanine (AAAA), measured at acidic pD = 1 (taken from ref(*69*)). The solid line results from a simulation based on a three-state (per residue) model, encompassing PPII, β, and helix-like conformations as described in the text. The band of the C-terminal carbonyl stretching modes has been modeled for the sake of completeness by fitting a Gaussian profile to the experimental data with the wavenumber position and half-width as free parameters.

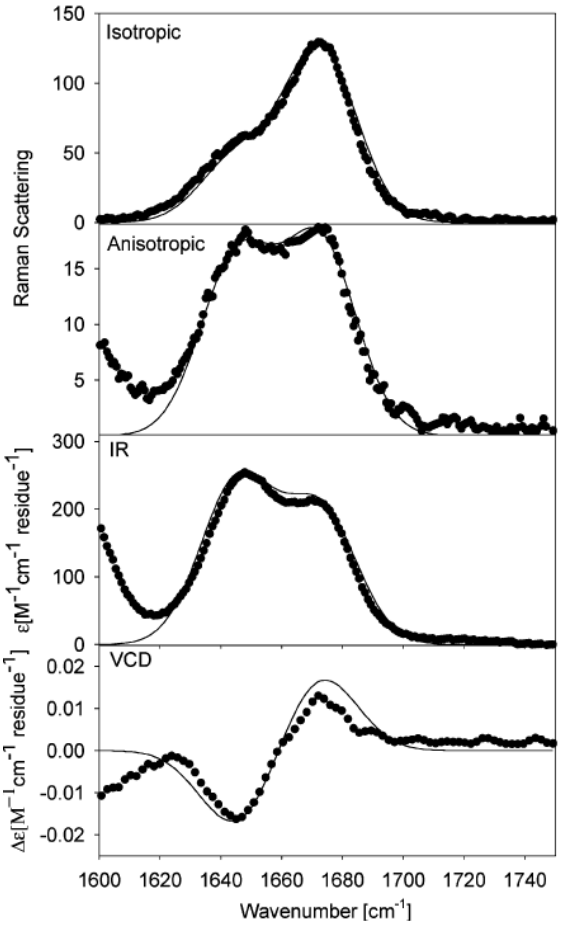


Figure 13: Amide I' band profile of the isotropic Raman, anisotropic Raman, IR, and VCD spectrum of AAA measured at neutral pD = 6 [taken from ref(*79*)]. The solid line results from a simulation based on a three-state (per residue) model, encompassing PPII, β, and helix-like conformations as described in the text.

Residue No.	$^3J_{C\alpha HNH}$ (Hz)
2	5.50
3	6.03
4	6.92

Table 2: $^3J_{C\alpha HNH}$ NMR coupling constants of the 2nd, 3rd, and 4th amino acid residues of tetraalanine (AAAA).

For tetraalanine, NOESY cross-peaks are observed in the H^N-H^α, H^N-H^{Me} and H^α-H^{Me} regions, the first one being the most relevant for conformational analysis. In fact, as far as the internal residues are concerned, the H^N-H^α intra-residue cross-peaks are much less intense than the inter-residue ones, suggesting that the latter are characterized by (φ, ψ) angles in the PPII and/or in the β sheet Ramachandran region, for which $d_{H^{\alpha i}H^{N(i+1)}}$ distances are shorter than $d_{H^{\alpha i}H^{Ni}}$ (*91*). Moreover, the absence of measurable NOE's in the amide-amide region excludes the presence of appreciable amounts of α-helix conformers. These results corroborate in general terms the analysis of the vibrational spectra.

With respect to trialanine, Eker et al. reported a 50:50 mixture of PPII and β (*79*), but this cannot be reconciled with the $^3J_{C\alpha HNH}$ coupling constant of 5.6 Hz. The rather satisfactory simulation depicted in **Figure 13** was obtained with a slightly modified PPII conformation, namely $(\phi,\psi)_{11}$= (-60°, 150°), while the other coordinates of the 'coil library model' remained unchanged. This PPII conformation is very close to the values reported in the early study of Woutersen and Hamm (*43*). The helical fraction of χ_3 = 0.20 is a necessity. The respective PPII and β fractions are χ_1 = 0.60 and χ_2 = 0.20, respectively. Again, we also tried to simulate the profiles with the coordinates

representing the distributions of Gnanakaran and Garcia (*90*), but the result was less satisfactory.

It is noteworthy that the C-terminal $^{3}J_{C\alpha HNH}$ coupling constants of AAKA and AAAA (6.7 and 6.9 Hz) are slightly higher than the corresponding value obtained for AA (6.5 Hz) (*134*). All these values are higher than those of the central residues of the investigated peptides. We interpret this as indicating that terminal residues are more likely to sample β-strand like conformations, which yield an increase of $^{3}J_{C\alpha HNH}$. If one invokes the two-state coupling model of Shi et al. (*81*) and the $^{3}J_{C\alpha HNH}$ constants of Avbelj and Baldwin (*125*), all these values are still indicative of a high percentage of polyproline II, as argued by Dragomir et al. (*135*).

Taken together, all of these analyses demonstrate that the combination of our vibrational spectroscopy approach with NMR data (i.e. $^{3}J_{C\alpha HNH}$) is very powerful in identifying the conformational manifold of amino acid residues in peptides. This strategy takes advantage of the site specific information provided by NMR, along with the high sensitivity of the amide I band profiles concerning even modest variations of the secondary structure. In this context we would like to emphasize that this stems in part from the fact that second nearest neighbor vibrational coupling (i, i + 2) is not insignificant, particularly for the helical and the β-conformation.

1.3.7 Discussion and Comparison with Literature

As mentioned above, the UV-CD spectra of various unfolded alanine based peptides exhibit the characteristic feature indicative of a substantial fraction of PPII (*41*). The much discussed XAO peptide is somewhat of an exception from the rule in that its maximum at 216 nm is substantially reduced, which suggests the presence of other conformers (*128*). This and other experimental data for this particular peptide led Makowska et al. (*87*) and Vila et al. (*136*) to the conclusion that alanine does not have a substantial PPII propensity. The results of the present study strongly corroborate the opposite notion. The PPII mole fraction of alanine was found to vary between 0.60 and 0.75 for the peptides investigated. These values are somewhat lower than those reported by Shi et al for XAO (*45*), but still large enough to consider a PPII preference of alanine. Our analysis also corroborates the notion that the PPII propensity depends on the number of alanine residues, owing to a higher PPII propensity of alanine in AAAA compared with that of AAA (and AA (*134*)), in line with theoretical predictions of Garcia (*129*) and Raman optical activity (ROA) data reported by McColl et al. (*68*).

The sampling of helical conformations by alanine in short peptides has been a matter of debate. Thus far, it has been overlooked in the studies of Kallenbach, Schweitzer-Stenner, and their respective associates. MD simulations, however, generally reveal a sampling of the helical basin of the Ramachandran plot (mostly right handed), but the respective fraction does depend strongly on the choice of the force field (*85*). Woutersen et al. investigated the spectral broadening of amide I of trialanine and arrived at the conclusion that their sample of AAA contained 20% α_R and 80% PPII (*137*). Our

own analysis for this peptide yields the same helical fraction, but a lower PPII propensity. The mixture reported by Woutersen et al. would be inconsistent with our amide I' profiles and UV-CD data. Mu and Stock performed MD simulations with a GROMOS 43A1 force field and obtained χ_1 = 0.41 (PPII), χ_2 = 0.41 (β), and χ_3 = 0.16 (helical), for the central residue (*138*). This is close to our earlier reported 50:50 mixing of PPII and β, but, as argued above, inconsistent with the $^3J_{C\alpha HNH}$ coupling constant. Duan et al. (*91*) used the force field of Cornell et al. (*139*) to carry out a MD simulation for an alanine dipeptide and a blocked alanine tetrapeptide in explicit water. For the dipeptide they obtained comparable populations of PPII and the (right handed) helical fraction with a minor population of β. For the tetrapeptide, these three basins are nearly equally populated, so that their plot looks like the classical random coil distribution, which Ramachandran et al. (*88*) and Brant and Flory (*89*) obtained for the alanine dipeptide. All these results are inconsistent with our experimental data. The discrepancy is even more pronounced for simulations which Hu et al. (*140*) and Zagrovic et al. (*86*) performed with different Amber, Charm and Gromos force fields for an alanine dipeptide and XAO, respectively. Kentsis et al. used a CHARM27 force field to calculate the propensities of all natural amino acids in GGXGG peptides (*141*). The so estimated alanine propensity for PPII is quite low (30%). All these simulations yield an overestimation of the helical fraction and an underestimation of PPII. Tran et al. used their own potential function and Monte Carlo simulations to obtain a mixture of nearly 50% PPII, 30% β (i.e. with angles of a parallel β-sheet) and approximately 18% helical conformations (*36*). This is closer to our results, but their PPII fraction contains a substantial amount of a conformation termed P_{hyp}, which exhibits lower ψ and ϕ values

than the canonical PPII. Such conformations would exhibit a very low level of excitonic coupling for amide I (*109*) and are therefore not likely to contribute significantly to the observed amide I band profiles. Gnanakaran and Garcia used a modified AMBER force field (A94/MOD) to simulate the sampling of polyalanine peptides of different length (*90*). For the alanine dipeptide they observed a mixture of $\chi_1 = 0.59$ (PPII), $\chi_2 = 0.14$ (β), and $\chi_3 = 0.20$ (helical); for AAA they observed $\chi_1 = 0.80$ (PPII), $\chi_2 \approx 0.08$ (β), and $\chi_3 \approx 0.08$ (helical). Their result for the alanine dipeptide is very close to what we observed for the central residue of AAA, but even their result for AAAA is much closer to ours than all the other simulations discussed above. Finally, it deserves to be mentioned that our results are in very good agreement with the PPII, β and helix population of alanine in the (reduced) coil library of Serrano (*61*), but are inconsistent with those in the library of Fiebig et al. (*84*).

We would also like to note the remarkable similarity between the dihedral angles which were used in the current study to simulate the vibrational spectra of tetraalanine and those reported recently by Pizzanelli et al. (*142*). The authors performed a DFT calculation for tetraalanine in implicit and explicit solvent (14 water molecules). The most stable conformation obtained was clearly PPII like; the respective coordinates are $(\phi,\psi)_{11} = (-57.6^o, 145.0^o)$ and $(\phi,\psi)_{12} = (-68.8^o, 138.4^o)$ for the central residues. This is not only close to the values obtained in the present study, but also reproduces the small difference between the two PPII conformations of the central residues, which we inferred from our data. The predominance of PPII was then confirmed by ^{1}H NMR spectroscopy by measuring the dipolar couplings of tetraalanine in an oriented lyotropic liquid crystal.

The DFT calculations revealed a conformation with both residues in a helical conformation as the second one in the conformational hierarchy. The close resemblance between these values and those which we report affirm the merit of our current model, as well as support the PPII propensity of alanine.

The low PPII propensity observed for the lysine residue in AAKA is somewhat surprising since earlier studies on charged polylysine peptides of different lengths were all in agreement in suggesting a high PPII propensity (*41, 50, 95*). That this property might depend on the choice of the nearest neighbors has already been indicated by Eker et al., who showed that K in AKA adopts β-strand like conformations at neutral pD, whereas PPII is preferred at acidic conditions (*48*). More recently, Measey and Schweitzer-Stenner (*143*) showed that the average PPII fraction of the octa-peptide, $(AAKA)_2$, is 60%, which is lower than expected based on results of previous spectroscopic studies of short alanine-based peptides (*45, 69*). We investigated AAKA at acidic conditions and showed that K samples PPII, β and helical conformations. This resembles the statistical coil concept of Zimmerman and Scheraga (*34-35*). The reason for this behavior might be that lysine perturbs the order of the hydration shell which has been suggested to stabilize PPII in polyalanine peptides (*144*). Generally, our result shows that nearest neighbor effects are relevant for individual propensities of amino acid residues and that a PPII preferring neighbor does not necessarily support the PPII propensity of a given residue.

Taken together our study provides compelling evidence for a significant PPII propensity of alanine in AAKA, AAAA, and AAA, while it reveals a context dependent

behavior of lysine, which deserves further investigations. Furthermore, we showed that alanine samples, to some extent, also (right handed) helical conformations, in line with predictions of Gnanakaran and Garcia (*90*), Mu and Stock (*138*) and Woutersen et al (*137*). We would like to emphasize, however, that our results do not argue in favor of a somewhat simplistic notion that *the* structure of unfolded alanine peptides is PPII, as suggested in a recent paper of Asher et al. (*145*). If the individual PPII mole fraction of alanine is as high as 0.73 (as obtained for AAAA), the molar fraction of XAO with all alanines in PPII is 0.11. The corresponding fractions for six, five and four residues are 0.08, 0.095 and 0.12, respectively. Thus, nearly 40% of the peptides exhibit PPII sequences of 6 and more. This is not insignificant. One has to take into account, however, that conformational fluctuations are significant on the PPII potential surface, i.e. a PPII segment is still a dynamic entity. Nevertheless, it is justified to state that the classical random coil or even the statistical coil concept does not apply to alanine, and evidence exists that it does not apply to other types of residues either (*61-62*).

1.3.8 The Hepta-Alanine-Containing XAO Peptide

The amide I' band profiles of the IR, isotropic Raman, anisotropic Raman, and VCD spectra of XAO in D_2O (pD = 2.2) are shown in **Figure 14.** We observed a clear non-coincidence between the isotropic and IR-band profiles, which suggests that the peptide predominantly samples the upper left quadrant of the Ramachandran plot. The anisotropic band is very broad and does not display the clearly discernable peak position in between the wavenumber positions of the IR and isotropic Raman bands, which has

been shown to be diagnostic of PPII (*114, 123, 128*). The VCD spectrum displays the usual negative couplet of so called unfolded peptides and proteins, but it is less pronounced than those observed for e.g. AAA, AAAA, AAKA and $(AAKA)_2$ (*43, 48, 69, 143, 146*). The rotational strength per residue is substantially smaller than that observed for the amyloid peptide fragment $A\beta_{1\text{-}28}$, which has been shown to contain a substantial PPII content (*108, 123, 128*), but similar to that of salmon calcitonin, which more closely resembles a statistical coil (*128*). All of these observations suggest that XAO has much less PPII content than suggested by Kallenbach and associates (*45*). This notion is corroborated by the UV-CD spectra of XAO at 20°C, as shown in **Figure 15**. As already argued by Vila et al. (*136*), the spectra show the minimum at 195 nm, which is associated with PPII, but do not display the maximum at 216 nm, which is diagnostic of this structure. With respect to the minimum, the corresponding $\Delta\varepsilon$ value (per residue) is much less pronounced than that observed for AAAA (triangles in **Figure 2**) (See **Section 1.3.6**), and slightly smaller than that obtained for $A\beta_{(1\text{-}28)}$ (*123, 146*).

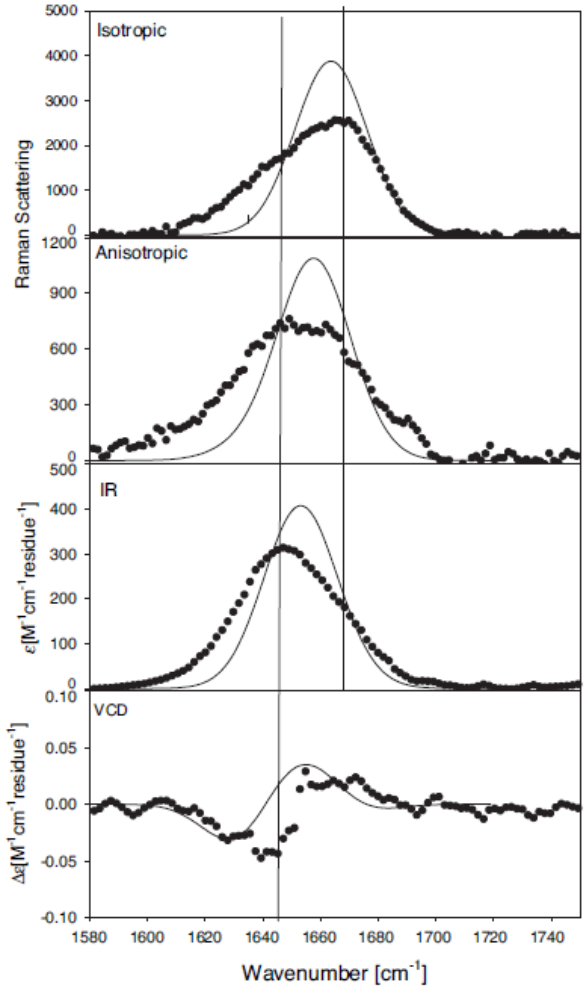


Figure 14: Experimental isotropic and anisotropic Raman, FTIR, and VCD spectra of the amide I' band region of XAO, pD = 2.2 (dotted), and band profiles simulated by using the two-state per residue model (PPII/β), in line with the results of Shi et al. (*45*).

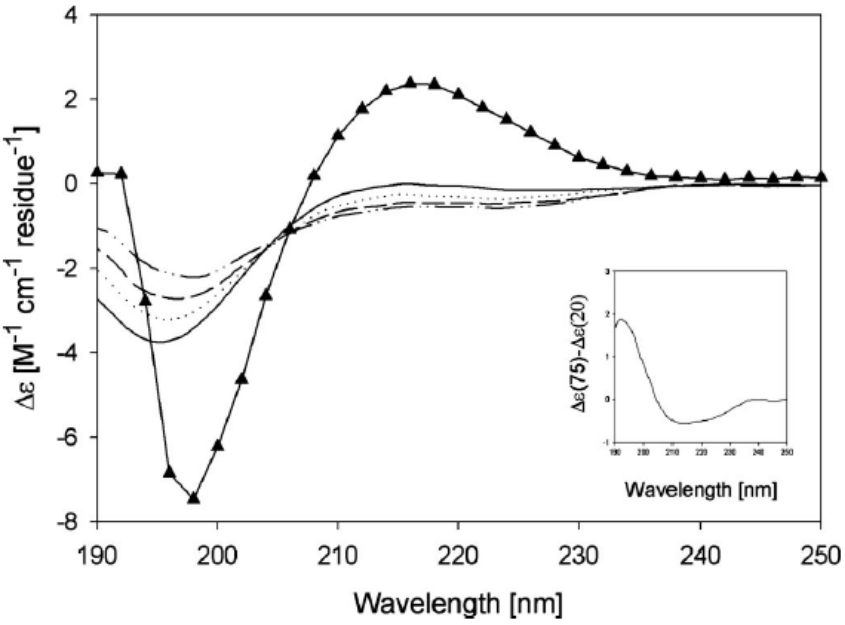


Figure 15: Temperature-dependent UV-CD spectra of XAO (pD = 2.2) at 20°C (solid), 40°C (dotted), 60°C (dash), and 80°C (dash dot dot dash). For comparison, the spectrum of tetraalanine (pD = 1) at 20°C is also shown (triangles). The inset shows the difference spectrum of XAO obtained by subtracting the spectrum at 20°C from that measured at 80°C.

We performed a detailed analysis of the amide I' band profiles based on a statistical model which assumes a blend of different conformations per residue. This approach is a substantial refinement of more crude models which were used previously to analyze the profiles of longer peptides (*95*, *123*), and an extension of a recently applied simple statistical model (*146*). The model is based on the following representative conformations: (a) PPII with $(\phi,\psi) = (-68^o, 150°)$, (b) β-strand with $(\phi,\psi) = (-119^o, 113^o)$, (c) various β-turn conformations, i.e. type I' $((\phi,\psi)_i = (60^o, 30^o), (\phi,\psi)_{i+1} = (90^o, 0^o))$, type II $((\phi,\psi)_i = (-60^o, 100^o), (\phi,\psi)_{i+1} = (80^o, -10^o))$, type III $((\phi,\psi)_i = (-60^o, -30^o), (\phi,\psi)_{i+1} = (-60^o, -30^o))$, type III' $((\phi,\psi)_i = (60^o, 30^o), (\phi,\psi)_{i+1} = (60^o, 30^o))$, type V $((\phi,\psi)_i = (80^o, -80^o), (\phi,\psi)_{i+1} = (-80^o, 80^o))$ and a conformation representing the bridge region between

the upper left quadrant and the helical region of the Ramachandran plot ($(\phi,\psi)_i$ = (-20^o,20^o), $(\phi,\psi)_{i+1}$ = (20^o,20^o)), which emerged from the simulations of Makowska et al. as part of the heterogeneous manifold of so called type IV turns (*87*). The coordinates for PPII, β-strand, and 3_{10} (type III) correspond to maxima of distributions inferred from coil libraries (*125*). The turn structures represent the structural manifolds sampled by MD simulations of Makowska et al. (*87*).

The total intensity of an amide I' band profile at a given wavenumber is written in terms of the above considered statistical ensemble, and calculated according to eq. 19 (Section 1.3.4). In our analysis we considered PPII (*j=1*) and β-strand (*j=2*) for all residues. For residues 3, 4, 7, and 8 (all alanines), we additionally considered type IV (*j*=3), type III (*j*=4) (type III' for residue 3), type V (*j*=5), type II (*j*=6), and type I' (*j*=7). For residues 2 (X), 5, 6, and 9 (all Ala) we allowed type III' and type IV turns as additional conformations. This blend represents, more or less, different clusters which emerged from the MD simulations of Makowska et al. (*87*). Their results suggest a limited sampling of conformations involving intrapeptide hydrogen bonding, hence this possibility was disregarded for the sake of simplicity. These authors did not obtain substantial fractions of right handed helix-like conformations for alanine residues, but some coil libraries and many MD simulations indicate that alanine significantly samples helix like conformations (*136*). We used the $^3J_{C\alpha HNH}$ coupling constants reported by Shi et al. (*45*) to restrict the choice of possible conformational mixtures by calculating:

$$^{3}J_k = \sum_{j_k}^{n_{c_k}} \chi_{j_k} \cdot J_{j_k} \quad (27)$$

where $^{3}J_k$ is the experimental coupling constant obtained for the *k-th* residue, χ_{j_k} is the mole fraction of the *j-th* conformation of the *k-th* residue and $^{3}J_{j_k}$ is the respective coupling constant.

For the simulation of the band profiles, we used the nearest neighbor coupling $\Delta_{i,i+1}$ constants which Torii and Tasumi obtained from *ab initio* calculations on a glycine dipeptide (*109*). These coupling constants have been shown to agree well with experimentally derived values for short peptides (*43, 111*). The non-nearest neighbor coupling constants $\delta_{i,j}$ were calculated by using the transition-dipole-coupling (TDC) formalism. We used the intrinsic amide I' wavenumber position of the central alanine residue of tetraalanine in water as a reference for the PPII conformation of alanine (*69*). For the two charged residues, X and O, we estimated the respective intrinsic wavenumbers by utilizing the wavenumber difference between K and A obtained from the anionic dipeptides AA and KA (*108*). The intrinsic wavenumbers for the other conformations considered in this study were obtained by considering their conformational dependence obtained by DFT studies on alanine dipeptides (*112*). The intrinsic dipole strengths of the individual amide I' modes were obtained from the dipeptide study of Measey et al. (*108*). Here, we used K as a model for X and O. Thus, we calculated the IR absorption and the vibrational circular dichroism in absolute units. Finally, we utilized the earlier obtained (relative) Raman tensors for AA (for A) and KA (for X, O) to

calculate the ratio of anisotropic and isotropic Raman scattering. Parameters used in the simulation can be found in **Table 3**.

Parameter	Value
μ_{Ala}	2.7×10^{-19} esu cm^{-1}
μ_{Lys}	3.0×10^{-19} esu cm^{-1}
$\tilde{\nu}_{Ala}$	1658 cm^{-1}
$\tilde{\nu}_{Lys}$	1664 cm^{-1}
Δ_{α}	12.0 cm^{-1}
$\Delta_{3_{10}}$	3.1 cm^{-1}
Δ_{PPII}	10.0 cm^{-1}
Δ_{β_s}	2.6 cm^{-1}
Δ_{TypeI}	7.0 cm^{-1}, 2.0 cm^{-1}
Δ_{TypeII}	-3.0 cm^{-1}, -2.0 cm^{-1}
$\Delta_{TypeIII}$	8.0 cm^{-1}, 8.0 cm^{-1}
Δ_{TypeIV}	-10.0 cm^{-1}
Δ_{TypeV}	-1.0 cm^{-1}, 20.0 cm^{-1}

Table 3: Parameters used in the simulation of XAO, including dipole moments (μ), wavenumber positions ($\tilde{\nu}$), and nearest-neighbor coupling constants (Δ). The two coupling constant values listed for the turn conformations, i.e. types I-V, are for residues *i* and *i* + 1 of the turn, respectively.

We started the analysis by first testing the model of Shi et al. (*45*), which invokes a two-state model per residue, namely PPII and β-strand. In this case the measured $^{3}J_{C\alpha HNH}$ coupling constants determine the respective molar fractions. The result of this simulation is depicted in **Figure 14**. Apparently neither the asymmetry of the IR nor that

of the isotropic Raman band profile is reproduced. The width of the anisotropic band profile is also not accounted for. Only the VCD signal is close to the experimentally obtained couplet. This simulation provides compelling evidence for the above formulated supposition that the two-state model of Shi et al. (*45*), which yields very high PPII fractions for the alanine residues (between 0.65 and 0.95), is not in agreement with our data. We calculated the corresponding average end to end distance by utilizing:

$$R = \sum_{m=1}^{Q} \sum_{k=1}^{10} \left(\vec{r}_{CN,k,m} + \vec{r}_{CC_\alpha,k,m} + \vec{r}_{C_\alpha C,k,m} \right) f_m \qquad (28)$$

where m labels a given sequence of PPII and β-strand residues, and f_m is the fraction of the sequence m. The total number of sequences for the two-state model is Q = 1024. The vectors denote the length and orientations of the CN, NC_α, and $C_\alpha C$ bonds with respect to a coordinate system defined in **Section 1.3.4**. **Eq. (26)** neglects the C-terminal residue for which we do not have any data. For the two-state model we obtained 30 Å, which far exceeds that of 18.1 Å, derived from the radius of gyration of 7.4 Å, as obtained from SAXS data (*86*).

In a second step, we performed a simulation which considered the entire above mentioned conformational manifold, containing representatives of all the local turn structures which emerged from the MD simulations of Makowska et al. (*87*). We used the result of these simulations to initially guess the mole fractions of the respective residue conformations and subsequently varied them to optimize the simulations with respect to the measured profiles. This yielded a reasonable reproduction of the IR band

profile, but the asymmetry of the isotropic band profile, the band profile of anisotropic scattering, and the observed non-coincidence between isotropic Raman scattering and IR absorption, were all not accounted for. Moreover, the VCD signal was nearly eliminated in all of these simulations. This is *per se* an important result in that it indicates that a pronounced amide I' VCD signal is inconsistent with a statistical coil sampling, in agreement with arguments by Dukor and Keiderling (*59*).

In the third and final step, we somewhat reduced the number of sampled conformations for all residues. **Table 4** lists the (representative) conformations and the respective optimized mole fractions considered for this simulation. With respect to the turn structures, we confined ourselves to β-turn type III' for residues 2 and 3 (XA), type III for residue 4 (A), type IV for residues 2-9 ($X_2A_7O_2$), and type V for residues 2 and 3 (XA). PPII and β-strand conformations were considered for all residues. Altogether, we thus considered 142,884 different peptide conformations. **Figure 16** shows the best set of profiles obtained from simulations based on this structural model. The two Raman profiles are nearly perfectly reproduced. The simulated IR-band profile is slightly broader than the experimental one and the negative part of the VCD couplet is somewhat underestimated. Nevertheless, the agreement between experiment and simulation is more than satisfactory, particularly in view of the fact that the utilized structural model is still crude, since it considers only representative conformations rather than distributions.

Residue	PPII	β-strand	β-turn type III	β-turn type III'	β-turn type IV	β-turn type V	β-turn type II	β-turn type I
1	0.64	0.35						
2	0.46	0.29			0.25			
3	0.21	0.09		0.20	0.25	0.15	0.05	0.05
4	0.03	0.32	0.15		0.25	0.15	0.05	0.05
5	0.68	0.12			0.2			
6	0.66	0.14			0.2			
7	0.57	0.18			0.25			
8	0.46	0.29			0.25			
9	0.77	0.12			0.11			
10	0.55	0.35						

Table 4: Molar fractions of different residue conformations used to simulate the amide I' band profiles of XAO.

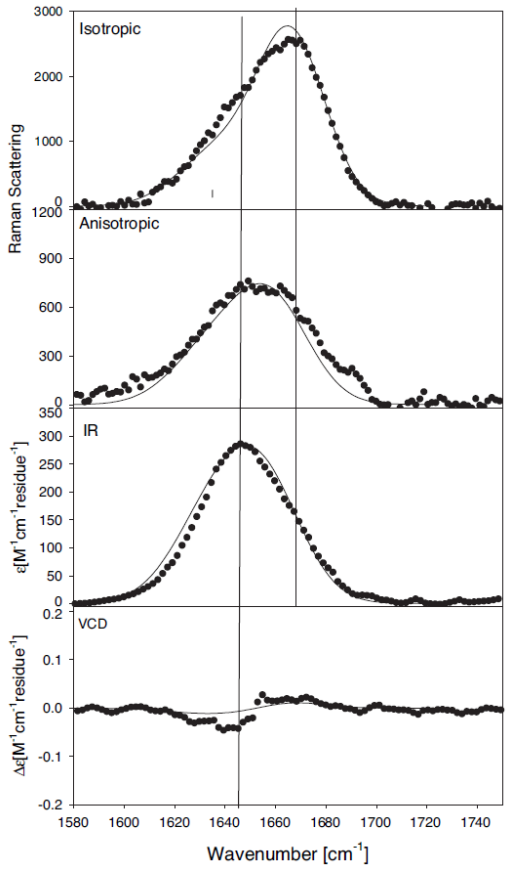


Figure 16: Experimental isotropic and anisotropic Raman, FTIR, and VCD spectra of the amide I' band region of XAO (pD = 2.2) (dotted) and the band profiles simulated by using the conformations listed in Table 4.

We used **eq. (26)** to calculate the end to end distance associated with the considered conformational blend, and obtained a value of 19.1 Å. This is in excellent agreement with the values derived from the SAXS experiment (*86*), and underscores the suitability of our analysis.

The temperature dependence of the UV-CD spectra (**Figure 15**), and of the $^{3}J_{C\alpha HNH}$ coupling constants, as observed by Shi et al. (*45*), has been interpreted as indicating a larger fraction of β-strand at higher temperatures. **Figure 17** compares the

IR, Raman, and VCD band profiles measured at 25° (black) and 65° C (red). In order to compare these profiles, we have to consider a wavenumber upshift of amide I' of 4 cm^{-1} for this temperature interval, which is due to anharmonic coupling to a low frequency mode involving the hydrogen bonds to water molecules (*67*), which was accounted for by downshifting the profiles measured at 65° C by this amount. The remaining difference between the experimental profiles at 25° and 65° should result from changes of the equilibrium between different conformations. Apparently, the first moments of the experimentally observed anisotropic Raman and IR band profiles, and to a lesser extent also that of the isotropic Raman band profile, are at lower wavenumbers than those of the respective solid line profiles in **Figure 17**. For IR, the high temperature band profile is more asymmetric. The non-coincidence between the peak positions of the IR and isotropic Raman profiles is reduced from 20 cm^{-1} (at 25° C), to 14 cm^{-1} (at 65° C). The VCD couplet is reduced at high temperatures. All these observations are consistent with the notion that a β-strand conformation is more populated at higher temperatures. This conformation gives rise to a reduced nearest neighbor coupling, which leads to the observed reduction of the IR/isotropic Raman non-coincidence and also of the VCD signal. The reduced wavenumber shift at high temperature is consistent with the results of DFT calculations, which predict that amide I' exhibits a slightly lower wavenumber position in the β-strand conformation than in PPII (*147*).

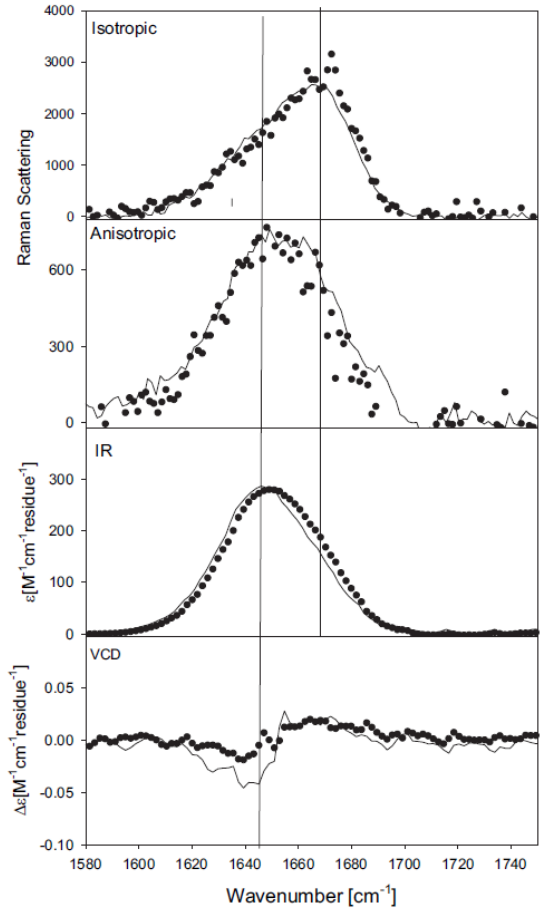


Figure 17: Comparison of the experimental isotropic and anisotropic Raman, FTIR, and VCD amide I' band profiles of XAO (pD = 2.2) at 25°C (solid line) and 65°C (dotted line). For comparison, the spectra at 65°C were downshifted by 4 cm^{-1}.

The result of their MD simulations led Makowska et al. to the conclusion that polyproline II is only one of many possible conformations sampled by alanine residues, and that the notion of a PPII propensity for alanine has to be rejected (*87*). This finding is at variance with results obtained from experimental investigations on short peptides (*48, 69, 95*), which are clearly indicative of a PPII propensity for alanine, in accordance with recent MD simulations of Garcia and coworkers (*85, 90*). On the other hand, it was clear that the radius of gyration value obtained from SAXS experiments was inconsistent

with a predominant PPII population for all alanine residues of the XAO peptide. Our results strongly indicate that the truth lies between the extremes invoked in the studies of Shi et al. (*45*) and Makowska et al. (*87*). The blend used for the simulation depicted in **Figure 16** still exhibits a substantial PPII fraction for residues 1 and 5-10. As a matter of fact, this is the only way to explain the significant non-coincidence between IR and isotropic Raman peak positions, which reflect the average nearest neighbor coupling. This can only be achieved by a substantive fraction of PPII like conformations (*124*, *143*). In particular, in the PPII conformation the amide I transition dipole moments are oriented nearly perpendicular to the helix axis, giving rise to strong IR intensity in the low-wavenumber amide I mode, E_1 (*109*). The high wavenumber in-phase amide I component, *A*, however, has a strong intensity in the Raman spectrum, resulting in a non-coincidence between the peaks of the amide I bands in the IR and isotropic Raman spectra (*109*).

Our results also agree to some extent with Makowska et al., in that the consideration of turn-like conformations is necessary to explain our experimental results (*87*). We invoked the results of these authors in assuming that type III' and V turn structures predominantly involve residues 2, 3 and 4, but we would have obtained similar results if we had selected another pair of residues. Makowska et al. obtained substantial type IV β-like conformations for residues 3-10 (*87*). In principle, this is a rather heterogeneous class of conformations, containing all turn like structures which are different from types I, II, III, and V, or their stereo-isomers. From the center of this region, we selected a representative conformation, i.e. $(\phi, \psi) = (-20^o, 20^o)$, which allows negative nearest neighbor coupling between amide I oscillators. The consideration of

negative coupling is absolutely necessary for reproducing the asymmetry of the isotropic Raman band.

Taken together, our results suggest that longer alanine peptides can be more heterogeneous than one would expect based on the analysis of short peptide fragments. A re-investigation of tri- and tetraalanine (see Section 1.3.6) suggested that a substantial population of turn structures is inconsistent with our spectroscopic data (*146*). The reason for the local turn propensity of alanine in XAO still needs to be explored. That the properties of XAO might reflect some general properties of unfolded alanine based peptides is underscored by a recent, very elegant ESR study on the unfolded state of $P_2HG_3WP(A_4K)_2CA_4KA$ by Jun et al. (*93*), which also indicates a much smaller end to end distance than expected for an extended molecule. The results of the present study might indicate that alanine shows a propensity for turn conformations in the vicinity of charged residues (e.g. X). The possibility of turn formations as initiation of protein folding had been discussed in detail by Wright et al. (*148*), in view of their NMR studies on unfolded peptides and proteins. The undeniable occurrence of turn structures, however, cannot obscure the fact that alanine still exhibits a substantial PPII propensity in longer peptides, which is in full accordance with findings derived from the coil libraries of Serrano (*61*) and Jha et al. (*80*), but at variance with the more ‘disordered’ library reported by Dobson and coworkers (*84*).

1.4 Materials and Methods

1.4.1 Materials

H-(Ala-Ala-Lys-Ala)-OH (AAKA) was custom synthesized by Celtek Peptides (Nashville, TN), with a purity of > 98%. The peptide contained a small amount of TFA, which was removed by dialysis in a Spectra/Por CE Float-A-Lyzer bag, and the peptide was lyophilized overnight. For IR, VCD, and Raman measurements, the peptide was dissolved at a concentration of 0.1 M, in 0.025 M $NaClO_4$ in D_2O (pD = 1), where the ClO_4^- Raman peak at 934 cm^{-1} was used as an internal standard (*149*), and the D_2O was made acidic by addition of DCl. Both D_2O and DCl were obtained from Sigma-Aldrich. After addition of the peptide, the solution had pD = 1.5. The pD values were determined by utilizing the method of Glasoe and Long to correct the values obtained from pH electrode measurements (*150*).

L-lysyl-L-alanine (KA) was purchased from Bachem Bioscience, Inc (> 96 % purity). The peptide was dissolved in D_2O at a concentration of 0.2 M (pD = 1.1).

Ac-$(Daba)_2$-$(Ala)_7$-$(Orn)_2$-NH_2 (XAO) was obtained as a gift from the laboratory of Dr. Pandé at Stanford University. In order to remove residual TFA, which absorbs in the vicinity of the amide I region, the peptide was dialyzed in a 1 mL Spectra/Por CE Float-A-lyzer dialysis bag , with a MWCO of 500, and lyophilized overnight. For Raman, FTIR, and VCD experiments, the peptide was dissolved at a concentration of 25.1 mg/mL in acidified D_2O. The pD of the resulting peptide solution was 2.2. For UV-CD measurements, the peptide solution was diluted 10-fold, with acidified D_2O.

1.4.2 Polarized Raman Spectroscopy

The polarized Raman spectra were obtained with the 442 nm (32 mW) excitation from a HeCd laser (Model IK 4601R-E, Kimmon Electric US). The laser beam was directed into a RM 100 Renishaw confocal Raman microscope, and focused onto a 1.0 mm Q Silica cell with a thin glass cover slip using a x50 objective. The scattered light was filtered with a 442 nm notch filter, dispersed by a single-grating 2400 l/mm grating and imaged onto a back-thinned Wright Instrument CCD. It was polarized by a combination of a linear polarizer and a λ/2 plate. The latter rotates the y-polarized light (perpendicular to the laser polarization) into the x-direction to achieve an optimal spectrometer transmission. All spectra were recorded in the 'continuous' mode. At least 5 scans were measured for each peptide sample for both polarization directions. All spectra were averaged for each polarization direction in order to eliminate some of the noise. The reference spectra were appropriately subtracted from the sample spectra. High temperature Raman measurements were carried out using an LTS-350 temperature controlled microscope slide holder from Linkam (UK).

1.4.3 FTIR/VCD Spectroscopy

The FTIR and VCD spectra were recorded with a Chiral IR^{TM} Fourier Transform VCD spectrometer from Bio Tools. The respective peptide samples were placed into a cell with a pathlength of 42.5 μm. The spectral resolution was 8 cm^{-1} for measurements of AAKA and 4 cm^{-1} for measurements of XAO. The VCD and IR were both collected as

one measurement for a combined total time of 720 minutes (648 min. for VCD, 72 min. for IR). The appropriate background spectrum was manually subtracted from each peptide spectrum using the program MULTIFIT (*127*). For XAO, the IR and VCD spectra were also obtained at 65°C, using a temperature controlled cell.

1.4.4 Spectral Analysis

All IR and Raman spectra were treated by using the program MULTIFIT (*127*). The calibration of the Raman spectrum was checked by using the $NaClO_4$ band at 934 cm^{-1}. To eliminate solvent contributions, we measured the solvent reference spectra for both polarizations, which were then subtracted from the corresponding peptide spectra. The isotropic and anisotropic Raman intensities were calculated as:

$$I_{iso} = I_x - \frac{4}{3} I_y$$
$$I_{aniso} = I_y \qquad (29)$$

where I_x and I_y denote the Raman scattering polarized parallel and perpendicular to the polarization of the exciting laser light.

1.4.5 UV-CD Spectroscopy

Temperature dependent UV-CD spectra of AAKA were obtained using an OLIS DSM-10 UV/Vis CD spectrophotometer, located in the laboratory of Prof. Kai Griebenow, at the University of Puerto Rico, San Juan. A 1.0 mm pathlength quartz cell was used, and the

spectra were collected at 2 nm resolution. A peptide concentration of 1.0 mM in D_2O (pD = 1) was used. The sample was placed in a nitrogen purged OLIS CD module. The sample temperature was controlled by means of a Peltier-type heating system (accuracy $\pm 1^{o}$C). Spectra were recorded from 20° – 80° C at 10° intervals, averaging 5 scans at each temperature. The sample was allowed to equilibrate for 5 minutes at a given temperature. The spectra were collected as milliabsorbance as a function of wavelength and converted to molar absorptivities via Beer's Law.

The temperature-dependent UV-CD spectra of XAO in the wavelength range of 190 – 240 nm of XAO were obtained using a JASCO J-810 spectrapolarimeter in a 0.1 mm quartz cell. Spectra were acquired with a resolution of 0.05 nm and a scan speed of 500 nm/min. The measurements were performed in D_2O rather than in H_2O in order to allow a direct comparison with structural data obtained by vibrational spectroscopies. For each measurement, the sample was allowed to equilibrate for 5 minutes at the adjusted temperature prior to acquisition. The spectra were obtained by averaging ten scans and were collected as ellipticity as a function of wavelength and converted to molar absorptivity per residue using the following relationship:

$$\Delta\varepsilon = \frac{\theta}{32980 \cdot c \cdot l \cdot (n-1)} \tag{30}$$

where θ is the ellipticity in [*mdeg*], c is the concentration in [M], l is the pathlength in [cm], and n is the number of residues.

1.4.6 NMR Spectroscopy

NMR samples of AAKA were prepared by dissolving the peptide at ca. 5 mM concentration in either D_2O (99.8 atom % D; Cambridge Isotope Laboratories) or a mixture of 90% H_2O/10% D_2O. One- and two-dimensional proton NMR spectra were acquired at 25 °C using a Varian INOVA NMR spectrometer (Varian Inc., Palo Alto, CA) operating at a proton frequency of 499.9 MHz and equipped with a 5 mm dual broad-band *z*-gradient probe.

Two-dimensional proton correlation spectroscopy (COSY) (*151-152*) and NOE spectroscopy (NOESY) (*153-154*) experiments were measured in Department of Chemistry at Rutgers University (Newark, NJ). Spectra were recorded in the pure absorption mode by employing the TPPI improvement (*155-156*) of the States-Haberkorn-Ruben hypercomplex method (*157*). Selection of desirable coherences and artifact suppression were accomplished by *z*-gradients (COSY) and phase cycles of 4 (COSY) or 32 (NOESY) steps. The NOESY data set was acquired using a 600 ms mixing time. Typically, 256 t_1 increments of 2K complex data points over a 6 kHz spectral width were collected with 4 (COSY) or 32 (NOESY) scans per t_1 increment, preceded by 4 or 32 dummy scans, and a relaxation delay of 1.5 s. All spectra were acquired with the carrier offset placed on the water resonance, which was reduced by either solvent presaturation (*158*) or tailored excitation, using WATERGATE (*159-160*). The peptide NH resonances must not be in fast exchange with water since they were readily observed in 90% H_2O.

Data sets were processed on a Sun Blade 100 workstation (Sun Microsystems Inc., Palo Alto, CA) using the VNMR software package (Varian Inc., Palo Alto, CA). In

order to decrease t_1 ridges arising from incorrect treatment of the first data point in the discrete Fourier transform (FT) algorithm, the spectrum corresponding to the first t_1 value was divided by 2 prior to FT along t_1 (*161*). Shifted (COSY) or unshifted (NOESY) Gaussian window functions were used in both dimensions. Data sets were zero-filled in the t_1 dimension, yielding 1K x 1K final matrices that have not been symmetrized. Spectra were referenced to the HDO resonance of the D2O sample [4.76 ppm vs 4,4-dimethylsilapentanesulfonic acid (deuterated at carbons 2 and 3) at 0.00 ppm].

The amino acid spin systems of AAKA were identified from the COSY spectra in D2O and 90% H2O/10% D_2O (*162*). The sequential assignments were made from the R CH_i-NH_{i+1} cross-peaks of the NOESY spectrum in 90% H2O/10% D2O. The $^3J_{C\alpha HNH}$ values were obtained from one-dimensional spectra of 32K complex data points over a 6 kHz spectral width.

^{1}H NMR experiments on AAAA (>98% purity; Bachem Bioscience Inc.) in aqueous solution (0.025 M) were measured at the Instituto per I Processi Chimico-Fisici, Consiglio Nazionale della Richerche Area della Ricerca di Pisa via G. Moruzzi (Pisa, Italy), and were performed at room temperature on a Bruker AMX-300 WB spectrometer equipped with a 5 mm reverse probe, using standard Bruker pulse programs. The $\pi/2$ pulse was 5.7 μs, and the relaxation delay was set to 4 s. The water signal was presaturated using a 4 s soft pulse (B_1 = 25 Hz). The assignment of the signals (**Table 5**) was performed through the cross-peaks of the 1H DQF-COSY and NOESY experiments in the amide-H^α proton region (*162*), the former yielding the *J* connectivity's and the latter allowing the identification of adjacent residues from the H^α and the amide proton of the subsequent residue cross-peak. The signal from the protons bound to the terminal

nitrogen could not be identified because of their rapid exchange with water. The $^{3}J_{C\alpha HNH}$ coupling constants were obtained from the DQF-COSY experiment. NOESY experiments with mixing times ranging from 200 to 400 ms were acquired.

Residue No.	NH	CH^{α}	CH^{β}
1	---	4.01	1.46
2	8.53	4.27	1.34
3	8.36	4.22	1.32
4	7.88	4.04	1.26

Table 5: ^{1}H chemical shifts (ppm) of AAAA in aqueous solution at 0.025 M and room temperature. The proton chemical shifts are referenced to the water signal, which was set at 4.75 ppm.

CHAPTER 2

Misfolding and Fibril Formation of a Short Alanine-Rich Peptide Probed by Vibrational Circular Dichroism (VCD) Spectroscopy

2.1 Synopsis

The aggregation of various misfolded peptides and proteins into well-ordered fibrils is linked to a number of debilitating human ailments. In addition to such disease-related proteins, short peptides with no pathological affiliations can form similar structures. Herein is presented a spectroscopic characterization of the misfolding and fibril formation of a short alanine-rich peptide, namely Ac-(AAAAKAAY)-NH_2 (AKY8). A detailed understanding of the interactions involved in peptide and protein misfolding is of utmost biomedical importance, so that short peptides such as AKY8 present convenient model systems for studying the mechanism of fibril formation.

The propensity of AKY8 to aggregate is found to be promoted by the C-terminal tyrosine residue, consistent with suggestions that aromatic interactions can play important roles in peptide and protein aggregation. Moreover, AKY8 fibrils exhibit an enhanced VCD couplet in the amide I' region. This chapter provides insight into the interactions involved in the misfolding and aggregation of short peptides, and highlights the use of vibrational circular dichroism (VCD) as a novel means to probe the kinetics of aggregation, which has the potential to provide information on the local structure formation during fibrillogenesis.

2.2 Background Information

2.2.1 Protein Misfolding

As the typical protein energy landscape is now considered heterogeneous and highly complex (see **Figure 2**, **section 1.2.4**), a polypeptide chain may inhabit various local minima along its journey down the energy funnel. Misfolded states are representative of these various local minima located along the pathway to the correctly folded state of a protein (4). Owing to the complexity of the protein energy landscape, various states are often accessible in the biological environment, which depend on specific conditions and various kinetic and thermodynamic factors (*4, 163*). These various accessible states of a protein molecule, as illustrated in **Figure 18**, include, in addition to the unfolded and native states, aggregated states such as fibrils, which are affiliated with various human pathologies. A complete and detailed description of the protein energy landscape requires an understanding of the energetics accompanying the interconversion between the assortment of accessible states, not only for understanding amyloid formation for disease treatment and prevention, but also for the exploitation of the peptide self-assembly process for various biotechnological applications.

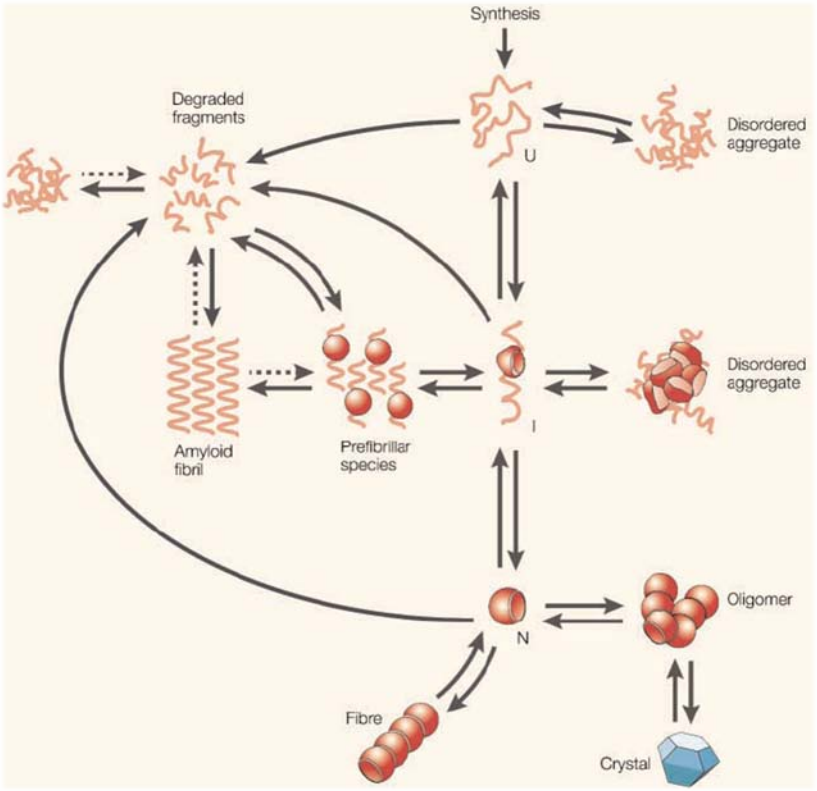


Figure 18: Various states accessible to a protein molecule, including the unfolded (U), intermediate (I), and folded (F) states. Adapted from ref(4).

2.2.2 Misfolding and Human Disease

The ability of a protein to achieve its native 3-D structure is important for maintaining proper biological function. It is not surprising, therefore, that the inability to achieve or maintain a native fold is linked to various debilitating human ailments (*164*). In particular, aggregation of misfolded peptides and proteins is affiliated with common neurodegenerative diseases such as Alzheimer's, Huntington's, Parkinson's, and the prion diseases, as well as the systemic amyloidoses (*164-166*). The largest class of 'protein misfolding' or 'conformational' diseases, as they are often referred, result from

the aggregation of specific misfolded or intrinsically disordered proteins (IDP's) into insoluble filamentous structures termed 'amyloid fibrils' (*4*, *164-165*, *167-168*). The number of conformational diseases currently known is quickly approaching 50. A list of some conformational diseases associated with amyloid deposition, including the culprit protein believed to result in the formation of insoluble cellular deposits is listed in **Table 6**, for the various neurodegenerative diseases, and systemic and localized amyloidoses.

Disease	Aggregating protein or peptide	Number of residues[a]	Native structure of protein or peptide[b]
Neurodegenerative diseases			
Alzheimer's disease	Amyloid β peptide	40 or 42	Natively unfolded
Spongiform encephalopathies	Prion protein or fragments thereof	253	Natively unfolded (residues 1-120) and α-helical (residues 121-230)
Parkinson's disease	α-Synuclein	140	Natively unfolded
Dementia with Lewy bodies	α-Synuclein	140	Natively unfolded
Frontontemporal dementia with Parkinsonism	Tau	352-441	Natively unfolded
Amyotrophic lateral sclerosis	Superoxide dismutase 1	153	All-β, Ig like
Huntington's disease	Huntingtin with polyQ expansion	3144	Largely natively unfolded
Spinocerebellar ataxias	Ataxins with polyQ expansion	816	All-β, AXH domain (residues 562-694); the rest are unknown
Spinocerebellar ataxia 17	TATA box-binding protein with polyQ expansion	339	α+β, TBP like (residues 159-339); unknown (residues 1-158)
Spinal and bulbar muscular atrophy	Androgen receptor with polyQ expansion	919	All-α, nuclear receptor ligand-binding domain (residues 669-919); the rest are unknown
Hereditary dentatorubral-pallidoluysian atrophy	Atrophin-1 with polyQ expansion	1185	Unknown
Familial British dementia	ABri	23	Natively unfolded
Familial Danish dementia	ADan	23	Natively unfolded
Nonneuropathic systemic amyloidoses			
AL amyloidosis	Immunoglobulin light chains or	»90	All-β, Ig like

	fragments		
AA amyloidosis	Fragments of serum amyloid A protein	76-104	All-α, unknown fold
Familial Mediterranean fever	Fragments of serum amyloid A protein	76-104	All-α, unknown fold
Senile systemic amyloidosis	Wild-type transthyretin	127	All-β, prealbuminn like
Familial amyloidotic polyneuropathy	Mutants of transthyretin	127	All-β, prealbuminn like
Hemodialysis-related amyloidosis	β2-microglobulin	99	All-β, Ig like
ApoAI amyloidosis	N-terminal fragments of apolipoprotein AI	80-93	Natively unfolded
ApoAIV amyloidosis	N-terminal fragment of apolipoprotein AII	98	Unknown
ApoAIV amyloidosis	N-terminal fragment of apolipoprotein AIV	»70	Unknown
Finnish hereditary amyloidosis	Fragments of gelsolin mutants	71	Natively unfolded
Lysozyme amyloidosis	Mutants of lysozyme	130	α+β, lysozyme fold
Fibrinogen amyloidosis	Variants of fibrinogen α-chain	27-81	Unknown
Icelandic hereditary cerebral amyloid angiopathy	Mutant of cystatin C	120	α+β, cystatin like
Nonneuropathic localized diseases			
Type II diabetes	Amylin, also called islet amyloid polypeptide (IAPP)	37	Natively unfolded
Medullary carinoma of the thyroid	Calcitonin	32	Natively unfolded
Atrial amyloidosis	Atrial natriuretic factor	28	Natively unfolded
Hereditary cerebal haemorrhage with amyloidosis	Mutants of amyloid β peptide	40 or 42	Natively unfolded
Pituitary prolactinoma	Prolactin	199	All-α, 4-helical cytokines
Injection-localized amyloidosis	Insulin	21 + 30	All-α, insulin like
Aortic medial amyloidosis	Medin	50	Unknown
Hereditary lattice corneal dystrophy	Mainly C-terminal fragments of kerato-epithelin	50-200	Unknown

Corneal amyloidosis associated with trichasis	Lactoerrin	692	α+β, periplasmic-biding protein like II
Cataract	γ-Crystallins	Variable	All-β, g-crystallin like
Calcifying epithelial odontogenic tumors	Unknown	~46	Unknown
Pulmonary alveolar proteinosis	Lung surfactant protein C	35	Unknown
Inclusion-body myositis	Amyloid β peptide	40 or 42	Natively unfolded
Cutaneous lichen amyloidosis	Keratins	Variable	Unknown

Table 6: Human diseases associated with formation of extracellular amyloid deposits or intracellular inclusions with amyloid-like characteristics. Adapted from ref(*165*).

A defining feature of the various pathologically diverse diseases which arise from the misfolding and subsequent aggregation of specific proteins is the accumulation of intra- or extra-cellular proteinaceous deposits, consisting primarily of fibrillar species. In particular, the brains of patients with Alzheimer's disease contain deposits of the Aβ protein known as amyloid plaques (*169*), while Parkinson's patients develop Lewy bodies within neurons consisting predominantly of the α-synuclein protein (*170*). It is typically believed that, e.g. amyloid plaque build-up inhibits proper neuronal signal transmission, thus causing neurodegeneration (*171*). A recent study, however, suggests that the deposits, instead, seek to protect neurons from stressful conditions which often result in pathogenesis (*172*).

Amyloid-like plaques and Lewy bodies are formed from the accumulation of unbranched fibrils which adopt a characteristic cross-β core, where so-called protofilaments intertwine with their constituent β-strands aligned perpendicular to the fibril axis, as shown in **Figure 19** (*173*). The thickness of the fibrils is determined by the number of protofilaments. Typical amyloid-like fibrils are 6-12 nm in diameter and

usually consist of 2-6 protofilaments (3). These fibrils show well-defined diffraction patterns, and specifically bind dyes such as Congo red and Thioflavin T. Despite the sequence diversity of the many known fibril-forming polypeptides and proteins, and their associated diseases, the architecture of the resulting fibrils is remarkably similar in all cases. Even short peptides and globular proteins not affiliated with any known disease can form amyloid-like fibrils under certain conditions. These results have led to the suggestion that amyloid fibril formation is a general feature of all polypeptides (*4, 164-165*).

It is now of current belief that the conformational disorder inherent in misfolded proteins is a prerequisite for aggregation and fibril formation. Disordered regions result in the exposure of non-native hydrophobic segments, which would otherwise be buried within the globular protein fold (*164, 174*). Disordered regions of a protein, by nature, are more solvated, allowing for aberrant interactions of exposed hydrophobic regions, such as non-native interactions which are believed to play a role in promoting aggregation of the diverse assortment of disease-related peptides and proteins (*175*).

Peptide and protein misfolding can be envisaged to result from two scenarios which produce partial conformational disorder, namely the partial unfolding of natively structured peptides and proteins, or the partial folding of intrinsically disordered/unfolded peptides and proteins (*164-165, 168, 176*). Large amounts of experimental data suggest that globular proteins require partial unfolding prior to aggregation (*166, 177-178*).

2.2.3 Functional Fibrils

In addition to the formation of toxic protein fibrils implemented in various human diseases, there are increasing reports of these structures being exploited by living systems to carry out important biological functions (*165*). A list of proteins from various bacteria which can form nonpathological amyloid-like fibrils exhibiting specific functions can be found in **Table 7**. For example, amyloid-like fibrils formed by the protein spidroin are the major constituent of spider silk (*179*). Other examples include the fibril-forming bacterial proteins which are deposited at the bacterial surface and promote binding to inert substrates (e.g. epithelial cells). In particular, the *E. coli* protein curlin forms amyloid-like fibrils which promote cell adhesion to the epithelial surface and favors colonization (*180*). Thus, if regulated under well-controlled conditions, fibril formation can be advantageous and exploited in living systems to carry out specific biological functions.

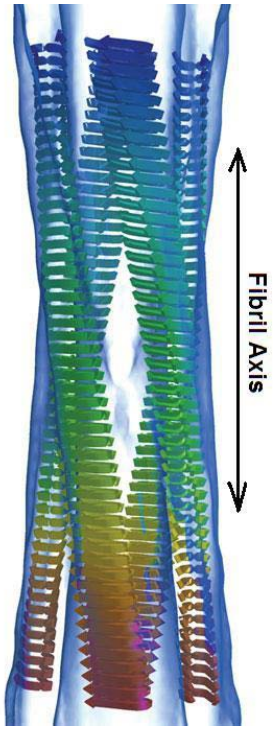


Figure 19: Cross-β structure of amyloid fibrils. Individual β-strands are arranged perpendicular to the fibril axis. Figure was adapted from ref(*181*).

Protein	Organism	Function of the resulting amyloid-like fybrils
Curlin	*Escherichia coli* (bacterium)	To colonize inert surfaces and mediate binding to host proteins
Chaplins	*Streptomyces coelicolor* (bacterium)	To lower the water surface tension and allow the development of aerial hypae
Hydrophobin EAS	*Neurospora crassa* (fungus)	To lower the water surface tension and allow the development of aerial hypae
Proteins of the chorion of the eggshell	*Bombyx mori* (silkworm)	To protect the oocyte and the developing embryo from a wide range of environmental hazards
Spidroin	*Nephila edulis* (spider)	To form the silk fibers of the web
Intralumenal domain of Pmel17	*Homo sapiens*	To form , inside melanosomes, fibrous striations upon which melanin granules form
Ure2p (prion)	*Saccharomyces cervisiae* (yeast)	To promote the uptake of poor nitrogen sources ([URE3])
Sup35p (prion)	*Saccharomyces cervisiae (yeast)*	To confer new phenotypes ([PSI+]) by facilitating the readthrough of stop codons on mRNA
Rnq1p (prion)	*Saccharomyces cervisiae (yeast)*	Not well understood ([RNQ+]), also known as [PIN+], phenotype
HET-s (prion)	*Podospora anserina* (fungus)	To trigger a complex programmed cell death phenomenon (heterokaryon incompatibility)
Neuron-specific isoform of CPEB (prion)	*Aplisia californica* (marine snail)	To promote long-term maintenance of synaptic changes associated with memory storage

Table 7: Proteins forming naturally nonpathological amyloid-like fibrils with specific functional roles. Adapted from ref(*165*).

2.2.4 Fibril Formation

Fibril formation, or fibrillogenesis, is believed to be pathogenic (*177*, *182*). Therefore, uncovering ways to inhibit the unwanted processes is necessary to combat various amyloid diseases (*183*). In this regard, elucidation of the mechanism of the self-assembly of polypeptides is of fundamental importance in structural biology (*164*), and of utmost biomedical significance. A complete understanding of fibril formation requires a detailed structural and dynamic characterization of the various prefibrillar species (e.g. oligomers and protofibrils), and the principles which govern their transition into amyloid fibrils (*165*). These prefibrillar species are believed to be in equilibrium with both the monomeric and fibrillar states, as illustrated in **Figure 20**. Moreover, identification of

specific residues and regions in the primary amino acid sequence of amyloidogenic peptides and proteins is necessary (*165*).

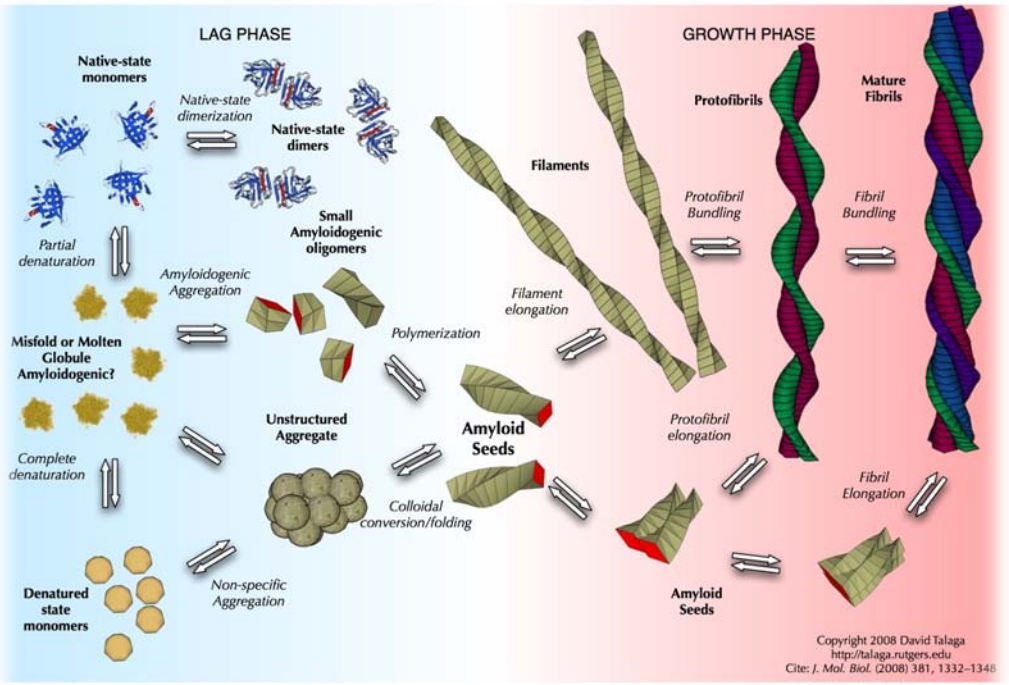


Figure 20: Various protein species which can be populated and co-exist during protein aggregation/fibril formation. Adapted from ref(*184*).

Amyloid formation usually proceeds via a "nucleation growth mechanism," in which a lag phase is typically observed prior to an exponential growth phase (*165, 185-186*), as shown in **Figure 21**. The lag phase is believed to harbor the formation of oligomeric species, which act the 'nuclei' for fibril formation. Many studies now indicate that it is the soluble oligomeric species that are pathogenic, and not the insoluble fibrils (*3, 166, 187-188*). In such cases, the formation of fibrils of such toxic species might be considered as a cell defense mechanism.

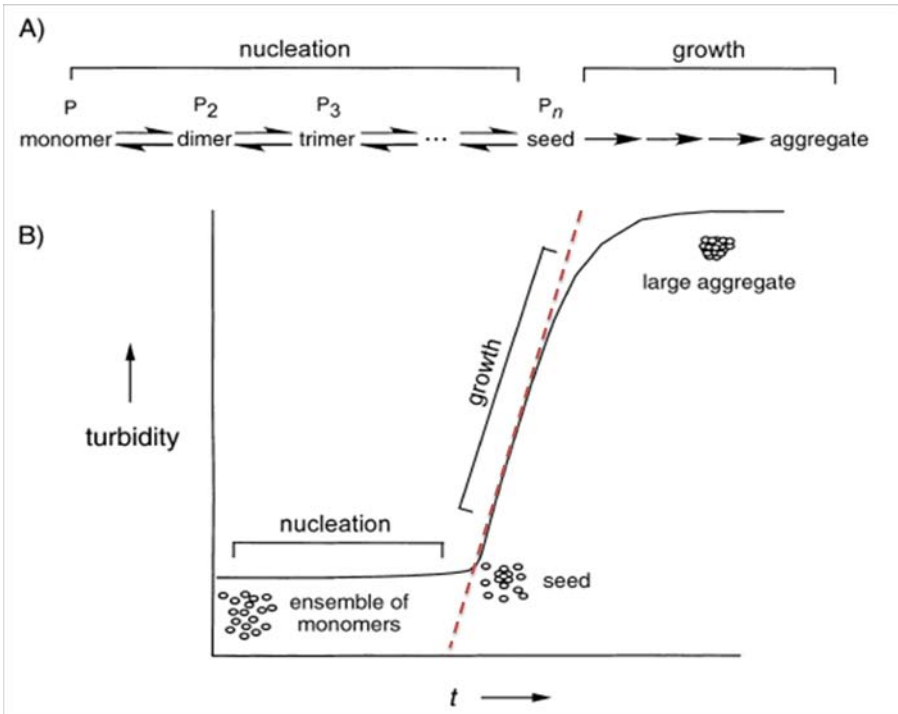


Figure 21: (A) Two stages of protein aggregation, i.e. nucleation and aggregate/fibril growth (B) Typical aggregation curve of nucleated aggregation, depicting turbidity of solution as a function of time, where turbidity is proportional to the formation of insoluble fibril aggregates. The dotted red line indicates the maximal growth rate. Figure was taken from ref(*186*) and adapted.

2.2.5 Cellular Control of Misfolded Species

In vivo, misfolded peptides and proteins often result from either amino acid mutations or improper translation (*3*, *165*). Mutated proteins are not always able to adopt their native fold. Since structure often dictates function, such improperly folded species usually only exhibit a fraction of their ability to carry out their proper functions. Pathological states resulting from folding inefficiency, for example in cystic fibrosis, ensue from the reduction in the amount of correctly folded protein necessary for normal functioning, due

to an increased amount of misfolded species which are targeted for degradation *in vivo* (*3, 165*).

Since adoption of a correct 3-D fold is often necessary for proper biological functioning, living systems have evolved ways in which to prevent the proliferation of incompletely or improperly folded proteins, which can favor inappropriate interactions and increase the likeliness of aggregation (*164*). These cellular control processes are a driving force of protein evolution (*189*), and include the 'unfolded protein response' (UPR) in the endoplasmic reticulum (ER) (*190*) and the heat shock response (HSR) in the cytoplasm (*164*). These regulatory processes typically rely on the performance of chaperones to guide rogue protein species to proper locations for their ultimate degradation and disposal (*164, 190*). The duties of protein chaperones include the oversight of correct folding of nascent polypeptide chains and refolding of misfolded chains to guide various later stages of folding, and to protect incompletely folded proteins from participating in inappropriate interactions which might favor aggregation (*3, 191*). In particular, some chaperones interact with nascent chains as they are being synthesized on the ribosome, while others work by interact with proteins at later stages of the folding process (*192-193*). **Figure 22** shows a schematic illustrating the cellular balance between proper folding vs. misfolding/disposal in the ER.

Despite the attempts made by a cell's regulatory machinery to properly dispose of misfolded and partially folded proteins, such aberrant species do thrive. The regulatory processes are not fool-proof, and rogue species often elude these quality-control checks. Some estimates have suggested that nearly half of newly synthesized polypeptides are targeted for disposal (4). It has been speculated that perhaps cells expend such large

quantities of energy by synthesizing and then destroying proteins for as yet unrecognized purposes (4).

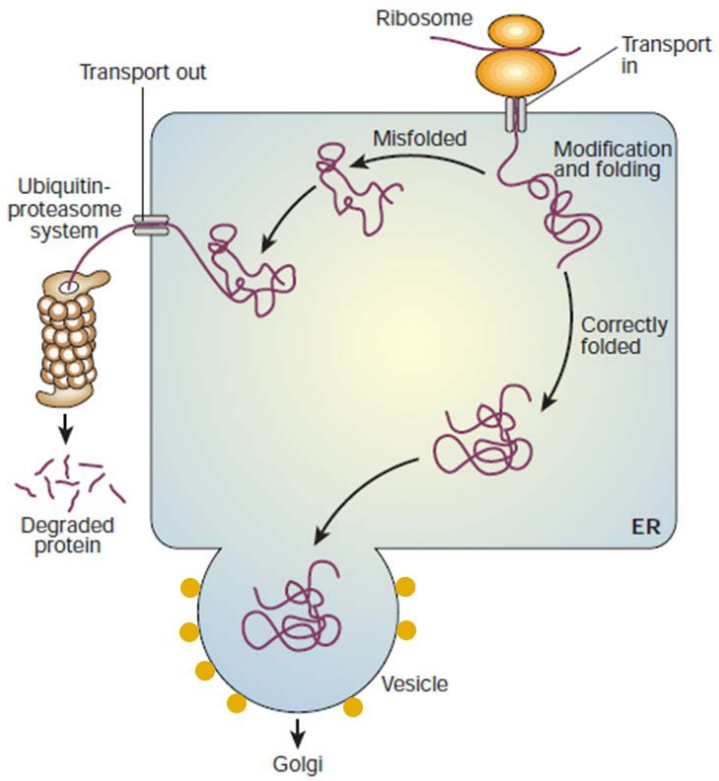


Figure 22: Regulation of protein folding in the endoplasmic reticulum. This figure illustrates two distinct pathways that a newly synthesized polypeptide chain may take after translation from the ribosome. These are: (1) proper folding, where the protein is sent further downstream to carry out its normal function in its destined are, or (2) degradation, where a mutated or improperly folded protein is ultimately chaperoned to the ubiquitin machinery. This figure was taken from ref(*164*) and modified.

2.2.6 Intrinsically Disordered Proteins (IDP's)

In addition to the misfolded species of globular proteins which expose various buried hydrophobic regions, or those mutated species which elude cellular regulatory processes *in vivo*, there also exist 'intrinsically disordered' or 'natively unfolded' peptides and proteins, which, owing to their inherent disorder, are also linked to various human diseases (*168*). The existence of intrinsic disorder challenges the well-established structure/function dogma of protein science, namely that the 3-dimensional fold of a protein dictates its function. The discovery of IDP's with specific and important functions clearly contradicts such a notion (*168*).

The term natively unfolded was introduced in 1994 to describe the nature of the tau protein, which has been linked to Alzheimer's disease (*176*). In contrast to molten globule proteins, which, although somewhat flexible, are still rather compact and contain elements of secondary structure, IDP's are flexible and non-compact, and exhibit very little secondary structure under physiological conditions (*176*). Although lacking in any major elements of secondary structure, most of these IDP's carry out important functions, such as molecular recognition and protein modification (*98, 174, 176*). A few examples of disordered proteins and their respective functions are listed in **Table 8** (*82*). From this point on, the term intrinsically disordered proteins (IDP's) will be employed when referred to those proteins lacking in any of the major structural content (α-helix, β-strand, β-sheet, β-turn, γ-turn, or PPII).

Protein	Function
fd Phage	Membrane penetration
Histone octamer	DNA replication, Transcription, Chromosome remodeling
Clusterin	Protein detergent
Calcineurin	Display of calmodulin target
Calsequestrin	Ca^{++} binding

Table 8: Examples of disordered proteins with well-defined biological functions. This table was adapted from ref(*82*).

Both misfolded and IDP's result in the exposure of residues and regions that promote aggregation. The distinction between the two classes is attributed to the fact that while misfolding often results in partial *unfolding*, IDP's typically undergo partial *folding*. In particular, transformations that result in favorable non-native interactions, including hydrogen and hydrophobic bonding, and favorable electrostatic interactions, can promote the aggregation and fibril formation of such disordered segments (*176*).

2.2.7 Fibril Formation is A General Feature of all Polypeptides

Owing to the similarity of structural features of amyloid-like fibrils formed by various disease- and non-disease-related peptides and proteins, it has been suggested that aggregation and fibril formation is an inherent property of all polypeptides and proteins (*3, 164, 166, 191*). Fibril formation is not restricted only to those peptides and proteins which are linked with various diseases. In fact, many studies have demonstrated the potential of various peptides and proteins that are not affiliated with any known disease to form amyloid-like fibrils *in vitro* (*191, 194*). Even myoglobin can form fibrils under

well-defined conditions (*195*). A list of non-disease related amyloidogenic peptides and proteins is given in **Table 9**.

Non-Disease-Associated Amyloidogenic Proteins or Peptides

Protein or Peptide
SH3 p85 α PI3 kinase (*Bovine*)
Fibronectin type III module (*Mouse*)
Betabellins 15D, 16D (*Synthetic*)
GAGA factor (*Drosophila*)
Acylphosphatase (*Horse, S. solfataricus*)
Monellin (*Dioscoreophyllum camminsii*)
CspB (1–22 fragment) (*Bacillus subtilis*)
Glycoprotein B (fragment) (*Herpes simplex virus*)
Fiber protein peptide 355–396 (*Adenovirus*)
B1 Ig binding domain (*Streptococcus*)
Phosphoglycerate kinase (*Yeast*)
Apolipoprotein C-II (*Human*)
ADA2H (*Human*)
Met aminopeptidase (*Pyrococcus furiosus*)
Lysozyme (*Hen*)
Apocytochrome c_{552} (*Haemophilus thermophilus*)
HypF N-terminal domain (*E. coli*)
Apomyoglobin (*Horse*)
Amphoterin (*Human*)
Polyamino acids (Synthetic)
Ure2P (*Yeast*)
Complement receptor (18–34 fragment) (*Human*)
Curlin CgsA subunit (*E. coli*)
Vl domain (*Mouse*)
Acidic FGF (*Notophthalmus viridescens*)
Stefin B (*Human*)
α-Lactalbumin (*Bovine*)
Prothymosin α (*Human*)
Barstar (*Bacillus amyloliquefaciens*)
Endostatin (*Human*)
Pmel17 (*Human*)
Albumin (*Bovine, glycated*)
Albebetin (*Synthetic*)
Acetylcholinesterase (586–599 fragment) (*Human*)
Core histones (*Bovine*)
α-Chymotrypsin

Table 9: **Non-disease-associated amyloidogenic proteins or peptides. Adapted from ref(3).**

This generic feature of amyloid-like fibrils, and the aggregated protein state in general, affords the extension of the typical protein energy landscape (see **Section 1.2.4**) to include such species, as shown in **Figure 23**. It was even suggested by Gazit (*186*) that the aggregated state may be the most stable proteins structure, consistent with the energy funnel depicted in **Figure 23**. From a fundamental viewpoint, the generic ability to form fibrils is intriguing, as it suggests that the many seemingly unrelated protein misfolding diseases share common molecular principles, so that therapeutic means to treat or prevent such diseases could result from common targets (*196*).

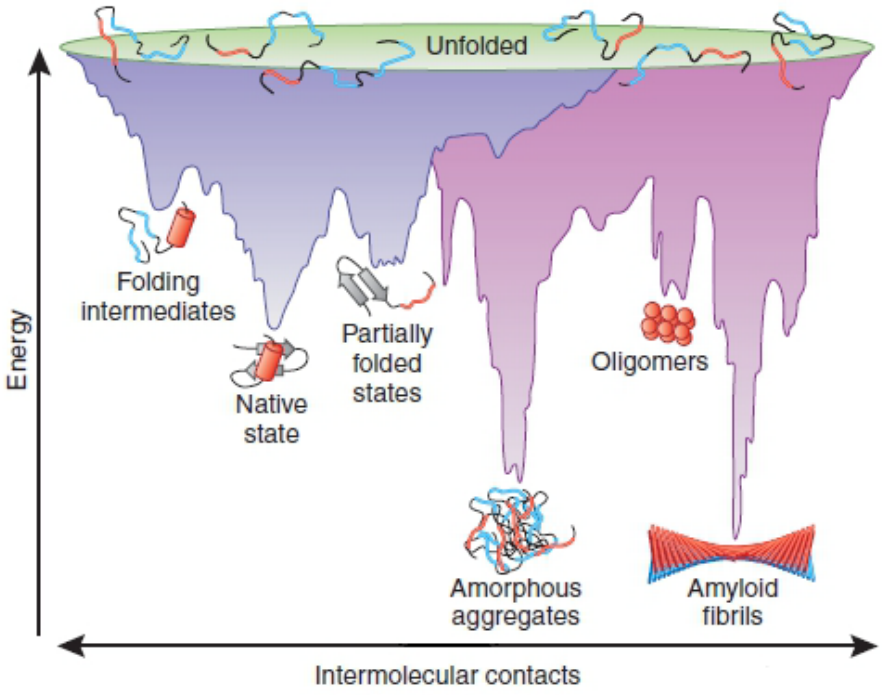


Figure 23: Extended protein energy landscape, exhibiting the amyloid fibril state as the lowest energy state of the landscape extension. This figure was adapted from ref(*197*).

2.2.8 Techniques Used to Probe Amyloid Fibril Formation

The most commonly used technique to confirm the presence of amyloid fibrils is the detection of a distinct green birefringence when viewed under crossed polarizers, after staining the fibrils with the diazo dye Congo red, but the reliability of this method of detection has been questioned (*198*). In particular, Kurana et al. recently suggested that Congo red does not bind only to amyloid fibrils (*198*). Various biophysical techniques are now commonly used to detect the presence of amyloid fibrils and probe fibril formation. These include x-ray crystallography (*199-200*) and x-ray diffraction, dynamic light scattering (DLS) (*201-203*), staining and fluorescence assays (*204-205*), solid-state NMR (*206-208*), and electron and atomic force microscopy (EM and AFM, respectively) (*209-210*). Vibrational spectroscopic techniques such as FTIR (*211*) and Raman (*212*), as well as UV-CD (3, 26) can also be used to monitor the increase in β-sheet marker bands, which occurs concomitantly with fibril formation.

While amyloid-like fibrils are not amenable to single crystal x-ray diffraction, a few studies have reported that some short amyloidogenic peptides can also form nano- and micro-crystals, which exhibit structural characteristics similar to those of the non-crystalline fibril species (*199-200*). The first such peptide found to form microcystals amenable to x-ray crystallographic techniques was the heptameric motif from the N-terminal region of the yeast prion protein (Sup35), namely GNNQQNY (*200*). Results from this study indicate that GNNQQNY microcrystals have a cross-β spine, consisting of a double β-sheet, with the side chains of each sheet protruding in such as way so as to form a tightly packed steric zipper. Also, the *de novo* designed 12-mer KFFEAAAKKFFE (AAAK) was found to form amyloid fibril-like crystals, composed of

antiparallel β-sheets which form a cross-β arrangement, characteristic of amyloid fibrils (*199*). A higher degree of inter-sheet order relative to the less rigid fibrils is characteristic of these so-called fibrous crystals (*199-200*).

More recently, VCD spectroscopy has shown sensitivity to amyloid fibrils, and was used to probe fibrillogenesis (*213*). In particular, fibrils formed from lysozyme and insulin were shown to exhibit enhanced VCD signals in the amide I band. Ma et al. suggested that the origin of the intensity enhancement might be due to coupling between neighboring β-sheets or from the twisting of individual sheets or braiding of multiple protofibrils (*213*). Since VCD can provide information on local structure, the use of this technique offers a novel and convenient approach to obtain information on local secondary structure of the amyloid fibrils as they form, especially when combined with other complementary techniques such as FTIR and Raman spectroscopy.

2.2.9 The Role of PPII in Misfolding and Aggregation

Conformational disorder is believed to be a prerequisite for fibril formation. Disordered regions of a protein are more solvent-accessible allowing for interactions of exposed hydrophobic regions, which can favor aggregation via non-native interactions. The PPII conformation (see **Section 1.2.10**), which is known to be prevalent in short, unfolded alanine-based polypeptides, has been dubbed the 'killer' conformation by Blanch et al. (*214*). These authors demonstrated that the prefibrillar species of lysozyme at low pH and a temperature of 57°C, exhibits characteristics of a PPII conformation in its Raman optical activity (ROA) spectrum (*214*). The PPII conformation is well-suited to act as a

'transitional' conformation to the β-sheet-rich aggregated state, due to its flexibility and extended and fully hydrated structure, lacking in intra-molecular hydrogen bonds. Another study illustrated the population of PPII-like conformations by PLK prior to β-sheet formation (*215*). Moreover, many disease-related peptides and proteins, such as tau (Alzheimer's) (*60*), Aβ (Alzheimer's) (*123*), and α-synuclein (Parkinson's) (*60*) have been shown to adopt elements of PPII structure in their monomeric states.

In contrast to the suggested precursor role for the PPII conformation in fibril formation, other studies suggest that the presence of PPII conformations inhibits fibril formation. In the huntington protein, for instance, which is affiliated with the disease of the same name, tracts of poly-L-glumatine (PLQ) residues are often found flanked by tracts of PLP residues. Darnell et al. have provided conclusive evidence which suggest that the flanking PLP residues promote the PLQ tracts to form PPII helices, which prevent β-sheet and thus fibril formation (*216*). Moreover, these authors have demonstrated that when the central PLQ tracts exceeds a certain threshold number of glutamine residues, the flanking PLP tracts can no longer stabilize PPII helices of the central glutamine residues, which then form β-sheet conformations, ultimately leading to the formation of toxic fibrils. In addition, a more recent study by these authors on short tandem PLQ-PLP peptides, which are too short to form fibrils, instead form small soluble oligomers with a predominantly PPII structure (*217*).

2.2.10 Self-Recognition Elements (SRE's)

Short peptide motifs (usually containing 5 – 15 amino acid residues) of various amyloidogenic proteins can also yield fibrillar structures, which are nearly indistinguishable from those formed from the full-length parent protein (*218*). These findings suggest that such short segments contain all the necessary elements for fibril formation, and possibly form the core of the resulting fibrils produced from the full-length proteins. Such short peptide segments have been dubbed "self-recognition elements" (SRE's) (*218*). Some examples of SRE's from various amyloidogenic proteins are listed in **Table 10**. Studies of the fibril formation of short peptides and proteins and, in particular, those which are fragments of disease-related proteins, have provided insight into the molecular interactions which promote aggregation, as well as those which stabilize β-sheet packing of such motifs (*219-221*). In addition to the SRE's of disease-related peptides and proteins, numerous short peptides with no disease affiliation have also been reported to form amyloid-like fibrils.

Protein	Amyloidogenic Sequence(s)/Peptide(s)	Location in Native Mature Protein Sequence
α-Synuclein	(68)GAVVTGVTAVA(78) (51)GVATVA(56) (66)VGGAVVTVG(74) (86)GSIAAAT(92) (71)VTGVTAVAQKTV(82) (77) VAQKTV	68 – 78 51 – 56 66 – 74 86 – 92 71 – 82 77 - 82
β-Lactoglobulin	(11)DIQKVAGTWY(20) (101)KYLLFCMENS(110) (116)SLVCQCLVRTP(126) (146)HIRLSFN(152)	11 – 20 101 – 110 116 – 126 146 - 152
Lysozyme, Human	(56)IFQINS(61)	56 – 61
Myoglobin, Horse Heart	(1)GLSDGEWQQVLNVWGKVEADIAGHGQEVL(29)	1 - 29
Prion Protein, Human, hPrP	(113)AGAAAAGAVVGGLGG(127)	113 - 127

Table 10: Proteins of which peptide fragments have been shown to form amyloid fibrils *in vitro*. Adapted from ref(*218*).

2.2.11 Role of Aromatic Residues in Peptide and Protein Aggregation

In addition to the many extrinsic factors that can favor protein and peptide aggregation, such as elevated temperature, high concentration, and exposure of hydrophobic patches, physicochemical properties such as hydrophobicity and net charge can also stabilize the aggregated state (*222*). Moreover, a growing number of experimental studies have highlighted the importance of aromatic residues in promoting peptide and protein self-assembly and fibril formation (*199*, *223-227*). Many short sequences which are believed to be active in the fibril formation of various disease-related proteins contain aromatic residues (*228*), and the importance of aromatic interactions in inducing peptide self-assembly to form supramolecular hydrogels has also been documented (*225*, *229-230*).

The SRE of the amyloid-β (Aβ) peptide associated with Alzheimer's disease constitutes residues 16 – 21, namely KLVFFA. This segment ultimately forms the core of Aβ fibrils, but can also form fibrils with an antiparallel β-sheet arrangement. The diphenylalanine (FF) motif has also been shown to self-assemble into nanorods under very well-defined conditions. Moreover, π-π interactions resulting from the terminal Tyr residue of the active motif of the yeast prion protein Sup35, namely GNNQQNY has been suggested to promote the parallel arrangement of β-sheets, (*224*). Computational studies have indicated that the polypeptide backbone favor an antiparallel arrangement, so that the π-π interactions stabilizing the parallel arrangement of GNNQQNY is significant (*231*).

Despite the bounty of evidence pointing to the significance of aromatic interactions in stabilizing peptide and protein aggregation and fibril formation, there are reports which indicate that aromatic residues do not promote aggregation and fibril formation (*232-233*). In particular, Bemporad et al. recently suggested that the β-sheet propensity and not the aromaticity of the bulky aromatic amino acid residues promotes polypeptide aggregation (*233*).

Name of Parent Protein	Pathological or Physiological Condition	Short Active Sequence
Islet Amyloid Polypeptide	Type II Diabetes	1. FGAIL 2. TNVGSNTY* 3. QRLANFLVH*
β-Amyloid Peptide	Alzheimer's Disease	1. KLVFF (inhibitor) 2. LVFFA (inhibitor) 3. LPFFD (inhibitor) 4. KLVFFAE*
Lactadherin	Aortic Medial Amyloid	NFGSVQFV*
Gelsolin	Finnish Hereditary Amyloidosis	SFNNGDCCFILD*
Serum Amyloid A	Chronic Inflammation Amyloidosis	SFFSFLGEAFD*
PrP	Creutzfeldt-Jakob Disease (CJD)	PHGGGWGQ
Sup35p	Yeast Prion Protein	1. PQGGYQQYN 2. GNNQQNY

Table 11: Functional amyloid-related sequences that contain aromatic residues. The aromatic residues are underlined. * indicates that the minimal active fragment may actually be shorter.

2.3 Misfolding and Fibril Formation of a Short Alanine-Rich Peptide

2.3.1 Introduction

Self-assembly of polypeptides and proteins has become the subject of intense research activity for various reasons. First, an understanding of the underlying mechanism is necessary for the development of diagnostic and therapeutic strategies for a variety of diseases, including Alzheimer's, Huntington's, Creutzfeldt-Jakob's, and Parkinson's disease (*164-166*). These diseases are believed to result from misfolding and subsequent aggregation of specific proteins and peptides to form rigid insoluble fibrils. These fibrils then deposit in tissues to form insoluble plaques, e.g. amyloid plaques (*169*). Second, polypeptide/protein self-aggregation results in the formation of novel supramolecular structures (e.g., hydrogels) that are of biomedical and biotechnological relevance (*26, 234-235*).

Elucidating the mechanism of the self-assembly of peptides and proteins is of fundamental importance (*164*). The rules governing self-aggregation are still debated, however. Strong experimental evidence suggests that the aggregation process does not primarily reflect the conformational propensities of a peptide's amino acid residues (*236*). The pivotal role of aliphatic and aromatic residues in the aggregation of even small peptides suggests that aggregation is strongly driven by hydrophobic forces (*228, 236*). Gazit demonstrated the importance of aromatic residues for self-aggregation, but it is unclear whether this is due to aromaticity or the propensity of these peptides to adopt *β*-sheet-favoring conformations (*225*). In this regard, conformational disorder has been suggested to be a prerequisite for fibrillation (*164-165, 237*).

Herein is reported the fibril formation of an eight residue alanine-based oligopeptide, namely, Ac-A_4KA_2Y-NH_2 (AKY8). Alanine-rich peptides of this size are typically used as model systems for studying the unfolded state of peptides and proteins (see **Chapter 1**), and often adopt an ensemble of conformations in aqueous solution, with a predominant sampling of PPII-like conformations (*133*). Owing to the intrinsic disorder of short peptides, however, along with the currently accepted notion that β-sheet formation is general feature of the polypeptide backbone, it is not surprising that short alanine-based peptides can also form fibrils under well-defined conditions. In addition, PPII has been hypothesized to be a prerequisite for polypeptide/protein aggregation, and has even been dubbed the 'killer' conformation by Blanch et al. (*214*). This stems mainly from its flexibility and extended structural features, as well as its close location to the β basin in Ramachandran space.

Incubation in acidic media results in the spontaneous aggregation of AKY8, where an exponential fibril growth period is preceded by a lag phase, consistent with typical amyloid fibril formation (*165, 185-186*). AKY8 fibrils give rise to an enhanced couplet in the amide I' region of the corresponding VCD spectrum, and a much less intense band in the amide II' region, both of which can be utilized to probe the kinetics of fibril formation. Such enhanced amide I' VCD signals are hypothesized to originate from favorable stacking interactions between side chains of neighboring strands, as well as close inter-sheet packing (*213*). VCD can yield local secondary structural information, and can thus be considered a novel probe of fibril formation, especially when coupled with complementary methods, such as IR, UV-CD, fluorescence spectroscopies.

Promotion of AKY8 fibril formation is attributed to the C-terminal tyrosine residue of AKY8. This is suggested to be due to cation – π interactions between the lysine and tyrosine residues which stabilize the final aggregated structure. The role of aromatic residues in amyloid-fibril formation is becoming increasingly important, and the fibril formation of other short peptides has also been suggested to be at least stabilized by aromatic interactions (*220, 225*), as well as C-terminal aromatic residues (*224*).

The results presented in this chapter reinforce the link between the flexible PPII conformation and fibril formation, and offer further support for generality of amyloid fibril formation and the role of aromatic residues in promoting fibril formation, at least in short peptides. Moreover, the use of VCD as a novel probe of fibrillogensis is highlighted, which, when combined with other experimental and computational methods, may be able to yield local structural information about the fibril architecture.

2.3.2 Monomeric AKY8

Upon dissolution in deuterated water (D_2O), and in the absence of any added ions, AKY8 behaves as expected for such a short alanine-rich peptide (*48, 69, 146, 238*). The UV-CD spectrum exhibits a negative maximum at 196 nm, and a much less-pronounced shoulder at 218 nm, as shown in **Figure 24**. These spectral features are typical for short alanine-rich peptides and are indicative of an ensemble of conformations being populated with a predominance of PPII-like conformations admixed with β-strand and turn-like conformations (*146, 238*).

The temperature-dependent UV-CD spectra of AKY8 show a loss of intensity at both the negative and positive maxima at 198 and 220 nm, respectively, at higher temperatures. The difference spectrum between that measured at 85ºC and that measured at 25ºC, as shown in the inset of **Figure 24**, is indicative of an increase in β-strand conformations. Such temperature-dependent spectra have also been observed for the controversial XAO peptide (see **Figure 15**) (*238*).

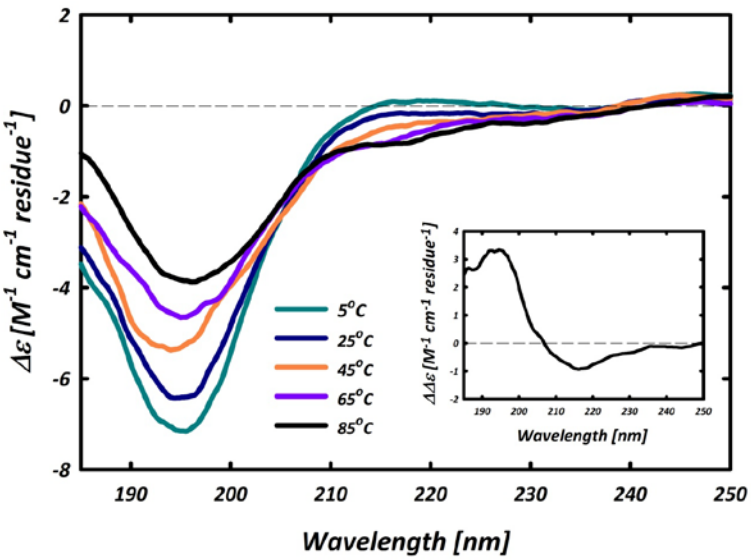


Figure 24: Temperature-dependent UV-CD spectra of 2 mM AKY8 in D_2O (pD = 1). Inset: difference spectrum between that measured at 85° and that measured at 5°C, reflecting the increase in β-strand structures at higher temperatures.

The amide I' band in the FTIR spectrum of monomeric AKY8 shows a single band at 1645 cm^{-1}, while the respective VCD spectrum in this region shows a negative couplet. Both the FTIR and VCD spectra are indicative of a predominance of PPII-like

conformations (*124*), consistent with results reported for other short alanine-rich peptides (see Chapter 1) (*59*, *124*).

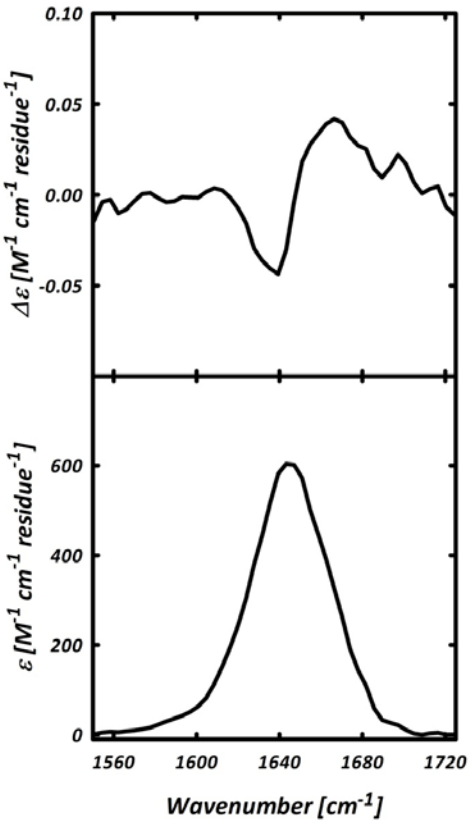


Figure 25: VCD (top panel) and FTIR (lower panel) spectra of 20 mM AKY8 in D_2O. These spectra are typical of short alanine-based peptides, and are attributed to an ensemble of conformations, with a predominate sampling of PPII-like conformations.

2.3.3 AKY8 Fibril Formation

In contrast to peptides of similar length and composition, (e.g. Ac-$(AAKA)_2$-NH_2 (*143*)), incubation of AKY8 at room temperature and in the presence of a small amount of

deuterium chloride (DCl) results in the slow and spontaneous aggregation into insoluble amyloid-like fibrils. The resulting fibrils are long, rigid, and un-branched, as indicated by atomic force microscopy (AFM) images shown in **Figure 26**. The fibrils consist of a mixture of rod-like and fibrillar structures, where the latter display a left-handed helical twist, consistent with amyloid-like fibrils (*183*). Moreover, the AKY8 fibrils are found to bind the amyloid-specific dye Congo red (*239*), as indicated by a red-shift of the visible absorption band at 498 nm, as shown in the lower panel of **Figure 27**. This, combined with the spectroscopic results and AFM images suggests that AKY8 aggregates are of the amyloid-fibril type.

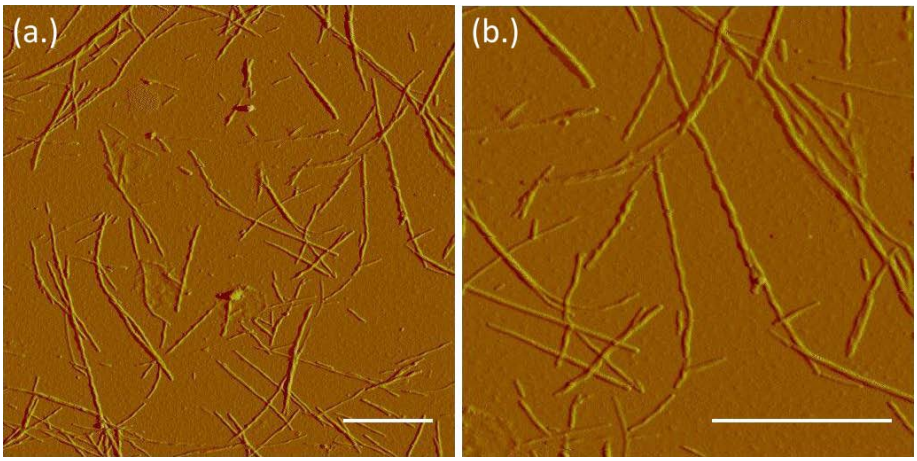


Figure 26: Amplitude AFM images of AKY8 fibrils. A mixture of twisted amyloid-like fibrils and more crystalline species can be discerned. Scale bar represents 1 μM in each image.

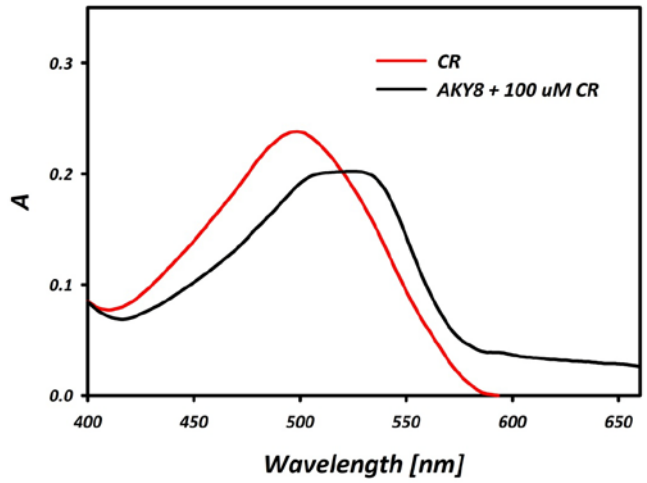


Figure 27: Visible absorption spectra of the amyloid-specific dye Congo red in the absence (red) and presence (black) of AKY8 fibrils.

Relative to the IR spectrum of the monomeric state, the amide I' band in IR spectrum of the AKY8 fibril solution is downshifted to 1616 cm^{-1}, as shown in **Figure 28**. This is indicative of a β-sheet-like structure, where the presence of a much less intense peak at ~1695 cm^{-1} is consistent with an anti-parallel sheet arrangement (*103, 240*). The broad IR band centered at ~1650 cm^{-1} in the AKY8 fibril solution is blue-shifted by ~7 cm^{-1} with respect to the amide I' band of the monomer, and coincides with the amide I' bands in the isotropic and anisotropic Raman spectra of AKY8, which are shown in **Figure 29**. This suggests that both IR and Raman bands should be assigned to a distorted β-sheet conformation, in which the two Raman active modes become IR active. That is, in an ideal anti-parallel β-sheet with 4 amide I modes, two of the modes are IR active and the other two are Raman active (*103*). The Raman active bands in an

ideal anti-parallel β-sheet occur at wavenumber positions just in between those which are IR active (*103*). Symmetry lowering via a sheet distortion results in the Raman active modes becoming IR active.

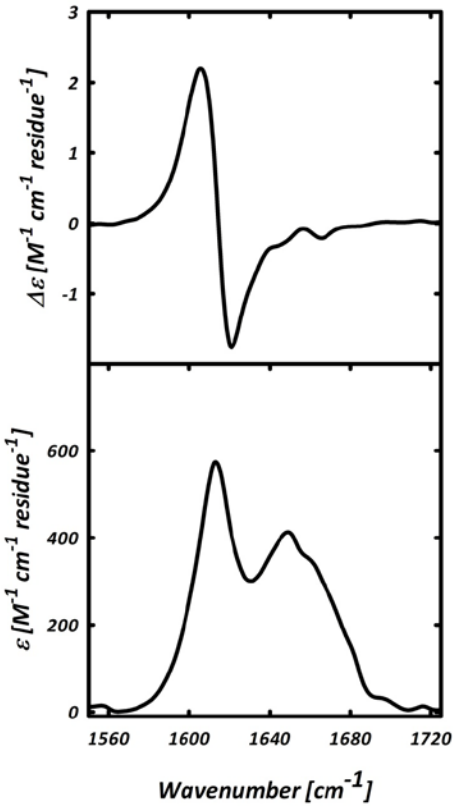


Figure 28: VCD (top panel) and FTIR (bottom panel) of AKY8 fibril (20 mM) solution, resulting from the overnight incubation at room temperature after the addition of a small amount of DCl.

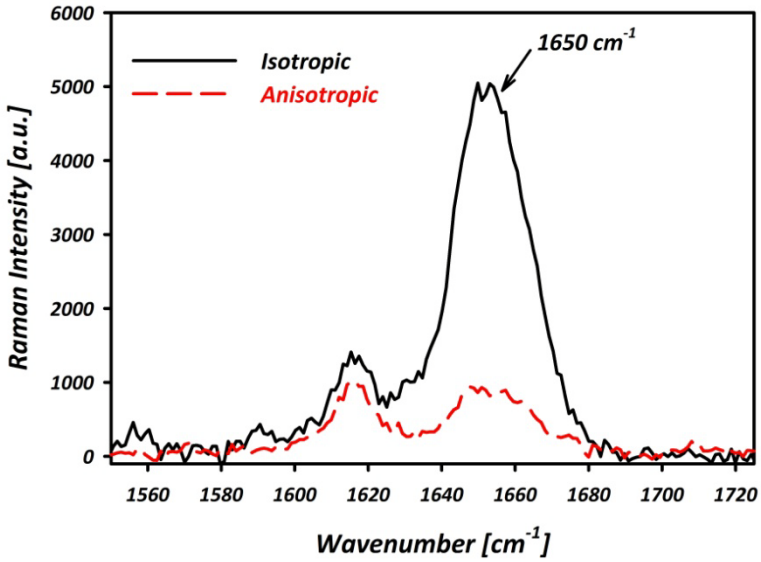


Figure 29: Amide I' profile in the isotropic and anisotropic Raman spectra of the AKY8 fibril solution.

Interestingly, the VCD spectrum of the AKY8 fibril solution shows an enhanced couplet in the amide I' region, with an inflection point coinciding with the peak in the amide I' peak IR spectrum (**Figure 28**). The VCD couplet is nearly 2 orders of magnitude larger than signals observed for typical peptides and proteins. Enhanced VCD signals of this magnitude were recently reported for fibrils of lysozyme and insulin, where it was suggested that the intensity enhancement could be the result of either inter-sheet coupling, fibril braiding, or fibril twisting (*213*). The origin of the enhancement has yet to be confirmed, however.

It should be noted that the less intense, high-frequency shoulders of the enhanced VCD couplet of the AKY8 fibril solution coincide with peaks in the corresponding IR

band profile. These spectral features suggest a heterogeneous population of β-sheets of various lengths (*240*).

2.3.4 VCD as a probe of fibril formation

To illustrate the potential of VCD as a means to probe fibril formation kinetics, both the IR and VCD spectra were monitored as a function of time for different concentrations of AKY8. A small amount of DCl was added to each peptide solution, which was then immediately placed into the spectrometer for measurement. After a period of time (lag phase), features in the amide I' and amide II' regions of the VCD spectra begin to develop, which reflect the formation of β-sheet-rich fibrils. In particular, a negative band appears in the amide II' region, as shown in **Figure 30,** while the symmetric couplet observed for the AKY8 fibril solution (**Figure 28**) emerges in the amide I' region.

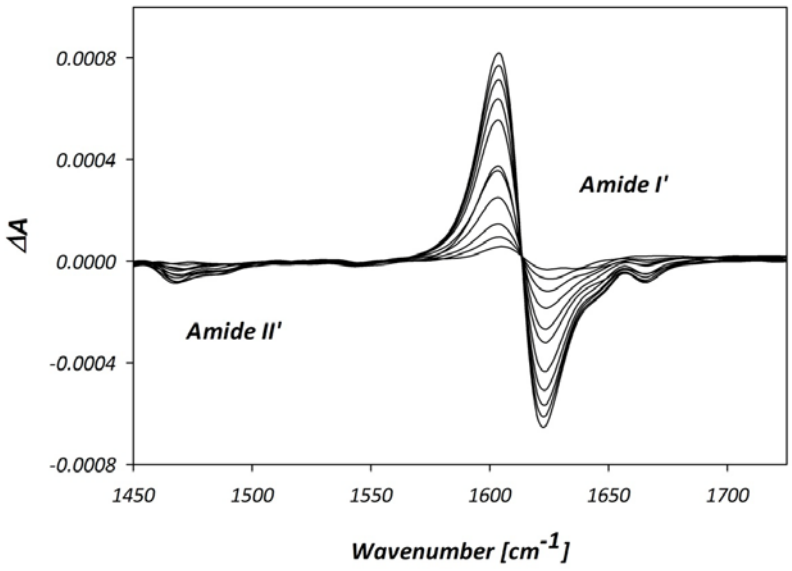


Figure 30: Time dependent VCD spectra of a 30 mM AKY8 solution, illustrating the increase in signals in the amide I' and amide II' regions attributed to AKY8 fibril formation. These are representative spectra obtained from the kinetic experiment described in the Methods section (Section 2.5.3). Starting with the spectrum displaying the least intensity in the amide I' couplet, these spectra were acquired at times of approx. 17, 18.5, 20, 22, 24, 26.5, 28.5, 30.5, 33.5, 37.5, and 44.5 hours, respectively.

Figure 31 displays the kinetics of various marker bands in the VCD spectra, which emerge during the incubation at room temperature of a 20 mM solution of AKY8 in acidic media. As shown in **Figure 31**, the various normalized marker bands, including a negative VCD maximum at 1469 cm^{-1} in the amide II' region, all overlap nearly perfectly. To our best knowledge, this is the first time that VCD signals have been observed in the amide II' region for amyloid-like fibrils, and thus suggest that this band may prove useful for delineating structural information of the resulting fibrils. Moreover, the sigmoidal shape of the kinetic curves, which include both a lag phase and an exponential growth phase, is consistent with amyloid fibril formation (*165, 185-186*).

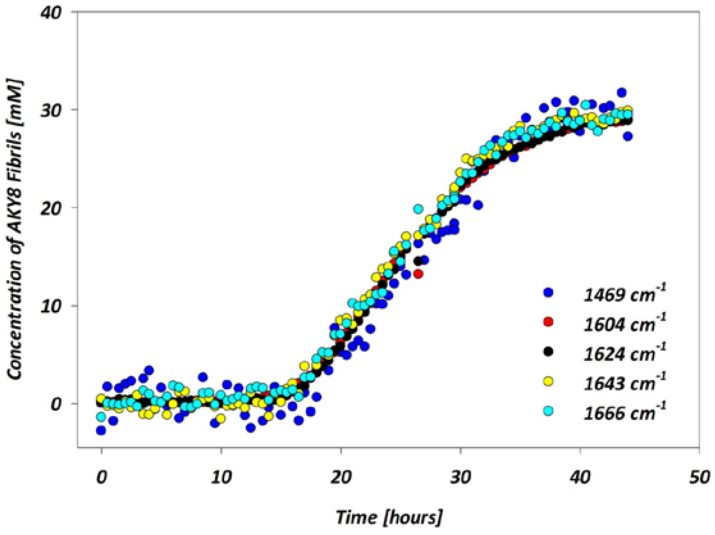


Figure 31: Kinetics of AKY8 fibril formation probed by various marker bands in the time-dependent VCD spectra.

The kinetics of fibril formation probed by the increase of the VCD couplet intensity in the amide I' region were measured for various AKY8 concentrations. Two representative sets of data are depicted in **Figure 32** for 20 and 30 mM AKY8 solutions.

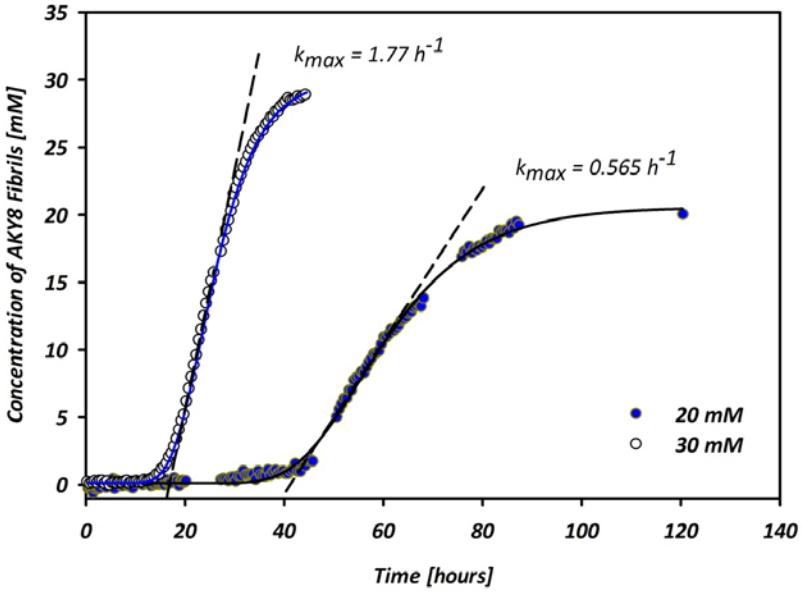


Figure 32: Kinetics of AKY8 fibril formation probed by the increase in the couplet intensity at 1604 cm^{1} in the VCD spectrum of AKY8 at two different concentrations, namely 20 and 30 mM.

As expected for aggregation via a nucleation mechanism, both the lag time and maximal rate of growth depend on the peptide concentration. In particular, the lag time decreases, while the maximal growth rate, k_{max}, increases upon increasing the peptide concentration from 20 to 30 mM. This is consistent with the aggregation of other peptides and proteins *in vitro* (*241-243*), which ultimately form un-branched, amyloid-like fibrils. Such behavior is typical of nucleated aggregation processes such as crystallization and amyloid fibril formation, where the lag phase can be eliminated by the addition of preformed aggregates to fresh solutions, a process known as seeding (*165, 185*). While the sequence of AKY8 is not linked with any particular disease, other non-

disease related protein aggregates have shown cytotoxic effects (*244*). Such systems can even form such toxic fibrils under the exact incubation conditions.

2.3.5 PPII Structure of Oligomers

During the lag phase, AKY8 maintains a PPII-like conformation, as indicated by VCD spectra acquired throughout this nucleation period shown in **Figure 33**. As mentioned in **Section 2.3.2**, a negative couplet in the amide I' region of the VCD spectrum is indicative of a predominant sampling of PPII-like conformations. The VCD spectra acquired after 20 min and 13 hours for a 20 mM AKY8 solution both reflect significant amounts of PPII-like structure is present during the nucleation phase. These results reinforce the hypothesis proposed by Blanch et al. that PPII may be the so-called 'killer' conformation, which precedes the conformational switch to β-sheet secondary structure and subsequent fibril formation (*214*).

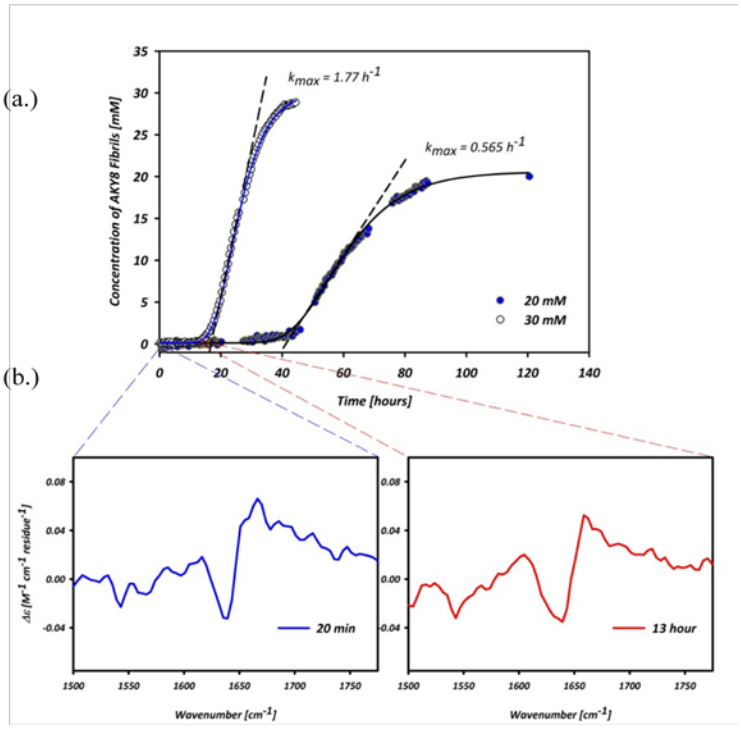


Figure 33: (a.) VCD kinetic data of 20 and 30 mM AKY8 solutions from Figure 32. (b.) Representative VCD spectra acquired during the nucleation phase of the 30 mM AKY8 solution after 20 min (blue, left) and 13 hour (red, right).

2.3.6 C-Terminal Tyrosine Promotes AKY8 Aggregation

The UV-CD spectrum of the AKY8 fibril solution is shown in **Figure 34**, along with that of monomeric AKY8. The spectrum exhibits a negative maximum at ~220 nm, indicative of a β-sheet conformation (*24*, *245*). Moreover, the presence of some CD intensity in the near-UV region suggests that the C-terminal tyrosine residue of AKY8 is in a chiral environment. Some scattering effects are also visible, attributed to negative sloping intensity extending from > 300 nm, where neither the peptide backbone nor the tyrosine side chain absorbs. It is plausible that the intensity in the region of the tyrosine

side chain absorption is due to π – π interactions between tyrosine residues of molecules arranged in a parallel β-sheet conformation. In fact, the N-terminal fragment of the yeast prion protein Sup35p, GNNQQNY, has been reported to form fibrils and microcrystals with the constituent molecules arranged as closely packed, dehydrated parallel β-sheets, stabilized by π – π interactions of neighboring strands (*200, 224*).

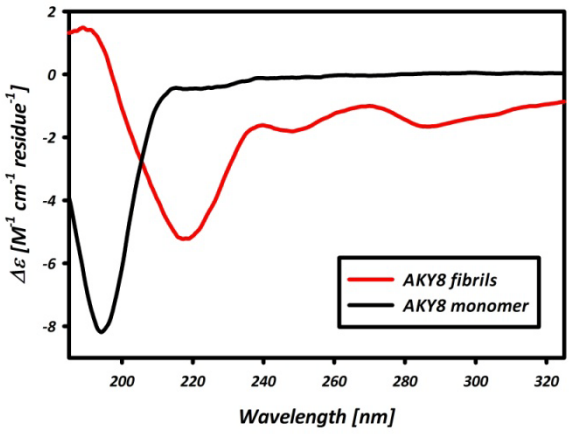


Figure 34: UV-CD spectra of a 2 mM AKY8 solution (black) and the AKY8 fibril solution (red).

Since the IR spectrum in the amide I' region clearly indicates an antiparallel conformation, interactions between the tyrosine residues in adjacent strands of the same sheet can be ruled out. Instead, it is proposed that sheet formation is caused by cation - π interactions between the lysine and tyrosine side chains that yield the following out-of-register arrangement:

AAAAKAAY

YAAKAAAA (Scheme I)

AAAAKAAY

YAAKAAAA

It has been shown that cation - π interactions, particularly those between a tyrosine and charged lysine, can be important for the stabilization of protein structures (*246*). The role of such an interaction in peptide aggregation, however, has not been proposed to date. The fibrils revealed by AFM images are consistent with the stacking of β-sheet tapes or protofibrils, which could be stabilized by π stacking between tyrosine residues or by addition cation – π interactions. First simulations of the IR and VCD profiles of the amide I band for a very simple two-dimensional model of amide I oscillators in which a one-dimensional chain represents a sheet indicate that efficient packing rather than tilting of the sheet structure might yield the observed enhancement of the VCD signal (*247*).

Intensity in the region of the UV-CD spectrum of AKY8 associated with tyrosine absorption prompted us to investigate the tyrosine-free peptide, namely Ac-(AAAAKAA)-NH_2 (AK7). In a first step, we measured the amide I' band profile of AK7 of the IR and VCD spectra in D_2O and in the absence of any added acid. **Figure 35** shows a comparison of the IR and VCD amide I' band profiles of monomeric AK7 and AKY8, prior to incubation in acidic media. Clearly, the spectra normalized with respect

to concentration, pathlength, and number of amino acid residues, indicate very little structural difference between the two peptides in the monomeric state.

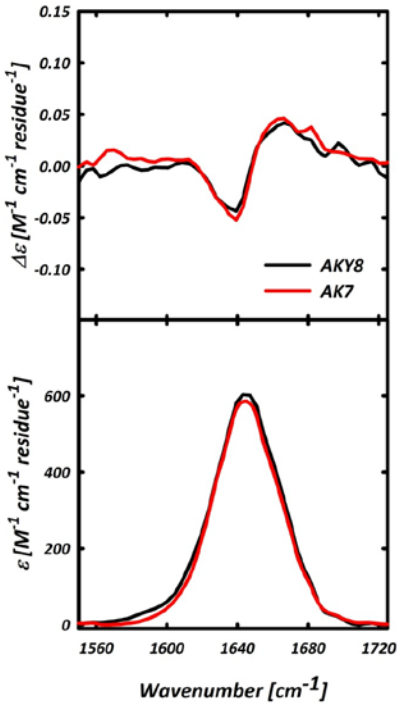


Figure 35: VCD (top) and FTIR (bottom) spectra of 30 mM AKY8 (black) and 50 mM AK7 (red) solution.

We also measured the UV-CD spectrum of AK7, which is depicted in **Figure 36**, along with the monomeric spectrum of AKY8 for comparison. These spectra offer further support for the notion that the monomeric states of both AK7 and AKY8 contain very similar structural characteristics.

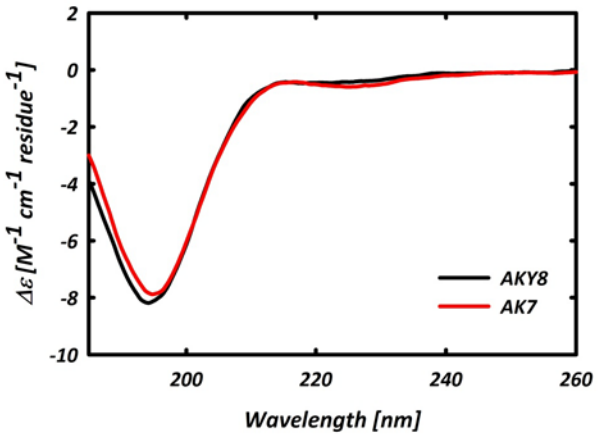


Figure 36: UV-CD spectra of monomeric solutions of AKY8 (2 mM) and AK7 (5 mM) in D_2O.

In contrast to AKY8, however, incubation of AK7 under conditions which promote the fibril formation of AKY8, does not result in any changes in the IR or VCD spectra (data not shown). That is, AK7 does not aggregate into the fibrillar structures observed for AKY8. Attempts were made to increase both the peptide and acid concentrations, as well as the incubation time, however, under no circumstance did AK7 aggregate. These results, in conjunction with the IR, VCD, and UV-CD spectra of AKY8 suggest that the C-terminal tyrosine residue of AKY8 promotes aggregation and fibril formation. These results thus provide insight into the mechanism of fibril formation by short segments of amyloidogenic proteins, such as Sup35p and Aβ, where in the former, the fibril formation of the N-terminal 7-residues motif, namely GNNQQNY has been suggested to be stabilized by π - π interactions involving the tyrosine residues from neighboring β-strands . In the latter, the SRE of Aβ comprising residues 16 – 21 (*248*), consists of an aromatic residue, F (residue 19), which is separated from a lysine residue

(residue 16), by two amino acids, and are therefore positioned in the primary sequence so as to allow for possible cation - π interactions, similar to those proposed herein for AKY8.

2.3.7 Yeast Sup35 Fragment

The fibril-forming motif from the N-terminal region of the yeast prion protein, Sup35p, namely GNNQQNY, has been shown to form fibrils with an underlying parallel β-sheet arrangement (*200*). The close packing of these 'dehydrated' sheets has been suggested to allow for aromatic interactions between the C-terminal tyrosine residues of consecutive strands (*249*). To further investigate the enhanced VCD phenomena and the significance of C-terminal aromatic residue, we measured the IR and VCD spectra in the amide I' band of the N-terminal fragment of the yeast prion protein. This fragment has been shown to form both amyloid-like fibrils and fibrous microcrystals (*200, 250*). More importantly, recent studies indicate a possible role of the C-terminal tyrosine residue in stabilizing the 'dehydrated' parallel β-sheets adopted by the fragment in fibrils (*224*). AFM images of the GNNQQNY aggregates reveal structures which more closely resemble microcrystals rather than typical amyloid-like fibrils (*200*, *224*).

The IR and VCD spectra of the amide I' band of a saturated solution of GNNQQNY in D_2O are shown in **Figure 37**. The IR shows a maximum at 1630 cm^{-1}, consistent with previous studies and indicative of a β-sheet conformation. The corresponding VCD spectrum in this region shows a negative couplet, with an inflection point coinciding with the maximum of the IR band. The negative couplet is in contrast to the positive couplet observed for AKY8 fibrils, but the magnitude of the enhanced VCD signal is even larger than that observed for AKY8 fibrils. Moreover, AFM images of the corresponding GNNQQNY solution reveal aggregates which are more crystalline, and similar to those previously reported, as shown in **Figure 38** (*200, 224*).

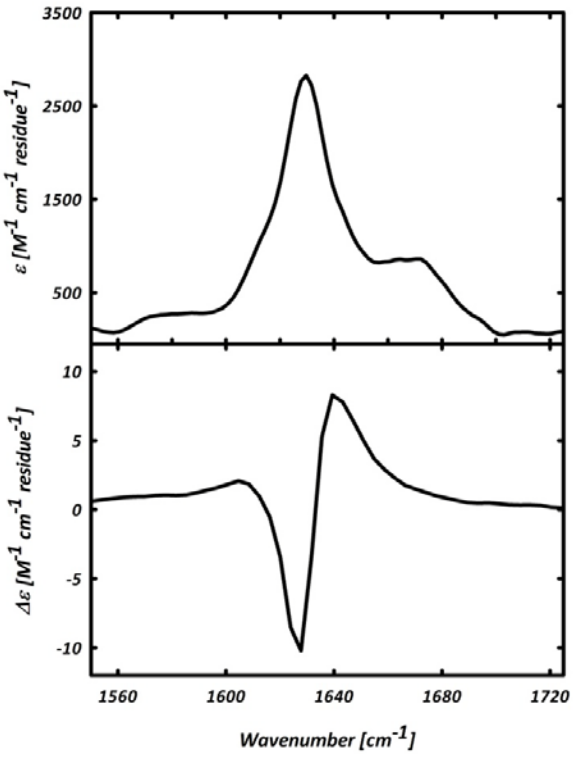


Figure 37: **Amide I' band profiles of the FTIR (top) and VCD (bottom) spectra of aggregated GNNQQNY.**

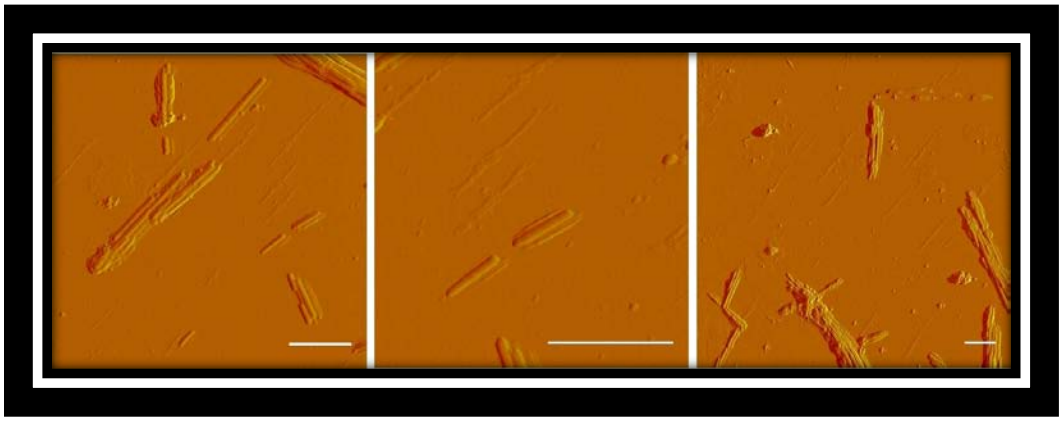


Figure 38: **AFM amplitude images of GNNQQNY. The scale bar in each panel represents 1 µM.**

The UV-CD spectrum of the GNNQQNY fibril solution exhibits a negative at ~ 220 nm, consistent with a β-sheet structure. Moreover, intensity in the near-UV region suggests that the C-terminal tyrosine residue is in a chiral environment. These results are consistent with earlier reports that GNNQQNY forms fibrils with an underlying parallel β-sheet secondary structure (*200*). The sheets are closely packed and stabilized by π - π interactions between the terminal tyrosine residues (*224*). These results might suggest that the arrangement of AKY8 fibrils is actually parallel, in contrast to the IR results reported above. In such a case, shielding of the positively charged lysine residue at acid pH by the Cl^- would be required.

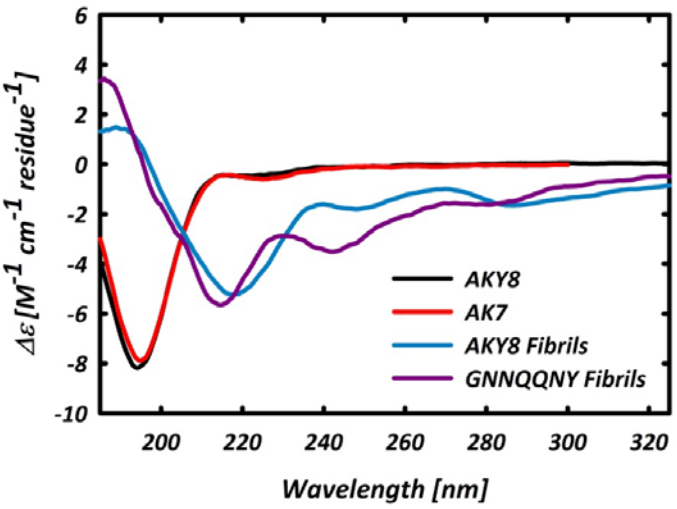


Figure 39: Comparison of the UV-CD spectra of monomeric AKY8 and AK7 with the UV-CD spectra of the fibril solution of AKY8 and GNNQQNY.

2.3.8 Conclusions

The fibril formation of a short alanine-rich peptide, namely AKY8, has been reported. AKY8 forms amyloid-like fibrils upon incubation at room temperature in the presence of a small amount of acid. Prior to aggregation, AKY8 adopts an ensemble of conformations in aqueous solution, with a predominance of PPII-like conformations, consistent with other short alanine-based peptides (e.g. see **Chapter 1**), as indicated by FTIR, VCD, and UV-CD spectra (*69, 146, 238*).

The formation of AKY8 fibrils results in a downshift of the amide I' peak in the amide I' band of the IR spectrum as well as an enhanced couplet in this region of the VCD spectrum. The increase in intensity of the amide I' couplet, as well as a smaller

band in the amide II' region of the VCD spectra can be used to probe the kinetics of fibril formation. These results suggest that VCD can complement typical aggregation studies using, e.g. IR spectroscopy, which can be employed to monitor the increase in the amide I' aggregation band. Thus, VCD spectroscopy is a convenient and novel probe of the fibril formation kinetics and structure of peptide fibrils. When coupled with computational results and simulations, the use of VCD for protein aggregation studies may be able to provide local structural information of fibrils. The progression of monomeric AKY8, containing a predominant fraction of PPII-like structure, into β-sheet-rich fibrils reinforces the hypothesis of Blanch et al. that PPII may serve as a transitional conformation between the unaggregated species and the toxic fibrillar aggregates linked to various human pathologies (*214-215*).

The C-terminal tyrosine residue plays a pivotal role in the aggregation of AKY8, as the heptameric analogue AK7 remains monomeric under conditions that promote AKY8 aggregation. Such terminal tyrosine residues have also been suggested to provide stabilization of other short fibril forming motifs contained in many disease related proteins (*224-225*). Results presented here further highlight the role of aromatic residues and aromatic interactions in promoting/stabilizing amyloid fibril formation of short peptides.

Taken together, these results reaffirm the notion that short peptides not affiliated with any known debilitating human ailments can be used as models for studying the mechanistic principles which govern the self-assembly of disease-related proteins. Moreover, these results offer further support for the hypothesis that fibril formation is an

inherent feature of the polypeptide backbone, so that a detailed understanding of the process will yield information which can be applied to developing therapeutic means for treating various seemingly unrelated diseases.

2.4 Materials and Methods

2.4.1 Peptide Synthesis and Purification

AKY8 and AK7 were synthesized in the laboratory of Sean Decatur at Mt. Holyoke College (South Hadley, MA), via Fmoc solid-phase peptide synthesis using a CEM Liberty microwave accelerated peptide synthesizer. An acetyl group was added to the N-terminus after the final deprotection. Cleavage of the peptide product from the resin was carried out in a cocktail consisting of 95% trifluoroacetic acid (TFA), 2.5% H_2O, and 2.5% isopropanol. Crude peptide was precipitated from the cleavage cocktail using cold t-butyl methyl ether, collected by centrifugation, and lyophilized. The crude peptide was purified using reverse phase high-performance liquid chromatography (RP-HPLC) on an AKTA system (Amersham Pharmacia Biotech) equipped with a C18 prep column (Vydac). The peptide was eluted using a linear gradient of acetonitrile, with 0.85% TFA. Electrospray mass spectrometry (performed at the University of Massachusetts at Amherst Mass Spectrometry Facility) was used to confirm the peptide identity and purity. Prior to use, all samples were dialyzed in a Spectra-Por Float-A-Lyzer Dialysis Bag, with a molecular weight cut-off of 500 Da (Spectrum Labs), to remove residual trifluoroacetic acid (TFA), which overlaps with the amide I' band in the FTIR spectrum. The peptides were then lyophilized overnight, and re-dissolved in D_2O. For AKY8, the exact peptide concentration was determined from the tyrosine absorption at 275 nm, which has an extinction coefficient of 1450 M^{-1} cm^{-1}, for experiments involving the monomeric, non-aggregated state. To form the fibril solution of AKY8, the peptide was

incubated overnight at room temperatures in the presence of deuterium chloride (DCl). This procedure resulted in a gelatinous solution, containing AKY8 fibrils.

The 7-residue peptide motif from the N-terminal region of the yeast prion protein, Sup35, namely GNNQQNY, was purchased from AnaSpec (Fremont, CA), and used without further purification. The peptide was dissolved in D_2O at a concentration of 25 mg/mL, which results in aggregation of GNNQQNY into amyloid-like fibrils and microcrystals (*224, 251*).

2.4.2 Congo Red Binding

Congo red was obtained from Fluka (> 97% purity). To determine if the fibrils formed by AKY8 were of the amyloid type, a previously prepared 100 μM Congo red in D_2O solution was added to the resulting gelatinous AKY8 solution (20 mM), and the subsequent UV/visible absorption spectrum was measured. Upon binding of Congo red to amyloid fibrils, a red-shift of the Congo red absorption band at ca. 498 nm is observed.

2.4.3 FTIR/VCD Spectroscopic Methods

Vibrational Circular Dichroism (VCD) and FTIR spectra were measured with a ChiralIR^{TM} spectrometer with a single PEM from BioTools (Jupiter, FL). The peptide solutions were placed in a 20 um CaF_2 BioCell^{TM} obtained from BioTools. The VCD and IR spectra of the unaggregated AKY8 and AK7 peptide were collected using the ChiralIR

software implemented in GRAMS/AI software v. 7.00 (Thermo Galactic), for 324 and 36 minutes respectively, to improve signal-to-noise ratio, while the VCD and IR spectra of the supernatant fibril solutions of both AKY8 and GNNQQNY were obtained for 18 and 2 minutes, respectively. Both VCD and IR spectra were collected using 8 cm^{-1} resolution.

Kinetic VCD and IR experiments were carried out using the features of the ChiralIR™ software, which allowed for arrayed or 'blocked' experiments. For all VCD and IR kinetic experiments, 20 minute blocks were measured, where, for each block, the VCD and IR spectra were acquired for 18 and 2 minutes, respectively.

2.4.4 Polarized Raman Spectroscopy

The polarized Raman spectra were obtained with the 514 nm excitation from a mixed-gas ArKr laser (Stabilite 2018-RM). The laser beam was directed into a RM 100 Renishaw confocal Raman microscope, and focused onto a 1.0 mm Q Silica cell with a thin glass cover slip using a x50 objective. The scattered light was filtered with a 514 nm notch filter, dispersed by a single-grating 2400 l/mm grating and imaged onto a back-thinned Wright Instrument CCD. It was polarized by a combination of a linear polarizer and a $\lambda/2$ plate. The latter rotates the y-polarized light (perpendicular to the laser polarization) into the x-direction to achieve an optimal spectrometer transmission. All spectra were recorded in the 'continuous' mode. 2 scans were measured for the AKY8 fibril solution

for both polarizations, and the spectra were averaged for each polarization direction in order to eliminate some of the noise.

2.4.5 Atomic Force Microscopy (AFM)

All AFM experiments were performed at room temperature, using a multimode atomic force microscope (Nanoscope IIIa; Digital Instruments, Santa Barbara, CA), equipped with an E-type piezoscanner. 5 µL of peptide solutions were applied to a freshly-cleaved mica surface for approx. 15 sec., at which point it was washed with 300 µL of deionized water and subsequently dried with a stream of N_2 gas. Tapping-mode imaging was carried out with a silicon probe (TESP) from Veeco (Camarillo, CA). Height and deflection images of the gelled and non-gelled samples were obtained with a scan rate of 1.0 Hz, 512 samples/line, an integral gain of 0.4, a proportional gain of 0.6, an amplitude setpoint of 0.8175, a drive amplitude of 10.53 V, and a lock-in BW of 1500Hz.

2.4.6 Ultraviolet Circular Dichroism (UV-CD) Spectroscopy

All UV-CD spectra were obtained on a Jasco J-810 spectropolarimeter, using a data pitch of 0.05 nm, a response time of 1 second, a bandwidth of 5 nm, a scan speed of 500 nm/min, and a wavelength window of either 180 – 300 or 180 – 500 nm. The instrument was purged with N_2 during the course of the measurements. For all UV-CD measurements, peptide stock solutions were diluted 10-fold with D_2O, and were placed in a 50 um Q Silica UV-Grade demountable cell (International Crystal Laboratories).

CHAPTER 3

Unexpected Aggregation and Hydrogel Formation of an Alanine-Rich Oligopeptide

3.1 Synopsis

A detailed understanding of the mechanistic principles which govern peptide and protein self-assembly is of considerable biomedical and biotechnological importance. Owing to the diversity of peptide and protein sequences which have been shown to aggregate, the ability to self-assemble is now recognized as an inherent feature of the polypeptide backbone. It is therefore of utmost importance to elucidate the rules governing the self-assembly process. *De novo* designed oligopeptides, based on alternating hydrophobic and hydrophilic sequences exhibiting repetitive complimentary charge distributions have shown the potential to self-assembly into hydrogels rich in β-sheet structure. Herein is presented the unexpected self-assembly of an alanine-rich oligopeptide containing only positively charged lysine residues, namely Ac-$(AAKA)_4$-NH_2 (AK-16). Such a sequence does not abide by typical rules which allow for peptide self-organization. AK-16 rapidly forms soluble thermodynamically unstable β-sheet structures in deuterated water at room temperature. These β-sheet-rich aggregates can be stabilized by salt addition to yield a self-supporting macroscopic hydrogel, exhibiting the ability to encapsulate and slowly release a model protein. The ability of AK-16 to self-assemble is attributed to the length of the lysine side chains, which allow for inter- and possibly intra-molecular hydrogen bonding to the peptide backbone. This one-charge-type system with the ability to adopt

unstable β-sheet structures prior to hydrogelation belongs to a novel class of self-assembling oligopeptides in which the conformational instability can be exploited to tune the viscosity of the resultant hydrogel. This study provides insight into the *de novo* design of self-assembling oligopeptide systems.

3.2 Background Information

3.2.1 Self-Assembly of Polypeptides

"Molecular self-assembly is the spontaneous association of molecules under equilibrium conditions into stable, structurally well-defined aggregates joined by non-covalent bonds (*252*)." Self-Assembly is ubiquitous throughout nature and technology, and results from the autonomous formation of non-covalent bonds, such as ionic and hydrogen bonds, as well as hydrophobic interactions (*253-254*). While such non-covalent interactions are inherently weak, they collectively govern the self-assembly process (*254*). A defining feature of all living systems is the precise organization of their constituent components (*164*). On the biomolecular level, self-assembly manifests itself as the characteristic tertiary fold of a protein, the intricate architecture of a cell membrane, and the organization of extra-cellular matrix proteins such as elastin and collagen, which are responsible for imparting elasticity and flexibility to various tissues and organs (*255*).

While the above-mentioned manifestations of biological self-assembly are necessary for the sustainability of living systems, other self-assembly processes are believed to be toxic. In particular, the formation of amyloid-like fibrils from the aggregation of misfolded or intrinsically disordered proteins and peptides (see Chapter 2) is believed to be responsible for the onset of various neurodegenerative diseases as well as the systemic amyloidoses (*4, 164-165*). Clearly, biological self-assembly results in diverse architectures displaying various and often complex functions (*253*).

3.2.2 Hydrogels

In addition to the biomedical relevance, peptide self-assembly has been exploited in recent years to create materials with inherent biofunctionality, such as hydrogels (*26*, *235*). Hydrogels are defined as a hydrophilic polymeric network exhibiting the unique ability to swell in the presence of water (*256-257*), and are thus able retain large amounts of fluids in their swollen states, which can affect different mechanical and surface properties, as well as permeability and biocompatibility (*257*).

Hydrogels were first discovered in 1960 by Wicherle and Lim (*258*), who realized the potential of poly(2-hydroxyethyl methacrylate) (PHEMA) gels as soft contact lens materials (*256-257*). All hydrogels have in common that they can trap water, and are amphiphilic, consisting of both hydrophilic and hydrophobic parts (*259*). Hydrogels are typically comprised of an underlying nanofiber scaffold, which allows for the entrapment of small clusters of water, such that convection and flow are reduced (*259*). Such properties are unique to nanoscale fibers, so that synthetic and biological polymeric gels which exhibit microcavities result in reduced surface area and entrapment of significantly less water (*259*).

While most synthetic hydrogels require the addition of chemical cross-linkers, peptide-based hydrogels rely on more physical cross-links, such as entanglements or weak intermolecular interactions (e.g. hydrogen bonds or van der Waals interactions) (*256*). This attribute makes peptide-based hydrogels appealing for biological use, owing to their inherent lower toxicity, due to the absence of any added chemical cross-linkers. The bioactivity of peptide-based hydrogels arises from their “living tissue”-like

properties. It has even been hypothesized that life began in a hydrogel environment, enabling pre-cells to form and carry out functions (*260*).

So-called responsive peptide-based hydrogels have shown sensitivity to various physiological and biological environments, such as external pH and temperature, which highlight their usefulness in various biological applications. Controlling/tailoring the primary amino acid sequence and environmental conditions may yield various novel biomaterials for specific biomedical and biotechnological uses (*256*).

3.2.3 Hydrogel Formation by Oligopeptides

De novo designed synthetic oligopeptides have shown the potential to self-assemble into macroscopic hydrogels with an underlying fibril-like network, rich in β-sheet content (*26, 183, 261-265*). Such oligopeptide-based hydrogels have been demonstrated to be useful as controlled release devices (*261, 264*) and as tissue-engineering scaffolds (*266-267*). Oligopeptides offer themselves as convenient systems to exploit the hydrogel formation as they can easily be fine-tuned to allow for specific functionality and conformational preference. In addition, large scale production of such oligo-systems has clear processing advantages over the production of longer poly-systems (*183*).

Certain general guidelines exist for the formation of oligopeptide hydrogels (*235*). In particular, oligopeptide-based hydrogels are typically formed from *de novo*-designed sequences with alternating hydrophobic and hydrophilic residues, often containing

complementary charge distributions recurring throughout (*235, 261, 264*). Charge complementarity provides for favorable electrostatic interactions which can stabilize both intermolecular and intramolecular hydrogen bonding, ultimately resulting in hydrogel formation with an underlying filamentous architecture, rich in β-sheet secondary structure (*261, 264, 268-270*).

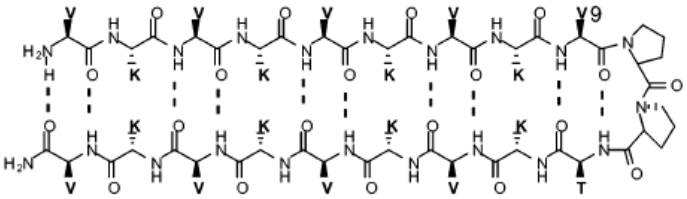
A few examples self-assembling oligopeptides that were designed under the consideration of the above-mentioned guidelines include KLD-12 ($AcN-(KLDL)_4-CNH_2$) (*268*), MAX-1 ($(VK)_4V^DPPT(KV)_4$ $-NH_2$) (*271*), RADA-16 ($AcN-(RARADADA)_2-CNH_2$) (*261, 264*), EAK-16 ($(AEAEAKAKA)_2$) (*269*), and EMK16-II ($(MEMEMKMK)_2$) (*270*). Each of these sequences possesses the ability to form a polymeric network of nanofibers under well-defined conditions, e.g. pH, temperature, etc, rich in β-sheet secondary structures. Most of these self-assembling oligopeptides form macroscopic hydrogels in the presence of added salt (*235*).

In particular, the oligopeptides KLD-12, RADA-16, EAK-16 and EMK-16 all possess charged side chains under neutral conditions, which are positioned in the sequence so as to allow for favorable electrostatic interactions which can stabilize β-sheet secondary structures. Recently, however, a comparison of the self-aggregation of EMK16-II and EAK-16 led to the conclusion that hydrophobicity rather than electrostatic interactions are the driving force of aggregation (*270*). To advance the exploitation of peptide self-assembly in the creation of novel biomaterials, it is necessary to have a complete and detailed understanding of the rules which govern the self-assembly process.

Viscoelastic properties of peptide-based hydrogels and the density of the underlying nanofiber scaffold can be controlled by varying both peptide and salt concentrations (*235, 262, 271*). Self-assembled gels typically exhibit shear thinning behavior, and are often quick in recovering their elastic properties once shearing has ceased (*271*). Shear thinning is defined as the drop in viscosity as a function of the rate of the applied shear. Such behavior results from breaking of physical cross-links, e.g. hydrogen bonds, by the applied strain (*271*).

3.2.4 MAX1

While polypeptides containing sequences of alternating hydrophobic and hydrophilic residues typically result in precipitation (*263*), the 20-residue-containing peptide $(VK)_4V^DPPT(KV)_4$ is known to form hydrogels at alkaline pH values. In this case, the short V^DPPT motif allows for the formation of a turn between the $(Val\text{-}Lys)_n$ repeating units of the amphiphilic segments, which are then driven to adopt an antiparallel β-sheet arrangement by hydrophobic interactions, as shown in **Figure 40** (*271*). While at pH values below the pKa of the lysine side chains the peptide is monomeric, the presence of the turn motif allows for the stabilization of the amphiphillicty of (VKVKVK) at pH values where the lysine side chains are deprotonated. In fact, UV-CD spectra indicate that upon increasing the pH, the peptide's secondary structure shows transition from a statistical-coil-like conformation to one with a predominance of β-sheet structure in the hydrogel phase (*245*).



MAX1: VKVKVKVKV^{D}PPTKVKVKVKV-NH_2

Figure 40: Sequence and antiparallel β-sheet formation by the *de novo* designed MAX1 oligopeptide. Figure was adapted from ref(*271*).

3.2.5 KLD-12

The oligopeptide AcN-$(KLDL)_3$-CHN_2 (KLD-12), designed by Kisiday et al., consists of alternating hydrophilic (K and D) and hydrophobic residues (L), and can form a macroscopic hydrogel under conditions providing favorable electrostatic interactions between the positively charged lysine residues and the negatively charged aspartic acid residues (*267-268*). As shown in **Figure 41**, the charged residues of a single charge type all reside on the same face of an extended β-sheet. Sheet stacking on one side via favorable electrostatic interactions between the negatively and positively charged faces of the sheet, and favorable hydrophobic interactions of the other side, results in the formation of a filamentous network, and ultimately hydrogelation.

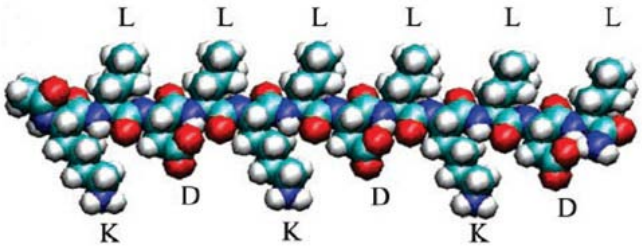


Figure 41: Representation of KLD-12 in a β-sheet conformation, illustrating the hydrophobic face containing all leucine residues and the hydrophilic face containing side chains with alternating charges (at neutral pH values). Adapted from ref(*268*).

3.2.6 RADA-16 and Controlled Delivery

The alternating complementary charges of the designed alanine-rich oligopeptide AcN-$(RARADADA)_2$-CNH_2 (RADA-16) allow for the stabilization of β-sheet structures, and the formation of a macroscopic hydrogel (*261*, *264*). These 'responsive' peptides form hydrogels under neutral conditions, which allow for favorable electrostatic interactions between the oppositely charged side chain residues, as in the case of KLD-12 (see preceding section). The resulting RADA-16 hydrogels have been shown to act as controlled release devices (*261*, *264*).

In particular, by tailoring the peptide concentration, Nagai et al. demonstrated the slow release of model dyes from the RADA-16 hydrogel matrix (*264*). The authors found that the diffusivity of the various dyes decreased with increasing hydrogel peptide concentration, and were thus able to tailor the release profile of dyes by controlling the molecular interactions between the peptide nanofiber scaffold and the diffusant (*264*). Koutsopoulos et al. incorporated various proteins of differing hydrodynamic radii into the RADA-16 hydrogels, and monitored the slow release of the proteins from the hydrogel

matrix using fluorescence correlation spectroscopy (FCS) (*261*). Using a variation of Fick's second law of diffusion, the authors were able to determine the diffusion coefficient of the various proteins encapsulated in the RADA-16 hydrogel matrix, which are reported in **Table 12**.

Protein	Mol Mass [kDa]	r_H [nm]	Diffusivity [10^{-10} m^2/s]
Lysozyme	14.3	1.9	0.50 ± 0.02
Trypsin Inhibitor	20.1	2.4	0.32 ± 0.01
BSA	66.0	5.3	0.24 ± 0.02
IgG	≈ 150.0	≈ 7.2	0.07 ± 0.01

Table 12: Molecular mass, radii of hydration (r_H) , and diffusivities of various proteins released from RADA-16, as reported in ref(*261*).

3.3 AK-16

3.3.1 Introduction

Molecular self-assembly is ubiquitous throughout nature and technology, and is governed by noncovalent interactions (*235, 252*). The onset of a variety of human diseases such as Alzheimer's, Huntington's, Parkinson's, and type II diabetes, is directly associated with the spontaneous aggregation of specific proteins into unbranched, filamentous structures, termed amyloid fibrils (*3-4, 165, 172*). Despite the sequence diversity of the disease-related proteins, the resultant fibril architectures are remarkably similar, consisting of a 'cross-β' core, with individual β-strands aligned perpendicular to the fibril axis (*173, 200*). The formation of these fibrils has recently been shown to be causative in at least one type of amyloidosis, suggesting that inhibition or prevention of fibril growth offers a therapeutic means to combat many amyloidoses (*272*). In this regard, elucidation of the mechanistic principles and contributing interactions which govern the fibril formation process is of utmost importance.

In addition to the biomedical relevance, peptide self-assembly has been exploited in recent years to create materials with inherent biofunctionality and biocompatibility. In particular, peptide hydrogels have shown potential to act as drug delivery systems (*261, 264*), tissue engineering scaffolds (*263, 269*), and to accelerate hemostasis (*262*). *De novo*-designed oligopeptide systems typically rely on a sequence of alternating hydrophobic and hydrophilic residues, often containing complimentary charge distributions alternating recurring throughout (*235*). A few examples self-assembling

oligopeptides include KLD-12 (AcN-$(KLDL)_4$-CNH_2) (*268*) (see **Section 3.2.5**), RADA-16 (AcN-$(RARADADA)_2$-CNH_2) (*261, 264*) (see **Section 3.2.6**), EAK-16 ($(AEAEAKAKA)_2$) (*269*), EMK16-II ($(MEMEMKMK)_2$) (*270*), and MAX-1 ($(VK)_4V^DPPT(KV)_4$ -NH_2) (*271*) (see **Section 3.2.4**). Each of these peptides can form a polymeric network of nanofibers rich in β-sheet secondary structures, under well-defined conditions, e.g. pH, temperature, etc.

In particular, the peptides KLD-12, RADA-16, EAK-16 and EMK-16 all posses charged side chains under neutral conditions, which are positioned in the sequence so as to allow for favorable electrostatic interactions which can stabilize β-sheet secondary structures. Recently, however, a comparison of the self-aggreagation of EMK16-II and EAK-16 led to the conclusion that hydrophobicity rather than electrostatic interactions is the driving force of aggregation (*270*). To advance the exploitation of peptide self-assembly in the creation of novel biomaterials, it is necessary to have a complete and detailed understanding of the rules which govern the self-assembly process.

In contrast to *de novo* designed oligopeptide sequences containing alternating complimentary charges, our group has recently reported the surprising ability of an alanine-based peptide, namely Ac-(AAKA)4-NH2 (AK-16) to form a hydrogel rich in β-sheet structures (*273*). This came as a surprise since alanine-rich peptides of similar length and amino acid composition typically adopt α-helical conformations in aqueous solutions (*65, 274*) and the positively charged lysine residues were expected to prevent self-aggregation. Instead, AK-16 forms thermodynamically unstable β-sheet structures, which decay into PPII- and turn-like conformations. Stabilization of the initial AK-16 β-

sheet structures via salt addition yields a macroscopic hydrogel which can slowly release encapsulated molecules. The viscosity and physicochemical properties of the AK-16 hydrogels can by tuned by varying the type of salt, the salt and peptide concentrations, and more importantly by exploiting the conformational instability of the β-sheet aggregates. These unique properties are, in part, promoted by the length of the side chain of the lysine side chains. Our results illustrate the unique properties of a novel class of self-assembling peptides, and provide insight into considerations for *de novo* design of self-assembling oligopeptides.

3.3.2 Time-Dependent UV-CD Reflects Conformational Instability of AK-16

Upon dissolution in deuterated water, AK-16 forms a thermodynamically unstable population of β-sheet-rich structures. This is observed even at concentrations as low as 7μM. These β-sheet rich aggregates organize into a filamentous network of structures, as shown in **Figure 42**. These filamentous aggregates exhibit the ability to bind the amyloid-specific dye Congo Red, as shown by the shift in the 498 absorption band of Congo red to approx. 541 nm (**Figure 43**) (*239*). This is consistent with other oligopeptide aggregates (*270*), and indicative of structures very similar to those of amyloid fibrils.

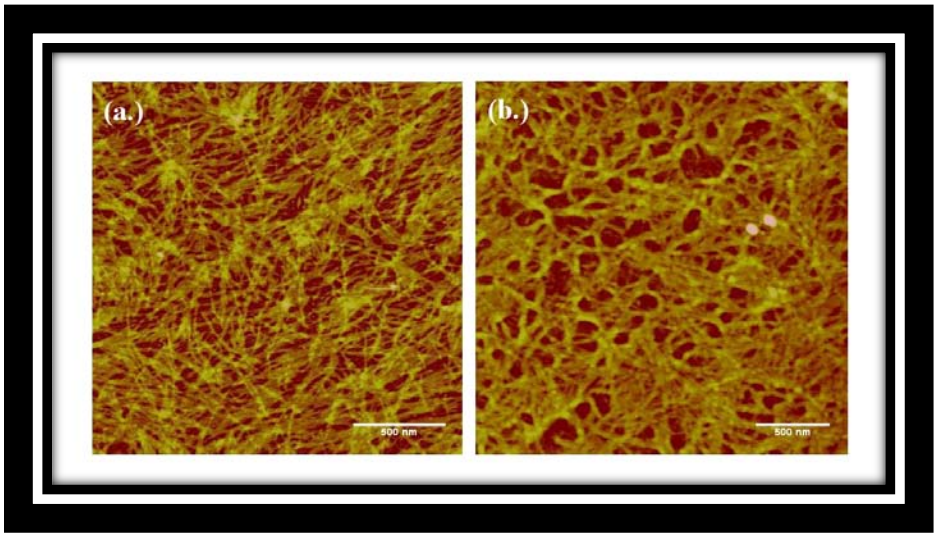


Figure 42: AFM Height Images of various samples deposited onto freshly cleaved mica. (a.) 10 mg/mL AK-16 solution, showing a filamentous network (b.) AK-16 hydrogel, prepared with AK-16 and NaCl concentrations of 10 mg/mL and 1.0 M, respectively, displaying a nanoweb-like network.

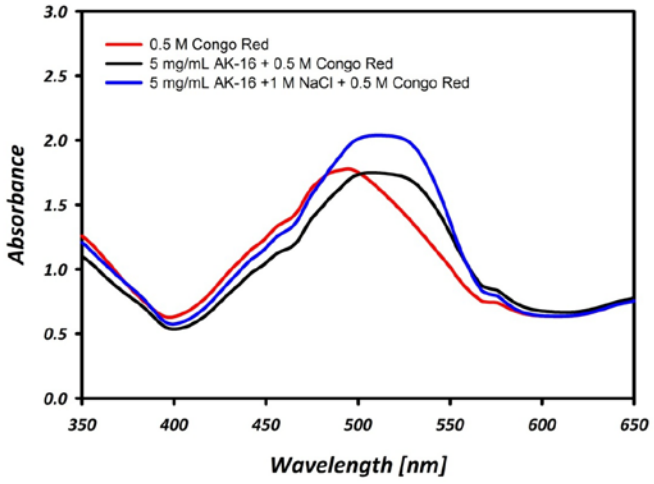


Figure 43: UV absorption spectra illustrating the binding of Congo red to AK-16 aggregates in a 5 mg/mL peptide solution in the absence (black) and presence (blue) of 1 M NaCl. Congo red (red) shows an absorption band at 498 nm. This is shifted when bound to amyloid-like fibrils, such as those formed by AK-16.

AK-16 forms filaments rich in β-sheet secondary structure, even at concentrations as low as 35 μM, where bundles of tangled fibers are observed, as shown in the AFM image in **Figure 44**.

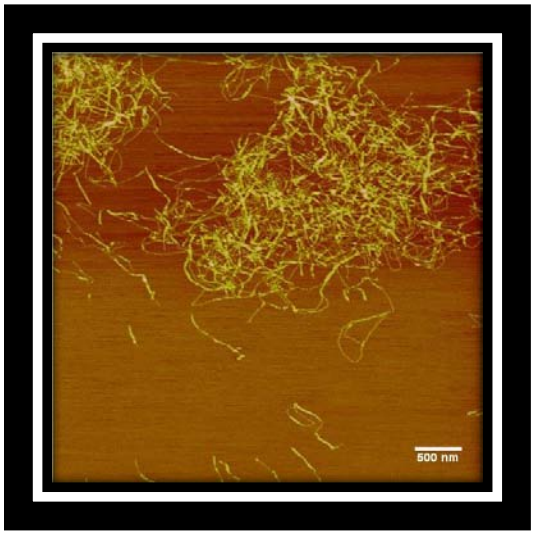


Figure 44: **AFM image (height) of 35 μM AK-16, showing a bundle of tangled filaments.**

In contrast to amyloid-like fibrils, however, the AK-16 aggregates are soluble and conformationally unstable. **Figure 45** exhibits the time-dependence of the UV-circular dichroism (UV-CD) spectra of AK-16 at various peptide concentrations. For each concentration, the spectrum measured after 5 minutes of sample preparation exhibits a negative maximum at ca. 215 nm, indicative of a β-sheet-conformation (*23-24*). As time progresses, this band slowly looses intensity at the expense of a negative maximum

appearing at approximately 198 nm and a negative shoulder near 220 nm, for concentrations ≤ 5 mg/mL (3.5 mM). These spectral changes reflect a conformational transition into more PPII- and possible turn-like conformations (*238*, *270*). The difference spectra between those measured at 300 min. and 5 min. exhibit a characteristic positive maximum at 218 nm and a negative maximum at 198 nm (see **Figure 46**), both of which are consistent with a predominance of PPII-like conformations being populated (*275*).

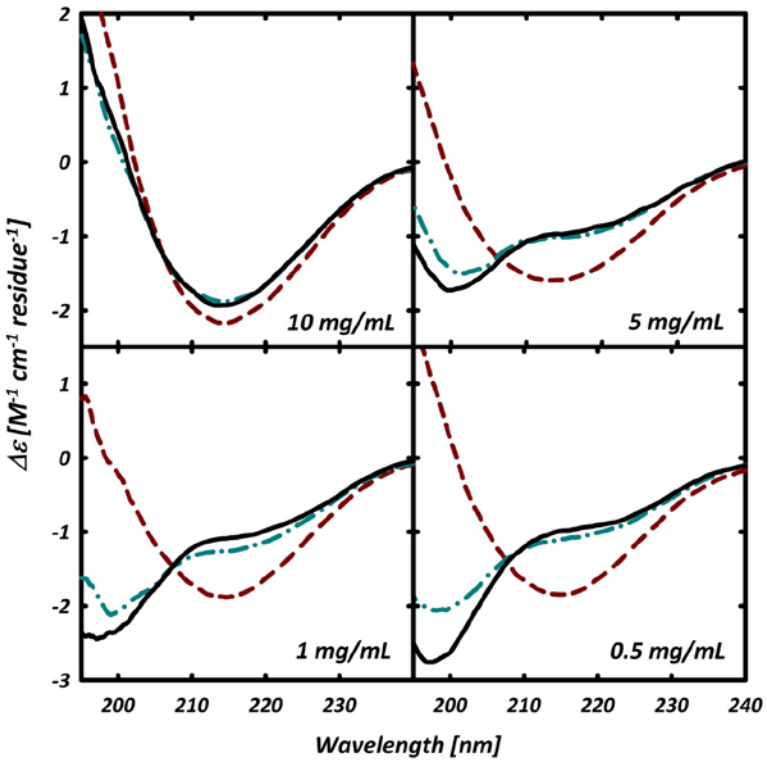


Figure 45: **Time-dependent UV-CD spectra of AK-16 at various concentrations. For clarity, only those spectra acquired at 5 (short dash), 150 (dash dot dash), and 300 (solid) minutes after dissolution in deuterated water are shown. The spectral changes reflect a decrease in β-sheet secondary structure, while the increase in the negative maximum at 198 nm reflects an increase in more flexible PPII- and turn-like conformations as described in the text.**

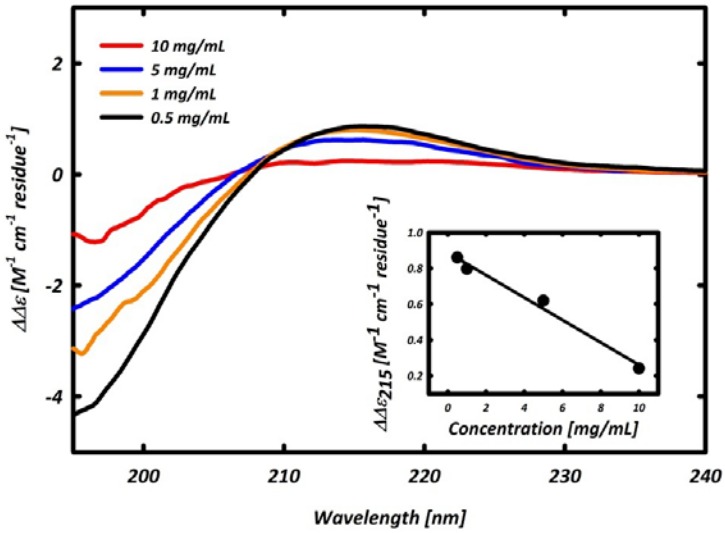


Figure 46: Time-dependent UV-CD spectra of AK-16 at various concentrations. For clarity, only those spectra acquired at 5 (short dash), 150 (dash dot dash), and 300 (solid) minutes after dissolution in deuterated water are shown. The spectral changes reflect a decrease in β-sheet secondary structure, while the increase in the negative maximum at 198 nm reflects an increase in more flexible PPII- and turn-like conformations as described in the text.

3.3.3 UV-CD Kinetics

Figure 47 displays the kinetics of the time-dependent UV-CD spectra of **Figure 47**, at three different peptide concentrations, namely 1.0, 5.0, and 10.0 mg/mL. The kinetics of the conformational transitions probed by recording the dichroism, $\Delta\varepsilon$, as a function of time at 215 nm exhibit a biphasic behavior. The fast phase is concentration independent and has a time constant of 18 ± 1 min., while the second phase proceeds an order of magnitude slower and yields a rate constant which is linearly proportional to the peptide concentration, as depicted in the inset of **Figure 47**. As will be discussed, the fast phase reflects the slow hydration of the AK-16 aggregates.

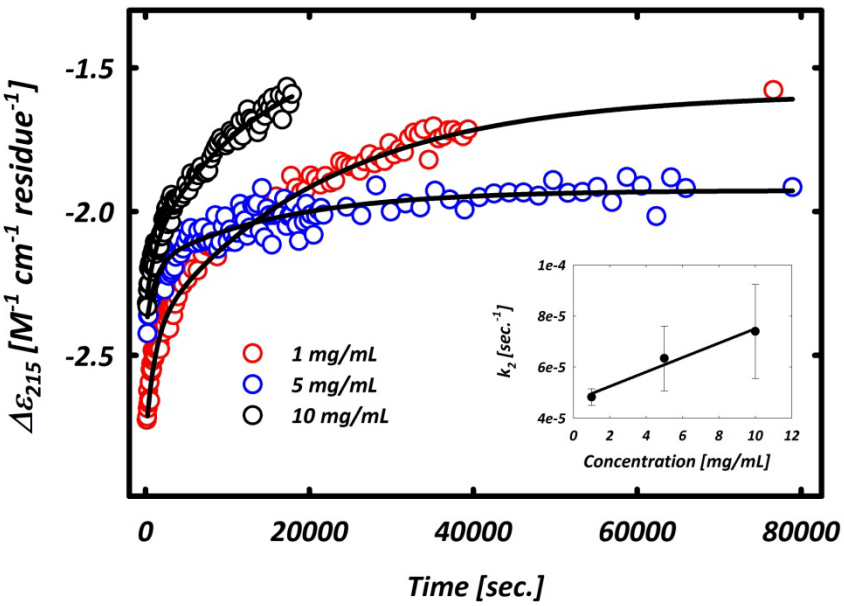


Figure 47: Kinetic Analysis of AK-16 at three different peptide concentrations, namely 1, 5, and 10 mg/mL, probed at 215 nm. The decay of the 215 nm signal reflects the loss of β-sheet structure. Analysis revealed biphasic kinetics, consisting of a concentration-independent fast phase with a time constant of 18 ± 1 min. attributed to the slow hydration kinetics of AK-16 β-sheet aggregates. The second phase proceeds an order of magnitude slower that the first, with the rate constant, k_2, exhibiting a linear correlation with the concentration of AK-16, as shown in the inset.

3.3.4 Time-Dependent FTIR Confirms Slow Hydration Kinetics of AK-16

To further elucidate structural characteristics of the unstable AK-16 aggregates, we measured the time-dependent FTIR spectra of AK-16 in the amide I' band region. The amide I' band is the most structurally sensitive IR band, resulting predominantly from backbone carbonyl stretching (*103*). The FTIR spectra of the amide I' band region of AK-16 (10 mg/mL), measured at 5 min. and 300 min. after dissolution are shown in **Figure 48**(a.). An intense band at 1616 cm^{-1} and a much weaker band at 1693 cm^{-1} are consistent with an antiparallel β-sheet conformation (*103, 240*). Further support for an

antiparallel arrangement has been confirmed by molecular dynamics (MD) simulations of AK-16 dimers and trimers (*276*).

Interestingly, the FTIR spectrum of AK-16 initially shows some intensity in the amide II band region (**Figure 48**(a)). This band typically occurs between 1500 and 1600 cm^{-1}, and results from a combination of C-N (s), C-C (s) (s = stretching mode), and N-H in-plane bending (*103*). Deuteration results in a downshift of the amide II band (termed amide II') by approx. 100 cm^{-1}, owing to the lower frequency of the N-D bending (*102*). This initial intensity in the amide II region of the initial AK-16 spectrum indicates a significant amount of amide protons involved in hydrogen bonds. These amide protons slowly exchange with deuterons of the solvent, resulting in a decrease in intensity of the amide II band and a concomitant increase in intensity of the amide II' band. A linear correlation exists between the amide II' and amide II peak intensities, suggesting a slow H-D exchange process, as show in the inset to **Figure 48**(b.) Interestingly, the kinetics of both the amide II' and amide II peak intensities **Figure 48**(b.) can be fit with a single exponential function with the time constant determined for the fast phase of the UV-CD kinetic analysis, namely 18 ± 1 min. This value is indicative of somewhat buried, immobile amide protons (*277*).

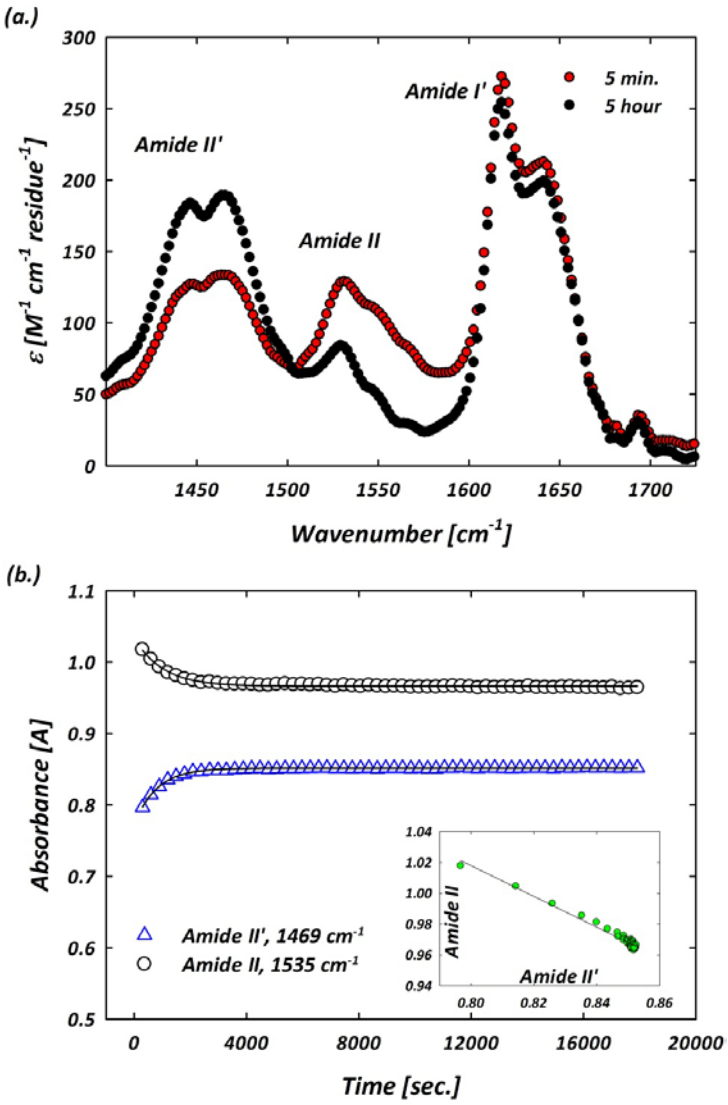


Figure 48: **Top Panel: FTIR Spectra of a 10 mg/mL AK-16 solution measured at both 5 min. (red) and 5 hour (black) after dissolving the peptide in D_2O. Bottom Panel: Kinetics of the amide II' (blue triangles) and amide II (black circles) bands probed at 1469 cm^{-1} and 1535 cm^{-1}, respectively, both of which can be fit with a single exponential function using the time constant reported for the fast phase of the UV-CD kinetic analysis, namely 18 min. Inset: a plot of amide II versus amide II' peak intensities. The negative linear correlation reflects H-D exchange of backbone amide protons.**

3.3.5 Temperature Dependence of AK-16

In addition to the time-dependent instability of the AK-16 β-sheet aggregates, these species also decay with increasing temperatures into conformations similar to those which form over time. **Figure 49** shows the temperature dependent UV-CD spectra of AK-16, which reflect a β-sheet to PPII conformational transition as the temperature is increased. Moreover, temperature-depending FTIR spectra are shown in **Figure 50**. The decrease in intensity of the band at 1616 cm^{-1} is also consistent with a loss of anti-parallel β-sheet secondary structure.

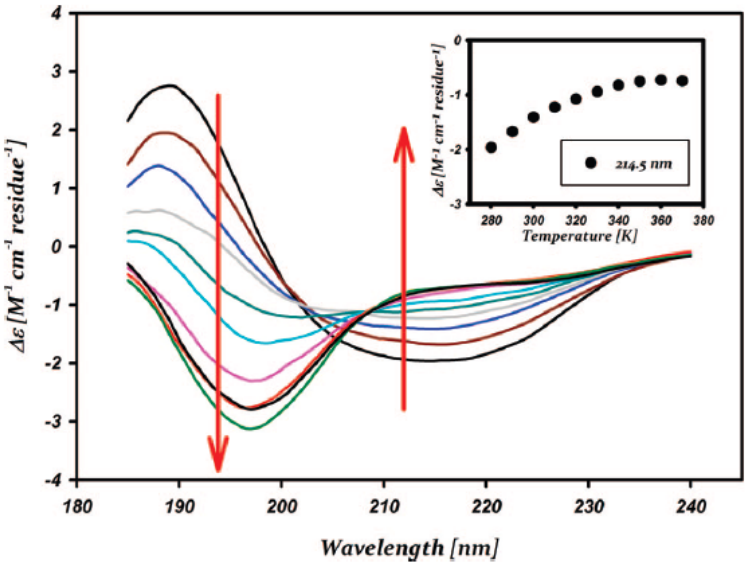


Figure 49: Temperature-dependent UV-CD spectra of a 5 mg/mL AK-16. The arrows indicate the direction of temperature increase. The inset displays the change of the molar absorptivity, Δε, at 214.5 nm as a function of temperature.

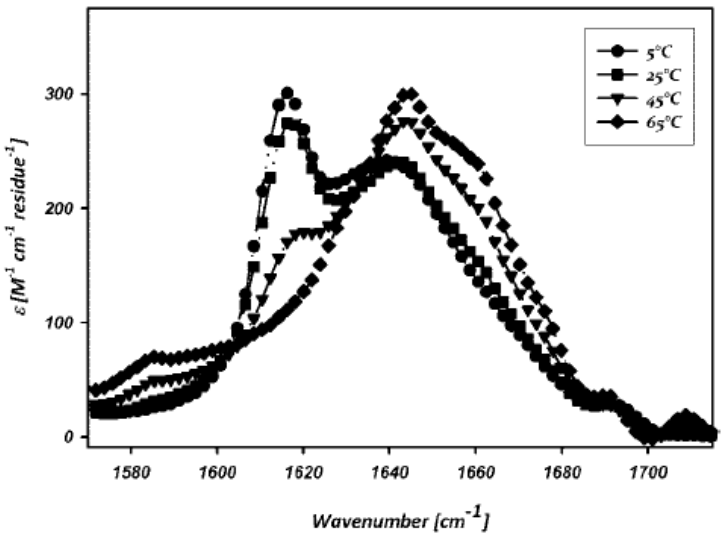


Figure 50: Temperature-dependent FTIR spectra of a 5 mg/mL AK-16 solution, reflecting the decay of the β-sheet secondary structure at increasing temperatures.

3.3.6 Side Chain Length Promotes Conformational Instability

In an attempt to elucidate the stabilizing elements of AK-16 β-sheet aggregates, we looked at various lysine side chain mutations. In particular, the four lysine residues in AK-16 were substituted by either glutamic acid (E), arginine (R), or ornithine (O), resulting in the following sequences: Ac-$(AAEA)_4$-NH_2 (AE-16), Ac-$(AARA)_4$-NH_2 (AR-16), and Ac-$(AAOA)_4$-NH_4 (AO-16), respectively. Ornithine is essentially a truncated lysine residue, with one less methylene ($–CH_2$) group. In addition, we also investigated Ac-$(AAKA)_3$-NH_2 (AK-12), to see if sequence length played a role in the aggregation propensity of AK-16. **Figure 51** shows the UV-CD spectra of AO-16, AE-16, and AK-12 in deuterated water at different pH values, all of which are at peptide concentrations of 5 mg/mL. It is clear that each of these peptides adopts a mixture of α-

helical and PPII conformations, depending on whether the side chains are uncharged or charged, respectively (*24*). This observation is consistent with homopolypeptides of amino acids containing ionizable side chains (*41*). AFM images show no indication of AO-16, AE-16, or AK-12 aggregation (not shown). In contrast to these results, AR-16 behaves in a way similar to that of AK-16 in solution. In particular, AR-16 forms soluble, thermodynamically unstable β-sheet-rich aggregates upon dissolution in deuterated water. AFM images of AR-16 depict a filamentous network similar to that which is observed for AK-16 (**Figure 52**).

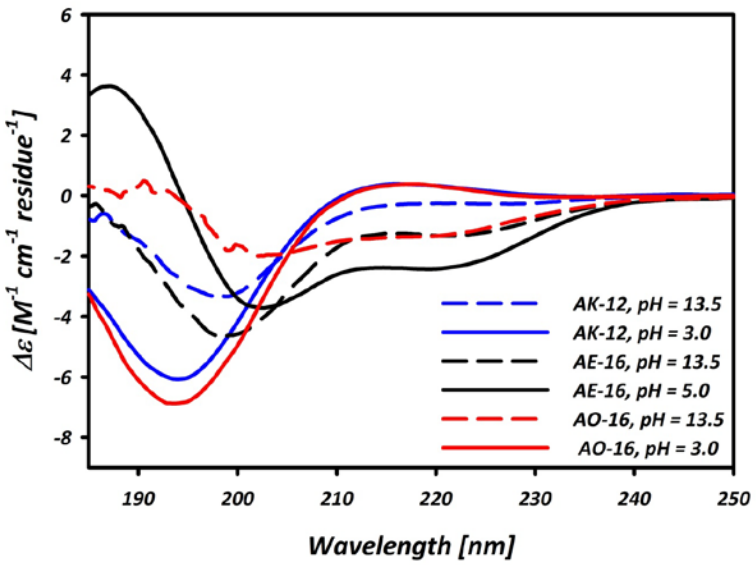


Figure 51: UV-CD spectra of AO-16 (dash dash), AE-16 (solid), and AK-12 (dash dot dash), measured at 25°C, at two indicated pH values. The negative maxima at 208 nm and 222 nm reflect α-helical conformation, while the positive and negative maxima at 220 and 198 nm, respectively, are indicative of PPII- like conformations.

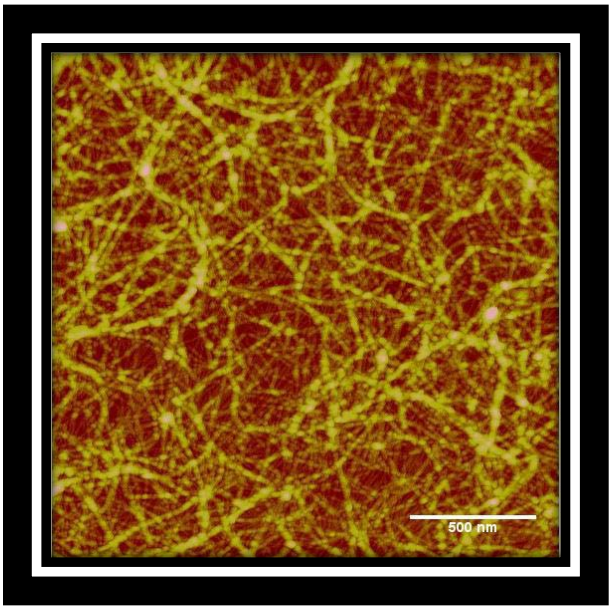


Figure 52: **AFM image (height) of 5 mg/mL AR-16.**

3.3.7 AK-16 Hydrogel Formation as a Function of Salt and Secondary Structure

Upon salt addition, AK-16 solutions instantaneously form macroscopic hydrogels at concentrations > approx. 5 mg/mL (**Figure 53**). This is consistent with hydrogel formation observed for other self-assembling oligopeptides (*235*). The AK-16 hydrogel shows a nanoweb-like architecture, as seen in the AFM image shown in **Figure 42**(b). To study the salt dependence of the physicochemical properties of the AK-16 hydrogels, we measured the the UV-CD spectra of AK-16 in the presence of various salts, including NaCl, CsCl, $MgCl_2$, and Na_3PO_4. As shown in **Figure 54**(a.), salt addition stabilizes the underlying β-sheet secondary structure, resulting in a narrowing of the negative maximum at 215 nm in the UV-CD spectrum. The valency of salt ions can have a large effect on the types of aggregates formed by oligopeptides (*270*). We found that the

position of the negative maximum in the UV-CD spectra is blue-shifted when multivalent salts are added, relative to the position in the presence of monovalent salts. Similar results have recently been observed for EMK16-II, where the blue-shift of the negative UV-CD band is attributed to the formation of more compact and amorphous aggregates observed in the presence of multivalent ions. The FTIR spectra of AK-16 as a function of the various salts (**Figure 54**(b.)), shows a downshift of the amide I' band from 1618 cm^{-1} in the absence of salt, to 1616 cm^{-1} in the presence of the monovalent salts NaCl and CsCl, at concentrations of 1 M, to 1614 cm^{-1} in the presence of the multivalent salts, $MgCl_2$ and Na_3PO_4. The amide I' downshift results from stronger coupling of the backbone carbonyl stretching vibrations (*240*), and is consistent with the formation of more compact β-sheet-rich aggregates (*270*).

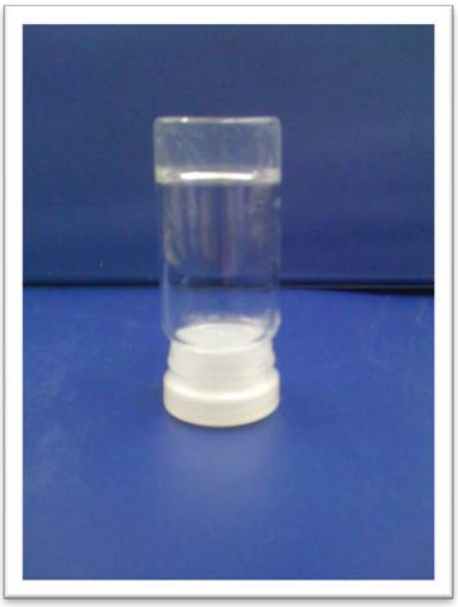


Figure 53: Self-suporting AK-16 hydrogel, prepared by adding NaCl to a 5 mg/mL solution of AK-16 to yield a 2 M NaCl AK-16 hydrogel.

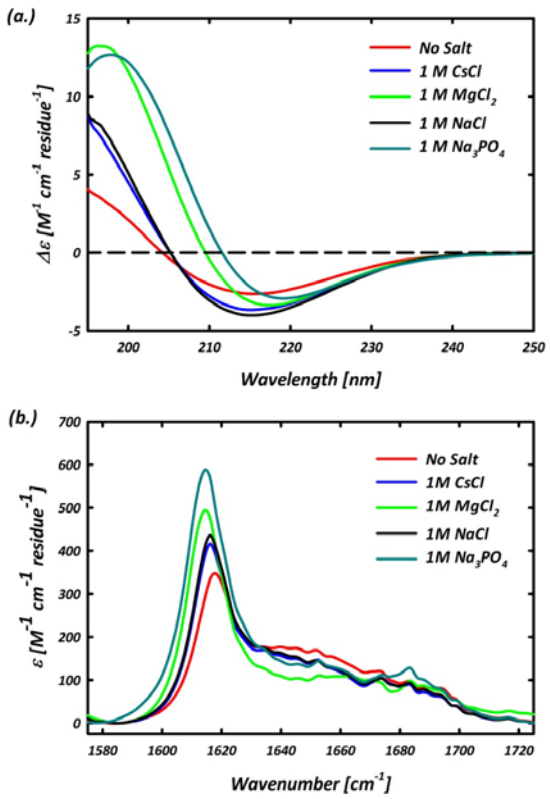


Figure 54: (a.) UV-CD and (b.) FTIR spectra of 5 mg/mL AK-16 solution with various salts (all salt concentrations were 1 M). The multivalent salts result in more compact β-sheet secondary structures, as reflected by the (a.) red-shift and narrowing of the UV-CD negative maximum (between 215 – 220 nm), and (b.) the downshift of the amide I' peak position.

To elucidate the rheological properties of AK-16 hydogels, we measured the viscosity as a function of shear rate for AK-16 hydrogels prepared with varying NaCl concentrations. At peptide concentrations of 5 and 10 mg/mL and NaCl concentrations ranging from 1 – 2 M, shear-thinning behavior is observed, in which the viscosity of the hydrogel decreases as a function of shear rate, as shown in **Figure 55**. Such behavior is a defining characteristic of oligopeptide hydrogels (see **Section 3.2.3**) (*271*). Higher salt

and peptide concentrations yield more viscous hydrogels. Based on the prerequisite of β-sheet structures being necessary for the formation of stable hydrogels, we attempted to exploit the conformational instability of the AK-16 β-sheets to essentially tune the viscosity of the resulting hydrogel. We prepared 2 separate AK-16 hydrogels from the same 5 mg/mL peptide stock solution, one of which was prepared with the fresh stock solution; the other from a portion of the solution which was allowed to incubate overnight at room temperature. As shown in **Figure 55**, the hydrogel prepared from the incubated solution exhibited a much lower viscosity than that which was prepared with a fresh sample, yet still resulted in shear-thinning behavior.

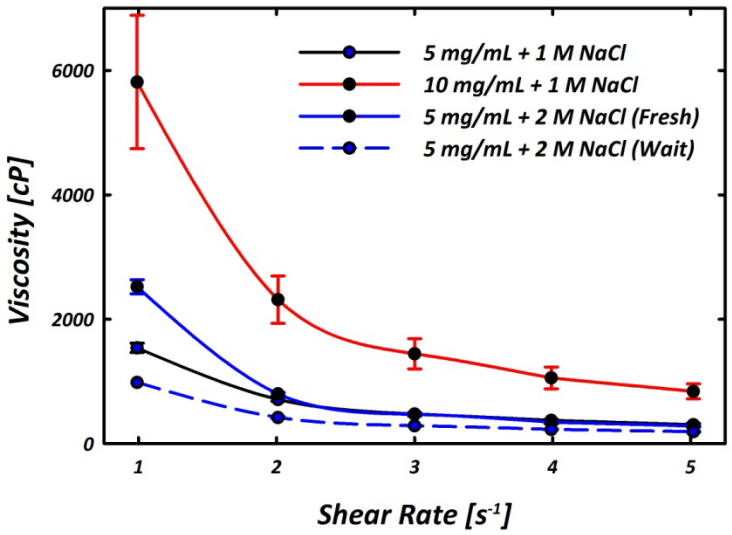


Figure 55: Viscosity of AK-16 hydrogels as a function of shear rate. Hydrogels were prepared with 5 mg/mL AK-16 and 1 M NaCl (black) and 10 mg/mL AK-16 and 1 M NaCl (red). To illustrate how β-sheet structure effects the viscosity of the resulting hydrogel, NaCl was added to both a freshly prepared portion of a 5 mg/mL AK-16 solution (blue, solid), and to a portion which was allowed to incubate overnight at room temperature (blue, dash), to yield as resulting salt concentration of 2 M. All samples were measured a total of three times. The decrease in viscosity vs. shear rate is indicative of shear-thinning behavior, and characteristic of peptide-based hydrogels.

3.3.8 Slow Release of a Model Protein from AK-16 Hydrogels

Capitalizing on the inherent biocompatibility and wide range of pH values at which the system should be operable (i.e. below the pK_a of the Lys side chain), we illustrated the application of AK-16 hydrogels as potential drug delivery systems by incorporating a model protein, namely cytochrome c, into the AK-16 matrix, and monitored the release of the protein from the hydrogel via UV absorption spectroscopy. The release profiles are shown in **Figure 56**. for different AK-16 and protein concentrations. The diffusion coefficient of cytochrome c was determined using the method of Koutsopoulos et al. (*261*), as described in **Section 3.4.5**. The diffusivity was found to be nearly independent of the protein concentration, and slightly dependent on the concentration of AK-16 in the hydrogel. In particular, we obtained a diffusion coefficient of $0.61 \pm 0.05 \times 10^{-10}$ m^2/s for cytochrome c released from the 10 mg/mL AK-16 hydrogel and $0.52 \pm 0.05 \times 10^{-10}$ m^2/s for cytochrome c released from the 5 mg/mL AK-16 hydrogel. The magnitude of these values is consistent with those of proteins of similar size and radii of hydration (*278*) released from RADA-16 hydrogels (*261*). The similarity of the UV-CD spectrum of cytochrome c released from the AK-16 hydrogel with that in aqueous solution (**Figure 57**) suggests that AK-16 hydrogels can be used to encapsulate and release model proteins while maintaining their native secondary structure.

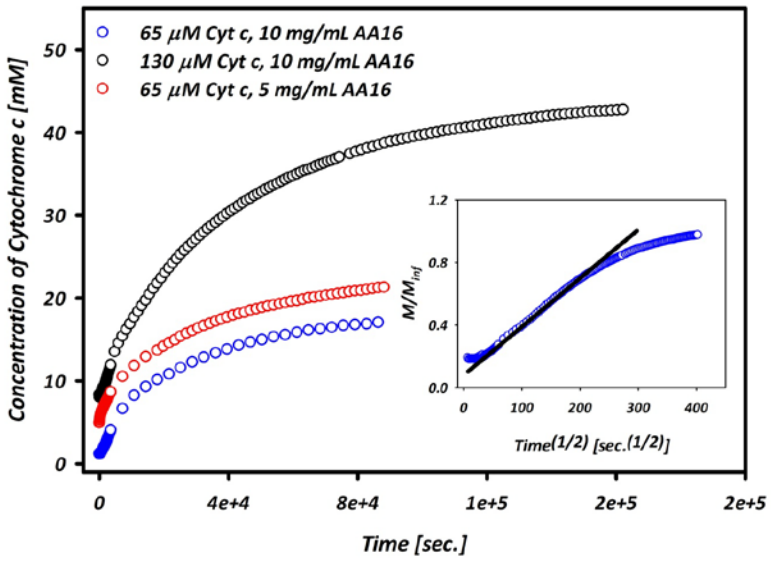


Figure 56: **Slow release experiments of cytochrome c (cyt. c) encapsulated in AK-16 hydrogels (prepared w 1 M NaCl) with the following peptide and protein concentrations: 10 mgmL AK-16 + 65 μM cyt. c (blue), 10 mg/mL AK-16 + 130 μM cyt. c (black), and 5 mg/mL AK-16 + 65 μM cyt. c (red). Cytochrome c release was monitored via the absorption at 275 nm. Inset: Plot of $\frac{[Cytc]_t}{[Cytc]_\infty}$ vs. $\sqrt{time}$. The slope of the linear portion is used to calculate the protein diffusion coefficient as described in Section 3.4.5.**

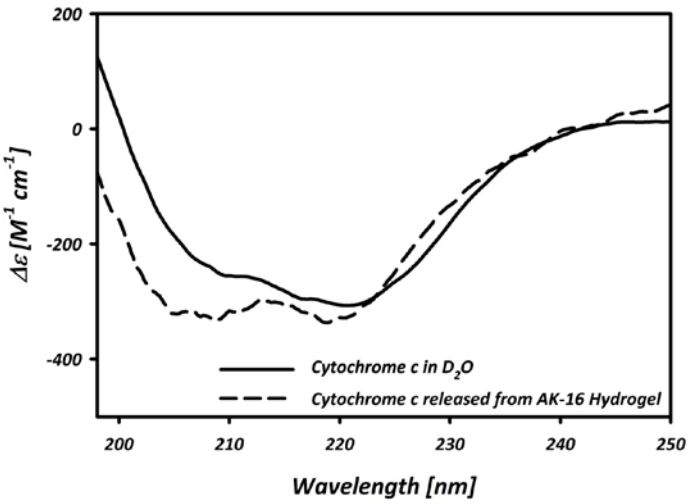


Figure 57: Comparison of UV-CD spectra of cytochrome c (equine) in D2O (solid), at a concentration of 130 µM, and after release from the AK-16 hydrogel (dashed line). Both samples had pH values between 7 and 8. The slight negative intensity bias at 208 nm in the spectrum of the cytochrome c released from the peptide hydrogel is most likely due to the presence of a small amount of AK-16, which may have become dislodged from the hydrogel matrix either during the addition of the top D2O layer (see Materials and Methods) or during the course of the release experiment.

3.3.9 Discussion

The ability of AK-16 to form aggregates rich in β-sheet secondary structures comes as a surprise since alanine-rich peptides of similar length and amino acid composition typically adopt α-helical conformations in aqueous solutions (*65*, *279*). The sequence positions of the positively charged lysine residues should prevent the formation of sheet structures via electrostatic repulsion. In contrast to these expectations, the time-dependent UV-CD and FTIR spectroscopic measurements suggest that AK-16 forms thermodynamically unstable β-sheet structures in water which spontaneously decay into

more flexible, PPII-like conformations. Biphasic kinetics are observed for the conformational decay, as probed by the UV-CD spectra, where the fast phase reflects the slow hydration of AK-16 aggregates, as confirmed by the H-D exchange process observed by the temporally resolved FTIR spectra.

These results suggest that preformed β-sheet aggregates in the solid state, which would be expected to dissociate upon dissolution in water, are somewhat stabilized and exhibit slow hydration kinetics. Indeed, the attenuated total reflectance (ATR) spectrum of AK-16 powder confirms the predominance of an antiparallel β-sheet conformation in the solid state (**Figure 58**). This behavior is different from what is generally reported in the literature about more conventional self-aggregating peptides, including both those which dissolve quickly into monomers in aqueous solution and recombine into aggregates only after a nucleation process has taken place (*280*), and those which assemble under well-defined and controlled conditions (*235*).

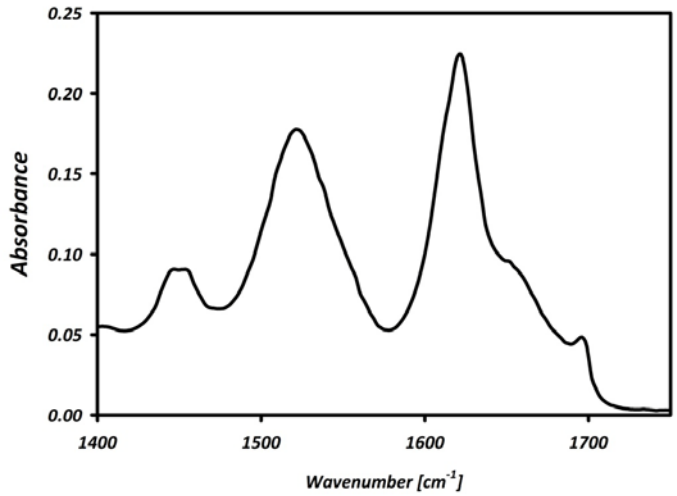


Figure 58: ATR spectrum of solid AK-16 measured on a Perkin Elmer Spectrum One FTIR Spectrometer equipped with a universal ATR sampling accessory, and acquired using the Spectrum software (v. 5.3, Perkin Elmer).

The percentage of PPII-like conformations present 5 hours after dissolution (relative to the initial 5 min. spectra) can be gauged by the intensity of the positive maximum at 220 nm (*275*). As shown in **Figure 46**, the PPII percentage exhibits a negative linear correlation with the AK-16 concentration. The decay of the AK-16 β-sheet conformation into PPII-like conformations suggests a link between these two conformations. This finding offers support for the hypothesis put forward by Blanch et al. suggesting that PPII may be the 'killer' conformation in the various 'conformational' diseases, since PPII conformations appear to precede the β-sheet conformational switch, which is concurrent with fibril formation (see **Section 2.2.9**) (*214*).

In addition to the spontaneous decay of the AK-16 β-sheets over time, these aggregated structures also decay with increasing temperature, as reflected by the temperature-dependent UV-CD and FTIR spectra shown in **Figure 49** and **Figure 50**, respectively. As the temperature is increased, both UV-CD and FTIR spectral changes reflect a melting of the β-sheet structures into more PPII-like conformations. Molecular dynamics simulations are also consistent with thermodynamically unstable AK-16 β-sheets (*276*). In particular, REMD simulations of the dimerization and trimerization of AK-16 were carried out for 64 nm, using 16 replicas with temperatures ranging from 290.5 to 527.5 K. Results of this study confirm that the population of the oligomer species drops sharply as the temperature is increased above approx. 320 K (*276*).

The ability of AR-16 to self-assemble into well-ordered structures similar to those of AK-16, together with the fact that neither AO-16 nor AE-16 aggregate under similar conditions suggests that the length of the side chains are defining features of $(AAxA)_4$ β-sheet aggregates. The side chains of both Lys and Arg are of similar lengths, while Orn and Glu are slightly shorter. Owing to the similar functionality of the Lys and Orn side chains, this finding provides strong evidence that inter- and possibly intra-molecular hydrogen bonds are initially formed between the Lys side chains and the peptide backbone, which can stabilize β-sheet structures. Without the addition of stabilizing elements, e.g. salt, electrostatic repulsion of the Lys/Arg side chains results in the spontaneous decay of the β-sheet structures, with the concurrent hydration of the backbone, until a slow-approaching equilibrium is reached. β-sheet stabilization and thus hydrogel formation results from the shielding of the lysine side chain charges, which can

then allow for such an extended structure, with the lysine side chains all residing on the same face of the sheet.

Sequence length is important for self-assembly and β-sheet formation. This is evidenced by the fact that AK-12 does not aggregate under conditions similar to those promoting AK-16 self-assembly, even at concentrations five times larger than those used for AK-16 and AR-16 (not shown). It is plausible that only the two central AAKA segments of AK-16 are involved in the β-sheet formation, so that end effects dominating in the case of AK-12, prevent such behavior. In fact, the aggregation and β-sheet formation of many disease-related peptides and proteins is thought to be initiated by short motifs, or "self-recognition elements" (SRE's), which contain only a few amino acid residues (*218*, *250*).

β-sheet structure is a prerequisite for hydrogel formation by AK-16. Stabilization of AK-16 β-sheet aggregates can be achieved via the addition of various salts to yield a self-supporting hydrogel with an underlying nanoweb-like architecture. The viscosity and physicochemical properties of the hydrogel depend on the type of salt used, the salt and peptide concentrations, as well as the amount of β-sheet secondary structure present. The underlying fibrillar network of AK-16 hydrogels allows for the encapsulation and slow release of a model protein, illustrating the potential of such a system for drug delivery. Since β-sheet structure is a prerequisite for AK-16 hydrogel formation, the initial conformational instability of such structures presents AK-16 as a unique class of hydrogel-forming oligopeptides, where the conformational instability can exploited as a means to tune the viscosity and physicochemical properties of the resultant hydrogel. In

contrast to other *de novo* designed oligopeptides which rely on stabilization by opposing charges present at neutral pH values, AK-16 hydrogels should, in principle, be stable at a wide range of pH values below the pK_a of the lysine side chains.

3.3.10 Conclusions

The 16-residue alanine rich peptide Ac-$(AAKA)_4$-NH_2 was found to form thermodynamically unstable β-sheet-rich aggregates in water. AK-16 disobeys typical sequence requirements that allow for peptide self-assembly. In particular, the presence of only positively-charged lysine residues spaced every three residues in the sequence, would be expected to disfavor β-sheet formation by electrostatic repulsion. The thermodynamically unstable AK-16 β-sheet aggregates decay as a function of both time and temperature, to yield a more heterogeneous population with a predominance of PPII-like conformations.

The kinetics of the temporal decay of the β-sheet aggregates indicates a biphasic behavior where the fast phase is independent of peptide concentration, yielding a value of 18 ± 1 min. This phase is attributed to the slow hydration kinetics of pre-formed β-sheet-rich aggregates in the solid state. The aggregation propensity of AK-16 is attributed to the length of the lysine side chain, which can allow for the formation of inter- and possibly intra-molecular hydrogen bonding to the peptide backbone upon dissolution. Stabilization of the unstable AK-16 β-sheet structures via salt addition results in the formation of a hydrogel for potential biotechnological and biomedical applications. The

viscosity and physicochemical properties of the resulting hydrogel can be tuned by factors such as salt type and concentration, peptide concentration, and β-sheet content. Taken together, these results illustrate the unique conformational properties of a novel class of self-assembling oligopeptides, and offer insight into *de novo* design of self-assembling peptides for the creation of novel biomaterials.

3.4 Materials and Methods

3.4.1 Materials

Ac-$(AAKA)_4$-NH_2 (AK-16, > 98 % purity), Ac-$(AAEA)_4$-NH_2 (AE-16, > 98 % purity), Ac-$(AARA)_4$-NH_2 (AR-16, > 98 % purity), Ac-$(AAOA)_4$-NH_2 (AO-16, > 98 % purity), and Ac-$(AAKA)_3$-NH_2 were synthesized by Celtek Peptides (Nashville, TN), and used without further purification. D_2O, CsCl, Na_3PO_4, and $MgCl_2$ were purchased from Sigma. NaCl was purchased from Acros Organics. Equine cytochrome c was purchased from Sigma and used without further purification. Peptide solutions for the UV-CD and FTIR spectroscopic time-dependent measurements were prepared as follows. A pre-weighed amount of solid peptide was dissolved in D_2O, and sample was vortexed for approximately 1 min. This resulted in non-turbid, homogenous samples. The peptide solution was added to an appropriate cell, and placed into the sample chamber of the instrument. All kinetic measurements were started approximately 5 minutes after sample preparation. Samples of AK-16 with varying salt concentrations were prepared by adding a freshly prepared solution of 5 mg/mL AK-16 to a pre-weighed amount of salt.

3.4.2 UV-CD Spectroscopy

All UV-CD spectra were measured using a Jasco J-810 spectropolarimenter, using a data pitch of 0.05 nm, a response time of 1 second, a bandwidth of 5 nm, a scan speed of 500 nm/min, and a wavelength window of 180 – 300 nm. The instrument was purged with N_2

during the course of the measurements. For all UV-CD measurements, peptide stock solutions were diluted 10-fold with D_2O, and were placed in a 50 um Q Silica UV-Grade demountable cell (International Crystal Laboratories). Temperature was controlled inside the chamber using a Peltier temperature controller (Jasco).

3.4.3 FTIR Spectroscopy

FTIR spectra were measured with a ChiralIR™ spectrometer BioTools (Jupiter, FL). All peptide solutions containing various peptide and salt concentrtaions were placed in a 100 µm CaF_2 circular BioCell™ obtained from BioTools. For kinetic FITR measurements, spectra were collected at 5 minute intervals over a course of 5 hours, using 8 cm^{-1} resolution, and were solvent corrected. 3 scans were averaged at each time interval. The instrument was purged with N_2 during the course of the measurements, and all spectra were collected using the GRAMS/AI software v. 7.00 (Thermo Galactic).

3.4.4 AFM

All AFM experiments were performed at room temperature, using a multimode atomic force microscope (Nanoscope IIIa; Digital Instruments, Santa Barbara, CA), equipped with an E-type piezoscanner. Peptide samples were applied to a freshly-cleaved mica surface for approx. 15 sec., and subsequently dried with a stream of N_2 gas. Tapping-mode imaging was carried out with a silicon probe (TESP) from Veeco (Camarillo, CA).

Height and deflection images of the gelled and non-gelled samples were obtained with a scan rate of 1 Hz.

3.4.5 Controlled Release Expeirments

Cytochrome c encapsulated AK-16 hydrogels were prepared as follows. A pre-weighed amount of AK-16 was dissolved in cytochrome c in D_2O solution with a protein concentration of either 1 or 2 mg/mL. This solution was transferred to a 10 mm quartz cell (Hellma), and a saturated NaCl solution (in D_2O) was added dropwise to yield a final AK-16 of either 5 mg/mL or 10 mg/mL, and a final NaCl concentration of 1.0 M. Due to dilution with the saturated NaCl solution, this procedure yielded an encapsulated cytochrome c concentration of either 65 or 130 µM, depending on the initial protein stock solution used. 500 µL of D_2O was carefully placed on top of the AK-16 hydrogel containing cytochrome c, and the system was placed into the nitrogen-purged sample chamber of a Jasco J-810 spectropolarimeter. Cytochrome c diffusion into the upper water layer was monitored via absorbance at 275 nm, resulting from the aromatic residues in the protein. A control experiment, containing no cytochrome c, was performed to establish a baseline for the release profiles.

The diffusion coefficient of cytochrome c was determined via the method reported in ref(*261*), using Fick's second law of diffusion, given as:

$$\frac{[Cytc]_t}{[Cytc]_\infty} = \left(\frac{16 \cdot D_{app.} \cdot t}{\pi \cdot H^2}\right)^{0.5} \quad (31)$$

Where $D_{app.}$ is the apparent diffusion coefficient, t is time, and $[Cytc]_t$ and $[Cytc]_\infty$ are the concentration of cytochrome c released at time, t, and infinite time, respectively. H is the height of the hydrogel, which was determined using the dimensions of the employed cuvette and the volume of the hydrogel. $D_{app.}$ was obtained from the slope of the linear portion of a plot of $\frac{[Cytc]_t}{[Cytc]_\infty}$ vs. $t^{0.5}$.

3.4.6 Rheological Measurements

AK-16 hydrogel samples for rheological measurements were prepared by adding a freshly prepared AK-16 in D_2O solution to a pre-weighed amount of salt. Additionally, a sample of 5 mg/mL AK-16 was left to sit overnight prior to hydrogelation, to investigate the effect of conformational relaxation, and in particular, β-sheet decay on the hydrogel viscosity.

Viscosity measurements of the AK-16 hydrogel samples were measured in the laboratory of Prof. Konstantin Kornev at the School of Materials Science and Engineering, Clemson University (Clemson, SC) using a Brookfield DV-III Ultra Viscometer/Rheometer using a Wells-Brookfield CPE-42 cone spindle. The absolute viscosity of the small amount (~1 mL) of the peptide hydrogel was obtained via 'cone-plate'-like method, with a cone angle of 1.5° and a spindle radius of 2.4 cm. Data was obtained using the Rheocalc software (Brookfield Engineering Laboratories), and automatically acquired over RS232.

After zeroing the instrument, approx. 1 mL of the peptide hydrogel sample was allowed to equilibrate on the plate at room temperature for 15 – 20 minutes. The shear rate was then varied from 1 s^{-1} to 5 s^{-1} and then back to 1 s^{-1}, at intervals of 1 s^{-1}. The shear stress [Dyne/cm^2] was measured after 3 minutes at each speed to ensure that data were collected at equilibrium. Viscosity was determined via:

$$V = 100 \cdot \left(\frac{Shear\ Stress}{Shear\ Rate}\right) \tag{32}$$

where V is the viscosity [cP], and the shear stress and shear rate have units of [Dyne cm^{-2}] and [s^{-1}], respectively. A total of three trials were performed for each sample, and the results were averaged.

REFERENCES

1. Dobson, C. M., Sali, A., and Karplus, M. (1998) Protein folding: A perspective from theory and experiment, *Angew. Chem. Int. Ed. 37*, 868-893.

2. Lehninger, A., Nelson, D. L., and Cox, M. M. (2008) *Lehninger Principles of Biochemistry*, 5th ed., W. H. Freeman.

3. Stefani, M. (2008) *Protein Folding and Misfolding, Relevance to Disease, and Biological Function*, CRC Press.

4. Dobson, C. M. (2003) Protein Folding and Disease: A View From the First Horizon Symposium, *Nat. Rev. Drug Discov. 2*, 154 - 160.

5. Anfinsen, C. B. (1973) Principles that Govern the Folding of Protein Chains, *Science 181*, 223 - 230.

6. Chou, P. Y., and Fasman, G. D. (1974) Conformational Parameters for Amino Acids in Helical, β-Sheet, and Random Coil Regions Calculated from Proteins, *Biochemistry 13*, 211 - 222.

7. Chou, P. Y., and Fasman, G. D. (1974) Prediction of Protein Conformation, *Biochemistry 13*, 222 - 245.

8. Shi, Z., Chen, K., Liu, Z., and Kallenbach, N. R. (2006) Conformation of the Backbone in Unfolded Proteins, *Chem. Rev. 106*, 1877 - 1897.

9. Levinthal, C. (1969) How To Fold Graciously, In *Mossbauer Spectroscopy in Biological Systems: Proceedings of a meeting held at Allerton House* (DeBrunner, J. T. P., and Munck, E., Eds.), pp 22 - 24, University of Illinois Press, Monticello, Illinois.

10. Eaton, W. A., Muñoz, V., Thompson, P. A., Henry, E. R., and Hofrichter, J. (1998) Kinetics and Dynamics of Loops, α-Helices, β-Hairpins, and Fast-Folding Proteins, *Acc. Chem. Res. 31*, 745 - 753.

11. Snow, C. D., Nguyen, H., Pande, V. S., and Gruebele, M. (2002) Absolute Comparison of Simulated and Experimental Protein-Folding Dynamics, *Nature 420*, 102 - 106.

12. Yang, W. Y., and Gruebele, M. (2003) Folding at the Speed Limit, *Nature 423*, 193 - 197.

13. Mayor, U., Guydosh, N. R., Johnson, C. M., Grossman, J. G., Sato, S., Jas, G. S., Freund, S. N. V., Alonso, D. O. V., Daggett, V., and Fersht, A. R. (2003) The Complete Folding Pathway of a Protein from Nanoseconds to Microseconds, *Nature 421*, 863 - 867.

14. Levinthal, C. (1968) Are There Pathways for Protein Folding, *J. Chim. Physique 65*, 44 - 45.

15. Dill, K. A., and Chan, H. S. (1997) From Levinthal to pathways to funnels, *Nat. Struct. Biol. 4*, 10-19.

16. Onuchic, J. N., and Wolynes, P. G. (2004) Theory of Protein Folding, *Curr. Opin. Struct. Biol. 14*, 70 - 75.

17. Onuchic, J. N., L.-Schulten, Z., and Wolynes, P. G. (1997) Theory of Protein Folding: The Energy Landscape Perspective, *Ann. Rev. Phys. Chem. 48*, 545 - 600.

18. Baldwin, R. L. (1995) The Nature of Protein Folding Pathways: The Classical Versus the New View, *J. Biomol. NMR 5*, 103 - 109.

19. Dinner, A. R., A. Šali, L. J. S., Dobson, C. M., and Karplus, M. (2000) Understanding Protein Folding via Free-energy Surfaces from Theory and Experiment, *Trends Biochem. Sci. 25*, 331 - 339.

20. Fersht, A. R. (2000) Transition-state Structure as a Unifying Basis in Protein-folding Mechanism: Contact Order, Chain Topology, Stability, and the Extended Nucleus Mechanism, *Proc. Natl. Acad. Sci. USA 97*, 1525 - 1529.

21. Bartlett, A. I., and Radford, S. E. (2009) An Expanding Arsenal of Experimental Methods Yields an Explosion of Insights Into Protein Folding Mechanisms, *Nat. Struct. Mol. Biol. 16*, 582 - 588.

22. Ramachandran, G. N., Ramakrishnan, C., and Sasisekharan, V. (1963) Stereochemistry of Polypeptide Chain Configurations, *J. Mol. Biol. 7*, 95 - 99.

23. Kelly, S. M., Jess, T. J., and Price, N. C. (2005) How to Study Proteins by Circular Dichroism, *Biochim. Biophys. Acta 1751*, 119 - 139.

24. Woody, R. W. (1996) Theory of Circular Dichroism of Proteins, In *Circular Dichroism and the Conformational Analysis of Biomolecules* (Fasman, G. D., Ed.), pp 25 - 67, Plenum Press, New York.

25. Woody, R. W. (1992) Circular Dichroism of Unordered Polypeptides, *Adv. Biophys. Chem. 2*, 37 - 79.

26. Hamley, I. W. (2007) Peptide Fibrillization, *Angew. Chem. Int. Ed. 46*, 8128 - 8147.

27. Hunte, C., and Michel, H. (2002) Crystallisation of Membrane Proteins Mediated by Antibody Fragments, *Curr. Opin. Struct. Biol. 12*, 503 - 508.

28. Protein Data Bank, Research Collaboratory for Structural Bioinformatics.

29. Zwanzig, R., Szabo, A., and Bagchi, B. (1992) Levinthal's Paradox, *Proc. Natl. Acad. Sci. USA 89*, 20 - 22.

30. Tanford, C. (1968) Protein Denaturation, *Adv. Protein Chem. 23*, 121 - 282.

31. Chan, H. S., and Dill, K. A. (1991) Polymer Principles in Protein Structure and Stability, *Annu. Rev. Biophys. Chem. 20*, 447 - 490.

32. Millett, I. S., Doniach, S., and Plaxco, K. W. (2002) Toward a Taxonomy of the Denatured State: Small Angle Scattering Studies of Unfolded Proteins, *Adv. Prot. Chem. 62*, 241 - 262.

33. Flory, P. J. (1969) *Statistical Mechanics of Chain Molecules*, Wiley, New York, NY.

34. Zimmerman, S. S., Pottle, M. S., Némethy, G., and Scheraga, H. A. (1977) Conformational Analysis of the 20 Naturally Occuring Amino Acid Residues Using ECEPP, *Macromolecules 10*, 1 - 9.

35. Vila, J. A., Ripoll, D. R., Baldoni, H. A., and Scheraga, H. A. (2002) Unblocked Statistical-Coil Tetrapeptides and Pentapeptides in Aqueous Solution: A Theoretical Study, *J. Biomol. NMR 24*, 245 - 262.

36. Tran, H. T., Wang, X., and Pappu, R. V. (2005) Reconciling Observations of Sequence-Specific Conformational Propensities with the Generic Polymeric Behavior of Denatured Proteins, *Biochemistry 44*, 11369 - 11380.

37. Kohn, J. E., Millett, I. S., Jacob, J., Zagrovic, B., Dillon, T. M., Cingel, N., Dothager, R. S., Seifert, S., Thiyagarajan, P., Sosnick, T. R., Hasan, M. Z., Pande, V. S., Ruczinski, I., Doniach, S., and Plaxco, K. W. (2004) Random-Coil Behavior and the Dimensions of Chemically Unfolded Proteins, *Proc. Natl. Acad. Sci. USA 101*, 12491 - 12496.

38. Fitzkee, N. C., and Rose, G. D. (2004) Reassessing Random-Coil Statistics in Unfolded Proteins, *Proc. Natl. Acad. Sci. USA 101*, 12497 - 12502.

39. Ohkubo, Y. Z., and III, C. L. B. (2003) Exploring Flory's Isolated-Pair Hypothesis: Statistical Mechanics of Helix-Coil Transitions in Polyalanine and the C-Peptide From RNase A, *Proc. Natl. Acad. Sci. USA 100*, 13916 - 13921.

40. Tiffany, M. L., and Krimm, S. (1969) Circular Dichroism of the "Random" Polypeptide Chain, *Biopolymers 8*, 347 - 359.

41. Tiffany, M. L., and Krimm, S. (1968) New Chain Conformations of Poly(glutamic Acid) and Polylysine, *Biopolymers 6*, 1379 - 1382.

42. Krimm, S., and Tiffany, M. L. (1974) The Circular Dichroism Spectrum and Structure of Unordered Polypeptides and Proteins, *Isr. J. Chem. 12*, 189 - 200.

43. Woutersen, S., and Hamm, P. (2000) Structure Determination of Trialanine in Water Using Polarization Sensitive Two-Dimensional Vibrational Spectroscopy, *J. Phys. Chem. 104*, 11316 - 11320.

44. Schweitzer-Stenner, R., Eker, F., Huang, Q., and Griebenow, K. (2001) Dihedral Angles of Trialanine in D_2O Determined by Combining FTIR and Polarized Visible Raman Spectroscopy, *J. Am. Chem. Soc. 123*, 9628 - 9633.

45. Shi, Z., Olson, C. A., Rose, G. D., Baldwin, R. L., and Kallenbach, N. R. (2002) Polyproline II Structure in a Sequence of Seven Alanine Residues, *Proc. Natl. Acad. Sci. USA 99*, 9190 - 9195.

46. Shi, Z., Woody, R. W., and Kallenbach, N. R. (2002) Is Polyproline II a Major Backbone Conformation in Unfolded Proteins?, *Adv. Prot. Chem. 62*, 163 - 240.

47. Keiderling, T. A., and Xu, Q. (2002) Unfolded Peptides and Proteins Studied with Infrared Absorption and Vibrational Circular Dichroism Spectra, *Adv. Prot. Chem. 62*, 111 - 161.

48. Eker, F., Griebenow, K., Cao, X., Nafie, L. A., and Schweitzer-Stenner, R. (2004) Preferred Peptide Backbone Conformations in the Unfolded State Revealed by the Structure Analysis of Alanine-Based (AXA) Tripeptides in Aqueous Solution, *Proc. Natl. Acad. Sci. USA 101*, 10054 - 10059.

49. Eker, F., Griebenow, K., and Schweitzer-Stenner, R. (2003) Stable Conformations of Tripeptides in Aqueous Solution Studies by UV Circular Dichroism Spectroscopy, *J. Am. Chem. Soc. 125*, 8178 - 8185.

50. Rucker, A. L., and Creamer, T. P. (2002) Polyproline II Helical Structure in Protein Unfolded States: Lysine Peptides Revisited, *Protein Sci. 11*, 980 - 985.

51. Chellgren, B. W., and Creamer, T. P. (2004) Short Sequences of Non-Proline Residues Can Adopt the Polyproline II Helical Conformation, *Biochemistry 43*, 5864 - 5869.

52. Poon, C.-D., and Samulski, E. T. (2000) Do Bridging Water Molecules Dictate the Structure of a Model Dipeptide in Aqueous Solution?, *J. Am. Chem. Soc. 122*, 5642 - 5643.

53. Barron, L. D., Blanch, E. W., and Hecht, L. (2002) Unfolded Proteins Studied by Raman Optical Activity, *Adv. Prot. Chem. 62*, 51 - 90.

54. Ding, L., Chen, K., Santini, P. A., Shi, Z., and Kallenbach, N. R. (2003) The Pentapeptide GGAGG Has PII Conformation, *J. Am. Chem. Soc. 125*, 8092 - 8093.

55. Ferreon, J. C., and Hilser, V. J. (2003) The Effect of the Polyproline II (PPII) Conformation on the Denatured State Entropy, *Protein Sci. 12*, 447 - 457.

56. Rucker, A. L., Pager, C. T., Campbell, M. N., Qualls, J. E., and Creamer, T. P. (2003) Host-Guest Scale of Left-Handed Polyproline II Helix Formation, *Proteins 53*, 68 - 75.

57. Chen, K., Liu, Z., and Kallenbach, N. R. (2004) The Polyproline II Conformation in Short Alanine Peptides is Noncooperative, *Proc. Natl. Acad. Sci. USA 101*, 15352 - 15357.

58. Ferreon, J. C., and Hilser, V. J. (2004) Thermodynamics of Binding to SH3 Domains: The Energetic Impact of Polyproline II (P_{II}) Helix Formation, *Biochemistry 43*, 7787 - 7797.

59. Dukor, R. K., and Keiderling, T. A. (1991) Reassessment of the Random Coil Conformation: Vibrational CD Study of Proline Oligopeptides and Related Polypeptides, *Biopolymers 31*, 1747 - 1761.

60. Syme, C. D., Blanch, E. W., Holt, C., Jakes, R., Goedert, M., Hecht, L., and Barron, L. D. (2002) A Raman Optical Activity Study of Rheomorphism in Caseins, Synucleins and Tau, *Eur. J. Biochem. 269*, 148 - 156.

61. Serrano, L. (1995) Comparison Between the ϕ Distribution of the Amino Acids in the Protein Database and NMR Data Indicates that Amino Acids have Various ϕ Propensities in the Random Coil Conformation, *J. Mol. Biol. 254*, 322 - 333.

62. Swindells, M. B., MacArthur, M. W., and Thornton, J. M. (1995) Intrinsic ϕ,ψ Propensities of Amino Acids, Derived From the Coil Regions of Known Structures, *Nat. Struct. Biol. 2*, 596 - 603.

63. Cowan, P. M., and McGavin, S. (1955) Structure of Poly-L-Proline, *Nature 176*, 501 - 503.

64. Marqusee, S., and Baldwin, R. L. (1987) Helix Stabilization by $Glu^{-} \cdots Lys^{+}$ Salt Bridges In Short Peptides Of De novo Designed, *Proc. Natl. Acad. Sci. USA 84*, 8898-8902.

65. Marqusee, S., Robbins, V. H., and Baldwin, R. L. (1989) Unusually Stable Helix Formation in Short Alanine-Based Peptides, *Proc. Natl. Acad. Sci. USA 86*, 5286-5290.

66. Scholtz, J. M., and Baldwin, R. L. (1992) The Mechanism of α-Helix Formation by Peptides, *Ann. Rev. Biophys. Biomol. Struct. 21*, 95 - 118.

67. Lednev, I. K., Karnoup, A. S., Sparrow, M. C., and Asher, S. A. (1999) α-Helix Peptide Folding and Unfolding Activation Barriers: A Nanosecond UV Resonance Raman Study, *J. Am. Chem. Soc. 121*, 8074 - 8086.

68. McColl, I. H., Blanch, E. W., Hecht, L., Kallenbach, N. R., and Barron, L. D. (2004) Vibrational Raman Optical Activity Characterization of Poly(L-proline) II Helix in Alanine Oligopeptides, *J. Am. Chem. Soc. 126*, 5076 - 5077.

69. Schweitzer-Stenner, R., Eker, F., Griebenow, K., Cao, X., and Nafie, L. A. (2004) The Conformation of Tetraalanine in Water Determined by Polarized Raman, FT-IR, and VCD Spectroscopy, *J. Am. Chem. Soc. 126*, 2768 - 2776.

70. Dill, K. A. (1999) Polymer Principles and Protein Folding, *Protein. Sci. 8*, 1166 - 1180.

71. Lifson, S., and Roig, A. (1961) On the Theory of Helix-Coil Transition in Polypeptides, *J. Chem. Phys. 34*, 1963 - 1974.

72. Zimm, B. H., and Bragg, J. K. (1959) Theory of the Phase Transition Between Helix and Random Coil in Polypeptide Chains, *J. Chem. Phys. 31*, 526 - 535.

73. Qian, H., and Schellman, J. A. (1992) Helix-Coil Theories: A Comparative Study for Finite Length Polypeptides, *J. Phys. Chem. 96*, 3987 - 3994.

74. van Gunsteren, W. F., Bürgi, R., Peter, C., and Daura, X. (2001) The Key to Solving the Protein-Folding Problem Lies in an Accurate Description of the Denatured State, *Angew. Chem. Int. Ed. 40*, 351 - 355.

75. Betancourt, M. R., and Skolnick, J. (2004) Local Propensities and Statistical Potentials of Backbone Dihedral Angles in Proteins, *J. Mol. Biol. 342*, 635 - 649.

76. Zaman, M. H., Shen, M.-Y., Berry, R. S., Freed, K. F., and Sosnick, T. R. (2003) Investigations into Sequence and Conforamtional Dependence of Backbone Entropy, Inter-basin Dynamics and the Flory Isolated-pair Hypothesis for Peptides, *J. Mol. Biol. 331*, 693 - 711.

77. Pappu, R. V., Srinivasan, R., and Rose, G. D. (2000) The Flory Isolated-Pair Hypothesis is not Valid for Polypeptide Chains: Implications for Protein Folding, *Proc. Natl. Acad. Sci. USA 97*, 12565 - 12570.

78. Fitzkee, N. C., and Rose, G. D. (2005) Sterics and Solvation Winnow Accessible Conformational Space for Unfolded Proteins, *J. Mol. Biol. 353*, 873 - 887.

79. Eker, F., Cao, X., Nafie, L. A., and Schweitzer-Stenner, R. (2002) Tripeptides Adopt Stable Structures in Water. A Combined Polarized Visible Raman, FTIR, and VCD Spectroscopy Study, *J. Am. Chem. Soc. 124*, 14330 - 14341.

80. Jha, A. K., Colubri, A., Zaman, M. H., Koide, S., Sosnick, T. R., and Freed, K. F. (2005) Helix, Sheet, and Polyproline II Frequencies and Strong Nearest Neighbor Effects in a Restricted Coil Library, *Biochemistry 44*, 9691 - 9702.

81. Shi, Z., Chen, K., Liu, Z., Ng, A., Bracken, W. C., and Kallenbach, N. R. (2005) Polyproline II Propensities from GGXGG Peptides Reveal an Anticorrelation with β-Sheet Scales, *Proc. Natl. Acad. Sci. USA 102*, 17964 - 17968.

82. Dunker, A. K., Lawson, J. D., Brown, C. J., Williams, R. M., Romero, P., Oh, J. S., Oldfield, C. J., Campen, A. M., Ratliff, C. M., Hipps, K. W., Ausio, J., Nissen, M. S., Reeves, R., Kang, C., Kissinger, C. R., Bailey, R. W., Griswold, M. D., Chiu, W., Garner, E. C., and Obradovic, Z. (2001) Intrinsically Disordered Protein, *J. Mol. Graphics Modelling 19*, 26 - 59.

83. Creighton, T. E. (1988) On the Relevance of Non-Random Polypeptide Conformations for Protein Folding, *Biophys. Chem. 31*, 155 - 162.

84. Fiebig, K. M., Schwalbe, H., Buck, M., Smith, L. J., and Dobson, C. M. (1996) Toward a Description of the Conformations of Denatured States of Proteins. Comparison of a Random Coil Model with NMR Measurements, *J. Phys. Chem. 100*, 2661 - 2666.

85. Gnanakaran, S., and Garcia, A. E. (2005) Helix-Coil Transition of Alanine Peptides in Water: Force Field Dependence on the Folded and Unfolded Structures, *Proteins 59*, 773 - 782.

86. Zagrovic, B., Lipfert, J., Sorin, E. J., Millett, I. S., Gunsteren, W. F. v., Doniach, S., and Pande, V. S. (2005) Unusual Compactness of a Polyproline Type II Structure, *Proc. Natl. Acad. Sci. USA 102*, 11698 - 11703.

87. Makowska, J., Rodziewicz-Motowidlo, S., Baginska, K., Vila, J. A., Liwo, A., Chmurzynski, L., and Scheraga, H. A. (2006) Polyproline II Conformation is One of Many Local Conformational States and is not an Overall Conformation of Unfolded Peptides and Proteins, *Proc. Natl. Acad. Sci. USA 103*, 1744 - 1749.

88. Ramachandran, G. N., Ramakrishnan, C., and Sasisekharan, V. (1962) Stereochemistry of Polypeptide Chain Configurations, *J. Mol. Biol. 7*, 95 - 99.

89. Brant, D. A., and Flory, P. J. (1965) The Configuration of Random Polypeptide Chains II. Theory, *J. Am. Chem. Soc. 87*, 2791 - 2800.

90. Gnanakaran, S., and Garcia, A. E. (2003) Validation of an All-Atom Protein Force Field: From Dipeptides to Larger Peptides, *J. Phys. Chem. B 107*, 12555 - 12557.

91. Duan, Y., Wu, C., Chowdhury, S., Lee, M. C., Xiong, G., Zhang, W., Yang, R., Cieplak, P., Luo, R., Lee, T., Caldwell, J., Wang, J., and Kollman, P. (2003) A Point-Charge Force Field for Molecular Mechanics Simulations of Proteins Based on Condensed-Phase Quantum Mechanical Calculations, *J. Comput. Chem. 24*, 2012 - 2012.

92. Gnanakaran, S., and Hochstrasser, R. M. (2001) Conformational Preferences and Vibrational Frequency Distributions of Short Peptides in Relation to Multidimensional Infrared Spectroscopy, *J. Am. Chem. Soc. 123*, 12886 - 12898.

93. Jun, S., Becker, J. S., Yonkunas, M., Coalson, R., and Saxena, S. (2006) Unfolding of Alanine-Based Peptides Using Electron Spin Resonance Distance Measurements, *Biochemistry 45*, 11666 - 11673.

94. Tucker, M. J., Oyola, R., and Gai, F. (2005) Conformational Distribution of a 14-Residue Peptide in Solution: A Fluorescence Resonance Energy Transfer Study, *J. Phys. Chem. B 109*, 4788 - 4795.

95. Eker, F., Griebenow, K., Cao, X., Nafie, L. A., and Schweitzer-Stenner, R. (2004) Tripeptides with Ionizable Side Chains Adopt a Perturbed Polyproline II Structure in Water, *Biochemistry 43*, 613 - 621.

96. Tompa, P. (2002) Intrinsically Unstructured Proteins, *Trends Biochem. Sci. 27*, 527 - 533.

97. Vila, J. A., Ripoll, D. R., and Scheraga, H. A. (2001) Influence of Lysine Content and pH on the Stability of Alanine-Based Copolypeptides, *Biopolymers 58*, 235 - 246.

98. Dyson, H. J., and Wright, P. E. (2004) Unfolded Proteins and Protein Folding Studied by NMR, *Chem. Rev. 104*, 3607 - 3622.

99. Bandekar, J. (1992) Amide Modes and Protein Conformation, *Biochim. Biophys. Acta 1120*, 123 - 143.

100. Surewicz, W. K., and Mantsch, H. H. (1988) New Insight in Protein Secondary Structure from Resolution-Enhanced Infrared Spectra, *Biochim. Biophys. Acta 952*, 115 - 130.

101. Susi, H., and Byler, D. M. (1986) Resolution-Enhanced Fourier Transform Infrared Spectroscopy of Enzymes, *Method. Enzymol. 130*, 290 - 311.

102. Schweitzer-Stenner, R. (2006) Advances in Vibrational Spectroscopy as a Sensitive Probe of Peptide and Protein Structure: A Critical Review, *Vib. Spectrosc. 42*, 98 - 117.

103. Krimm, S., and Bandekar, K. (1986) Vibrational Spectroscopy and Conformation of Peptides, Polypeptides, and Proteins, *Adv. Protein Chem. 38*, 181 - 364.

104. Schweitzer-Stenner, R. (2001) Visible and UV-Resonance Raman Spectroscopy of Model Peptides, *J. Raman Spectrosc. 32*, 711 - 732.

105. Mirkin, N. G., and Krimm, S. (1991) *Ab Initio* Vibrational Analysis of Hydrogen-Bonded *trans*- and *cis*-*N*-Methylacetamide, *J. Am. Chem. Soc. 113*, 9742 - 9747.

106. Chen, X. G., Schweitzer-Stenner, R., Asher, S. A., Mirkin, N. G., and Krimm, S. (1995) Vibrational Assignments of *trans*-*N*-Methylacetamide and Some of Its Deuterated Isotopomers from Band Decomposition of IR, Visible, and Resonance Raman Spectra, *J. Phys. Chem. 99*, 3074-3083.

107. Miyazawa, T., Shimanouchi, T., and Mizushima, S.-I. (1956) Characteristic Infrared Bands of Monosubstituted Amides, *J. Chem. Phys. 24*, 408 - 418.

108. Measey, T., Hagarman, A., Eker, F., Griebenow, K., and Schweitzer-Stenner, R. (2005) Side Chain Dependence of Intensity and Wavenumber Position of Amide I' in IR and Visible Raman Spectra of XA and AX Dipeptides, *J. Phys. Chem. B 109*, 8195 - 8205.

109. Torii, H., and Tasumi, M. (1998) *Ab Initio* Molecular Orbital Study of the Amide I Vibrational Interactions Between the Peptide Groups in Di- and Tripeptides and Considerations on the Conformation of the Extended Helix, *J. Raman Spectrosc. 29*, 81-86.

110. Choi, J.-H., Ham, S., and Cho, M. (2003) Local Amide I Mode Frequencies and Coupling Constants in Polypeptides, *J. Phys. Chem. B 107* 9132 - 9138.

111. Hamm, P., Lim, M., and Hochstrasser, R. M. (1998) Structure of the Amide I Band of Peptides Measured by Femtosecond Nonlinear-Infrared Spectroscopy, *J. Phys. Chem. B 102*, 6123 - 6138.

112. Choi, J.-H., and Cho, M. (2004) Amide I Vibrational Circular Dichroism of Dipeptide: Conformation Dependence and Fragment Analysis, *J. Chem. Phys. 120*, 4383 - 4392.

113. Eker, F., Cao, X., Nafie, L., Huang, Q., and Schweitzer-Stenner, R. (2003) The Structure of Alanine Based Tripeptides in Water and Dimethyl Sulfoxide Probed by Vibrational Spectroscopy, *J. Phys. Chem. B 107*, 358 - 365.

114. Schweitzer-Stenner, R. (2004) Secondary Structure Analysis of Polypeptides Based on an Excitonic Coupling Model to Describe the Band Profile of Amide I' of IR, Raman, and Vibrational Circular Dichroism Spectra, *J. Phys. Chem. B 108*, 16965 - 16975.

115. Schweitzer-Stenner, R., Measey, T. J., Hagarman, A. M., and Dragomir, I. C. (2010) The Structure of Unfolded Peptides and Proteins Explored by Raman and IR Spectroscopies, In *Assessing Structures and Conformation of Intrinsically Disordered Proteins* (Longhi, S., and Uversky, V. N., Eds.), Wiley.

116. Schweitzer-Stenner, R., Eker, F., Perez, A., Griebenow, K., Cao, X., and Nafie, L. A. (2003) The Structure of Tri-Proline in Water Probed by Polarized Raman, Fourier Transform Infrared, Vibrational Circular Dichroism, and Electric Ultraviolet Circular Dichroism Spectroscopy, *Biopolymers 71*, 558 - 568.

117. Schweitzer-Stenner, R. (2002) Dihedral Angles of Tripeptides in Solution Directly Determined by Polarized Raman and FTIR Spectroscopy, *Biophys. J. 83*, 523 - 532.

118. Long, D. A. (2002) *The Raman Effect: A Unified Treatment of the Theory of Raman Scattering by Molecules*, Wiley, New York.

119. Nafie, L. A., Dukor, R. K., and Freedman, T. B. (2002) Vibrational Circular Dichroism, In *Handbook of Vibrational Spectroscopy* (Chalmers, H. M., and Griffiths, P. R., Eds.), John Wiley & Sons Ltd, Chichester.

120. Schweitzer-Stenner, R. (2009) Distribution of Conformations Sampled by the Central Amino Acid Residue in Tripeptides Inferred From Amide I Band Profiles and NMR Scalar Coupling Constants, *J. Phys. Chem. B 113*, 2922 - 2932.

121. Holzwath, G., and Chabay, I. (1972) Optical Activity of Vibrational Transitions: A Coupled Oscillator Model, *J. Chem. Phys. 57*, 1632 - 1635.

122. Condon, E. U. (1937) Theories of Optical Rotatory Power, *Rev. Mod. Phys. 9*, 432 - 457.

123. Eker, F., Griebenow, K., and Schweitzer-Stenner, R. (2004) $A\beta_{1-28}$ Fragment of the Amyloid Peptide Predominantly Adopts a Polyproline II Conformation in an Acidic Solution, *Biochemistry 43*, 6893 - 6898.

124. Measey, T., and Schweitzer-Stenner, R. (2005) Simulation of Amide I' Band Profiles of *trans* polyproline Based on an Excitonic Coupling Model, *Chem. Phys. Lett. 408*, 123 - 127.

125. Avbelj, F., and Baldwin, R. L. (2003) Role of Backbone Solvation and Electrostatics in Generating Preferred Peptide Backbone Conformations: Distributions of phi, *Proc. Natl. Acad. Sci. USA 100*, 5742 - 5747.

126. Huggins, M. L. (1943) The Structure of Fibrous Proteins, *Chem. Rev. 32*, 195 - 218.

127. Jentzen, W., Unger, E., Karvounis, G., Shelnutt, J. A., Dreybrodt, W., and Schweitzer-Stenner, R. (1996) Conformational Properties of Nickel(II) Octaethylporphyrin in Solution. 1. Resonance Excitation Profiles and Temperature Dependence of Structure-Sensitive Raman Lines, *J. Phys. Chem. 100*, 14184 - 14191.

128. Schweitzer-Stenner, R., Measey, T., Hagarman, A., Eker, F., and Griebenow, K. (2006) Salmon Calcitonin and Amyloid β: Two Peptides with Amyloidogenic Capacity Adopt Different Conformational Manifolds in Their Unfolded States, *Biochemistry 45*, 2810 - 2819.

129. Garcia, A. E. (2004) Characterization of Non-Alpha Helical Conformations in Ala Peptides, *Polymer 45*, 669 - 676.

130. Vuister, G. W., and Bax, A. (1993) Quantitative *J* Correlation: A New Approach for Measuring Homonuclear Three-Bond $J(H^NH^\alpha)$ Coupling Constants in ^{15}N-Enriched Proteins, *J. Am. Chem. Soc. 115*, 7772 - 7777.

131. Sreerama, N., and Woody, R. W. (1994) Poly(Pro)II Helices in Globular Proteins: Identification and Circular Dichroism, *Biochemistry 33*, 10022 - 10025.

132. Hagarman, A., Measey, T. J., Mathieu, D., Schwalbe, H., and Schweitzer-Stenner, R. (2010) Intrinsic Propensities of Amino Acid Residues in GxG Peptides Inferred from Amide I' Band Profiles and NMR Scalar Coupling Constants, *J. Am. Chem. Soc. 132*, 540 - 551.

133. Graf, J., Nguyen, P. H., Stock, G., and Schwalbe, H. (2007) Structure and Dynamics of the Homologous Series of Alanine Peptides: A Joint Molecular Dynamics/NMR Study, *J. Am. Chem. Soc. 129*, 1179 - 1189.

134. Hagarman, A., Measey, T., Doddasomayajula, R. S., Dragomir, I., Eker, F., Griebenow, K., and Schweitzer-Stenner, R. (2006) Conformational Analysis of XA and AX Dipeptides in Water by Electronic Circular Dichroism and ^{1}H NMR Spectroscopy, *J. Phys. Chem. B 110*, 6979 - 6986.

135. Dragomir, I. C., Measey, T. J., Hagarman, A. M., and Schweitzer-Stenner, R. (2006) Environment-Controlled Interchromophore Charge Transfer Transitions in Dipeptides Probed by UV Absorption and Electronic Circular Dichroism Spectroscopy, *J. Phys. Chem. B 110*, 13235 - 13241.

136. Vila, J. A., Baldoni, H. A., Ripoll, D. R., Ghosh, A., and Scheraga, H. A. (2004) Polyproline II Helix Conformation in a Proline-Rich Environment: A Theoretical Study, *Biophys. J. 86*, 731 - 742.

137. Woutersen, S., Pfister, R., Hamm, P., Mu, Y., Kosov, D. S., and Stock, G. (2002) Peptide Conformational Heterogeneity Revealed from Nonlinear Vibrational Spectroscopy and Molecular-Dynamics Simulations, *J. Chem. Phys. 117*, 6833 - 6840.

138. Mu, Y., and Stock, G. (2002) Conformational Dynamics of Trialanine in Water: A Molecular Dynamics Study, *J. Phys. Chem. B 106*, 5294 - 5301.

139. Cornell, W. D., Cieplak, P., Bayly, C. I., Gould, I. R., Merz Jr., K. M., Ferguson, D. M., Spellmeyer, D. C., Fox, T., Caldwell, J., and Kollman, P. A. (1995) A Second Generation Force Field for the Simulation of Proteins, Nucleic Acids, and Organic Molecules, *J. Am. Chem. Soc. 117*, 5179 - 5197.

140. Hu, H., Elstner, M., and Hermans, J. (2003) Comparison of a QM/MM Force Field and Molecular Mechanics Force Fields in Simulations of Alanine and Glycine "Dipeptides" (Ace-Ala-Nme and Ace-Gly-Nme) in Water in Relation to the Problem of Modeling the Unfolded Peptide Backbone in Solution, *Proteins 50*, 451 - 463.

141. Kentsis, A., Mezei, M., and Osman, R. (2005) Origin of the Sequence-Dependent Polyproline II Structure in Unfolded Peptides, *Proteins 61*, 769 - 776.

142. Pizzanelli, S., Monti, S., and Forte, C. (2005) Conformation and Orientation of a Tetraalanine in a Lyotropic Liquid Crystal Studied by Nuclear Magnetic Resonance, *J. Phys. Chem. B 109*, 21102 - 21109.

143. Measey, T., and Schweitzer-Stenner, R. (2006) The Conformations Adopted by the Octamer Peptide $(AAKA)_2$ in Aqueous Solution Probed by FTIR and Polarized Raman Spectroscopy, *J. Raman Spectrosc. 37*, 248 - 254.

144. Pappu, R. H., and Rose, G. D. (2002) A Simple Model for Polyproline II Structure in Unfolded States of Alanine-Based Peptides, *Protein Sci. 11*, 2437 - 2455.

145. Asher, S. A., Mikhonin, A. V., and Bykov, S. (2004) UV Raman Demonstrates that α-Helical Polyalanine Peptides Melt to Polyproline II Conformations, *J. Am. Chem. Soc. 126*, 8433 - 8440.

146. Schweitzer-Stenner, R., Measey, T., Kakalis, L., Jordan, F., Pizzanelli, S., Forte, C., and Griebenow, K. (2007) Conformations of Alanine-Based Peptides in Water Probed by FTIR, Raman, Vibrational Circular Dichroism, Electronic Circular Dichroism, and NMR Spectroscopy, *Biochemistry 46*, 1587 - 1596.

147. Gorbunov, R. D., Kosov, D. S., and Stock, G. (2005) *Ab initio*-Based Exciton Model of Amide I Vibrations in Peptides: Definition, Conformational Dependence, and Transferability, *J. Chem. Phys. 122*, 224904 - 224904.

148. Wright, P. E., Dyson, H. J., and Lerner, R. A. (1988) Conformation of Peptide Fragments of Proteins in Aqueous Solution: Implications for Initiation of Protein Folding, *Biochemistry 27*, 7167 - 7175.

149. Sieler, G., and Schweitzer-Stenner, R. (1997) The Amide I Mode of Peptides in Aqueous Solution Involves Vibrational Coupling Between the Peptide Group and Water Molecules of the Hydration Shell, *J. Am. Chem. Soc. 119*, 1720 - 1726.

150. Glasoe, P. K., and Long, F. A. (1960) Use of Glass Electrodes to Measure Acidites in Deuterium Oxide, *J. Phys. Chem. 64*, 187 - 190.

151. Aue, W. P., Bartholdi, E., and Ernst, R. R. (1976) Two-Dimensional Spectroscopy. Application to Nuclear Magnetic Resonance, *J. Chem. Phys. 64*, 2229 - 2246.

152. Bax, A., and Freeman, R. (1981) Investigation of Complex Networks of Spin-Spin Coupling by Two-Dimensional NMR, *J. Magn. Reson. 44*, 542 - 561.

153. Jeener, J., Meier, B. H., Bachmann, P., and Ernst, R. R. (1979) Investigation of Exchange Processes by Two-Dimensional NMR Spectroscopy, *J. Chem. Phys. 71*, 4546 - 4553.

154. Macura, S., and Ernst, R. R. (1980) Elucidation of Cross-Relaxation in Liquids by Two-Dimensional NMR Spectroscopy, *Mol. Phys. 41*, 95 - 117.

155. Redfield, A. G., and Kunz, S. D. (1975) Quadrature Fourier NMR Detection: Simple Multiplex for Dual Detection and Discussion, *J. Magn. Reson. 19*, 250 - 254.

156. Marion, D., and Wüthrich, K. (1983) Application of Phase Sensitive Two-Dimensional Correlated Spectroscopy (COSY) For Measurements of ^{1}H-^{1}H Spin-Spin Coupling Constants in Proteins, *Biochem. Biophys. Res. Comm. 113*, 967 - 974.

157. States, D. J., Haberkorn, R. A., and Ruben, D. J. (1982) A Two-Dimensional Nuclear Overhauser Experiment with Pure Absorption Phase in Four Quadrants, *J. Magn. Reson. 48*, 286 - 292.

158. Hoult, D. I. (1976) Solvent Peak Saturation with Single Phase and Quadrature Fourier Transformation, *J. Magn. Reson. 21*, 337 - 347.

159. Piotto, M., Saudek, V., and Sklenár, V. (1992) Gradient-Tailored Excitation for Single-Quantum NMR Spectroscopy of Aqueous Solutions, *J. Biomol. NMR 2*, 661 - 665.

160. Sklenár, V., Piotto, M., Leppik, R., and Saudek, V. (1993) Gradient-Tailored Water Suppression for ^{1}H- ^{15}N HSQC Experiments Optimized to Retain Full Sensitivity, *J. Magn. Reson. A 102*, 241 - 245.

161. Otting, G., Widmer, H., Wagner, G., and Wüthrich, K. (1986) Origin of t_1 and t_2 Ridges in 2D NMR Spectra and Procedures for Suppression, *J. Magn. Reson. 66*, 187 - 193.

162. Wüthrich, K. (1986) *NMR of Proteins and Nucleic Acids*, Wiley-Interscience, New York.

163. Dobson, C. M., and Karplus, M. (1999) The Fundamentals of Protein Folding: Bridging Together Theory and Experiment, *Curr. Opin. Struct. Biol. 9*, 92 - 101.

164. Dobson, C. M. (2003) Protein Folding and Misfolding, *Nature 426*, 884 - 890.

165. Chiti, F., and Dobson, C. M. (2006) Protein Misfolding, Functional Amyloid, and Human Disease, *Ann. Rev. Biochem. 75*, 333 - 366.

166. Dobson, C. M. (1999) Protein Misfolding, Evolution and Disease, *Trends Biochem. Sci. 24*, 329 - 332.

167. Carrell, R. W., and Lomas, D. A. (1997) Conformational Disease, *The Lancet 350*, 134 - 138.

168. Uversky, V. N., Oldfield, C. J., and Dunker, A. K. (2008) Intrinsically Disordered Proteins in Human Diseases: Introducing the D2 Concept, *Annu. Rev. Biophys. 37*, 215 - 246.

169. Glenner, G. G. (1980) Amyloid Deposits and Amyloidosis, *N. Engl. J. Med. 302*, 1283-1291.

170. Engelender, S. (2008) Ubiquitination of α-Synuclein and Autophagy in Parkinson's Disease, *Autophagy 4*, 372 - 374.

171. Hardy, J., and Allsop, D. (1991) Amyloid Deposition as the Central Event in the Aetiology of Alzheimer's Disease, *Trends Pharm. Sci. 121*, 383 - 388.

172. Lee, H.-g., Zhu, X., Petersen, R. B., Perry, G., and Smith, M. A. (2003) Amyloids, Aggregates and Neuronal Inclusions: Good or Bad News for Neurons?, *Curr. Med. Chem. Immunol. Endocr. Metab. Agents 3*, 293 - 298.

173. Serpell, L. C., Fraser, P. E., and Sunde, M. (1999) X-Ray Fiber Diffraction of Amyloid Fibrils, *Method. Enzymol. 309*, 526 - 536.

174. Dyson, H. J., and Wright, P. E. (2005) Intrinsically Unstructured Proteins and Their Functions, *Nat. Rev. Mol. Cell Biol. 6*, 197 - 208.

175. Stefani, M., and Dobson, C. M. (2003) Protein Aggregation and Aggregate Toxicity: New Insights Into Protein Folding, Misfolding Diseases and Biological Evolution, *J. Mol. Med. 81*, 678 - 699.

176. Uversky, V. N. (2002) What Does It Mean To Be Natively Unfolded, *Eur. J. Biochem. 269*, 2 - 12.

177. Kelly, J. W. (1998) The Alternative Conformations of Amyloidogenic Proteins and Their Multi-Step Assembly Pathways, *Curr. Opin. Struct. Biol. 8*, 101 - 106.

178. Uversky, V. N., and Fink, A. L. (2004) Conformational Constraints for Amyloid Fibrillation: The Importance of Being Unfolded, *Biochim. Biophys. Acta 1698*, 131 - 153.

179. Kenney, J. M., Knight, D., Wise, M. J., and Vollrath, F. (2002) Amyloidogenic Nature of Spider Silk, *Eur. J. Biochem. 269*, 4159 - 4163.

180. Chapman, M. R., Robinson, L. S., Pinkner, J. S., Roth, R., Heuser, J., Hammar, M., Normark, S., and Hultgren, S. J. (2002) Role of *Escherichia coli* Curli Operons in Directing Amyloid Fiber Formation, *Science 295*, 851 - 855.

181. Johnston, N. (2005) γ-Secretase Makes a Splash, *The Scientist 19*, 24 - 25.

182. P. T. Lansbury, J. (1999) Evolution of Amyloid: What Normal Protein Folding May Tell Us About Fibrillogenesis and Disease, *Proc. Natl. Acad. Sci. 96*, 3342 - 3344.

183. Aggeli, A., Nyrkova, I. A., Bell, M., Harding, R., Carrick, L., McLeish, T. C. B., Semenov, A. N., and Boden, N. (2001) Hierarchical Self-Assembly of Chiral Rod-Like Molecules as a Model for Peptide β-Sheet Tapes, Ribbons, Fibrils, and Fibers, *Proc. Natl. Acad. Sci. USA 98*, 11857 - 11862.

184. Giurleo, J. T., He, X., and Talaga, D. S. (2008) β-Lactoglobulin Assembles into Amyloid Through Sequential Aggregated Intermediates, *J. Mol. Biol. 381*, 1332 - 1348.

185. Knowles, T. P. J., Waudby, C. A., Devlin, G. L., Cohen, S. I. A., Aguzzi, A., Vendruscolo, M., Terentjev, E. M., Welland, M. E., and Dobson, C. M. (2009) An Analytical Solution to the Kinetics of Breakable Filament Assembly, *Science 326*, 1533 - 1537.

186. Gazit, E. (2002) The "Correctly Folded" State of Proteins: Is it a Metastable State?, *Angew. Chem. Int. Ed. 41*, 257 - 259.

187. Brorsson, A.-C., Bolognesi, B., Tartaglia, G. G., Shammas, S. L., Favrin, G., Watson, I., Lomas, D. A., Chiti, F., Vendruscolo, M., Dobson, C. M., Crowther, D. C., and Luheshi, L.

M. (2010) Intrinsic Determinants of Neurotoxic Aggregate Formation by the Amyloid β Peptide, *Biophys. J. 98*, 1677 - 1684.

188. McLean, C. A., Cherny, R. A., Fraser, F. W., Fuller, S. J., Smith, M. J., Beyreuther, K., and Bush, A. I. (1999) Soluble Pool of Aβ Amyloid as a Determinant of the Severity of Neurodegeneration in Alzheimer's Disease, *Ann. Neurol. 46*, 860 - 866.

189. Monsellier, E., and Chiti, F. (2007) Prevention of Amyloid-Like Aggregation as a Driving Force of Protein Evolution, *EMBO Rep. 8*, 737 - 742.

190. Wiseman, R. L., Haynes, C. M., and Ron, D. (2010) Snapshot: The Unfolded Protein Response, *Cell 140*.

191. Dobson, C. M. (2001) The Structural Basis of Protein Folding and its Links with Human Disease, *Philos. Trans. R. Soc. Lond. B Biol. Sci. 356*, 133 - 145.

192. Hartl, F. U., and Hayer-Hartl, M. (2002) Molecular Chaperones in the Cytosol: From Nascent Chain to Folded Protein, *Science 295*.

193. Bukau, B., and Horwich, A. L. (1998) The Hsp70 and Hsp60 Chaperone Machines, *Cell 92*, 351 - 366.

194. Fändrich, M., and Dobson, C. (2002) The Behaviour of Polyamino Acids Reveals an Inverse Side Chain Effect in Amyloid Structure Formation, *EMBO J. 21*, 5682 - 5690.

195. Fändrich, M., Forge, V., Buder, K., Kittler, M., Dobson, C. M., and Diekmann, S. (2003) Myoglobin Forms Amyloid Fibrils by Association of Unfolded Polypeptide Segments, *Proc. Natl. Acad. Sci. USA 100*, 15463 - 15468.

196. Smith, A. (2003) Nature Insight: Protein Misfolding, *Nature 426*, 883.

197. Hartl, F. U., and Hayer-Hartl, M. (2009) Converging Concepts of Protein Folding *in vitro* and *in vivo*, *Nat. Struct. Mol. Biol. 16*, 574 - 581.

198. Khurana, R., Uversky, V. N., Nielsen, L., and Fink, A. L. (2001) Is Congo Red an Amyloid-Specific Dye?, *J. Biol. Chem. 276*, 22715 - 22721.

199. Makin, O. S., Atkins, E., Sikorski, P., Johansson, J., and Serpell, L. C. (2005) Molecular Basis for Amyloid Fibril Formation and Stability, *Proc. Natl. Acad. Sci. USA 102*, 315 - 320.

200. Nelson, R., Sawaya, M. R., Balbirnie, M., Madsen, A. Ø., Riekel, C., Grothe, R., and Eisenberg, D. (2005) Structure of the Cross-β Spine of Amyloid-Like Fibrils, *Nature 435*, 773 - 778.

201. Walsh, D. M., Lomakin, A., Benedek, G. B., Condron, M. M., and Teplow, D. B. (1997) Amyloid β-Protein Fibrillogenesis, *J. Biol. Chem. 272*, 22364 - 22372.

202. Sokolowski, F., Modler, A. J., Masuch, R., Zirwer, D., Baier, M., Lutsch, G., Moss, D. A., Gast, K., and Naumann, D. (2003) Formation of Critical Oligomers Is a Key Event During Conformational Transition of Recombinant Syrian Hamster Prion Protein, *J. Biol. Chem. 278*, 40481 - 40492.

203. Modler, A. J., Gast, K., Lutsch, G., and Damaschun, G. (2003) Assembly of Amyloid Protofibrils *via* Critical Oligomers - A Novel Pathway of Amyloid Formation, *J. Mol. Biol. 325*, 135 - 148.

204. LeVine III, H. (1993) Thioflavin T Interaction with Synthetic Alzheimer's Disease β-Amyloid Peptides: Detection of Amyloid Aggregation in Solution, *Protein Sci. 2*, 404 - 410.

205. Wood, S. J., Maleeff, B., Hart, T., and Wetzel, R. (1996) Physical, Morphological and Functional Differences Between pH 5.8 and 7.4 Aggregates of the Alzheimer's Amyloid Peptide Aβ, *J. Mol. Biol. 256*, 870 - 877.

206. Jaroniec, C. P., MacPhee, C. E., Astrof, N. S., Dobson, C. M., and Griffin, R. G. (2002) Molecular Conformation of a Peptide Fragment of Transthyretin in an Amyloid Fibril, *Proc. Natl. Acad. Sci. USA 99*, 16748 - 16753.

207. Petkova, A. T., Ishii, Y., Balbach, J. J., Antzutkin, O. N., Leapman, R. D., Delaglio, F., and Tycko, R. (2002) A Structural Model for Alzheimer's b-Amyloid Fibrils Based on Experimental Constraints from Solid State NMR, *Proc. Natl. Acad. Sci. USA 99*, 16742 - 16747.

208. Ritter, C., Maddelein, M.-L., Siemer, A. B., Lührs, T., Ernst, M., Meier, B. H., Saupe, S. J., and Riek, R. (2005) Correlation of Structural Elements and Infectivity of the HET-s Prion, *Nature 435*, 844 - 848.

209. Makin, O. S., and Serpell, L. C. (2004) Structural Characterisation of Islet Amyloid Polypeptide Fibrils, *J. Mol. Biol. 335*, 1279 - 1288.

210. Goldsbury, C. S., Cooper, G. J. S., Goldie, K. N., Müller, S. A., Saafi, E. L., Gruijters, W. T. M., Misur, M. P., Engel, A., Aebi, U., and Kistler, J. (1997) Polymorphic Fibrillar Assembly of Human Amylin, *J. Struct. Biol. 119*, 17 - 27.

211. Zandomeneghi, G., Krebs, M. R. H., McCammon, M. G., and Fändrich, M. (2004) FTIR Reveals Structural Differences Between Native b-sheet Proteins and Amyloid Fibrils, *Protein Sci. 13*, 3314 - 3321.

212. Shashilov, V. A., and Lednev, I. K. (2008) 2D Correlation Deep UV Resonance Raman Spectroscopy of Early Events of Lysozyme Fibrillation: Kinetic Mechanism and Potential Interpretation Pitfalls, *J. Am. Chem. Soc. 130*, 309 - 317.

213. Ma, S., Cao, X., Mak, M., Sadik, A., Walkner, C., Freedman, T. B., Lednev, I. K., Dukor, R. K., and Nafie, L. A. (2007) Vibrational Circular Dichroism Shows Unusual Sensitivity to Protein Fibril Formation and Development in Solution, *J. Am. Chem. Soc. 129*, 12364 - 12365.

214. Blanch, E. W., Morozova-Roche, L. A., Cochran, D. A. E., Doig, A. J., Hecht, L., and Barron, L. D. (2000) Is Polyproline II Helix the Killer Conformation? A Raman Optical Activity Study of the Amyloidogenic Prefibrillar Intermediate of Human Lysozyme, *J. Mol. Biol. 301*, 553 - 563.

215. JiJi, J. D., Balakrishnan, G., Hu, Y., and Spiro, T. G. (2006) Intermediacy of Poly(L-proline) II and β-Strand Conformations in Poly(L-lysine) β-Sheet Formation Probed by Temperature-Jump/UV Resonance Raman Spectroscopy, *Biochemistry 45*, 34 - 41.

216. Darnell, G., Orgel, J. P. R. O., Pahl, R., and Meredith, S. C. (2007) Flanking Polyproline Sequences Inhibits b-Sheet Structure in Polyglutamine Segments by Inducing PPII-like Helix Structure, *J. Mol. Biol. 374*, 688 - 704.

217. Darnell, G. D., Derrberry, J., Kurutz, J. W., and Meredith, S. C. (2009) Mechanism of *Cis*-Inhibition of PolyQ Fibrillation by PolyP: PPII Oligomers and the Hydrophobic Effect, *Biophys. J. 97*, 2295 - 2305.

218. Tzotzos, S., and Doig, A. J. (2010) Amyloidogenic Sequences in Native Protein Structures, *Protein Sci. 19*, 327 - 348.

219. de la Paz, M. L., Goldie, K., Zurdo, J., Lacroix, E., Dobson, C. M., Hoenger, A., and Serrano, L. (2002) *De novo* Designed Peptide-Based Amyloid Fibrils, *Proc. Natl. Acad. Sci. USA 99*, 16052 - 16057.

220. Madine, J., Copland, A., Serpell, L. C., and Middleton, D. A. (2009) Cross-β-Spine Architecture of Fibrils Formed by the Amyloidogenic Segment NFGSVQFV of Medin from Sold-State NMR and X-ray Fiber Diffraction Measurements, *Biochemistry 48*, 3089 - 3099.

221. Chaudhary, N., Singh, S., and Nagaraj, R. (2009) Morphology of Self-Assembled Structures Formed by Short Peptides from the Amyloidogenic Protein tau Depends on the Solvent in Which the Peptides are Dissolved, *J. Pept. Sci. 15*, 675 - 684.

222. Chiti, F., Stefani, M., Taddei, N., Ramponi, G., and Dobson, C. M. (2003) Rationalization of the Effects of Mutations on Peptide and Protein Aggregation Rates, *Nature 424*, 805 - 808.

223. Tartaglia, G. G., Cavalli, A., Pellarin, R., and Caflisch, A. (2005) Prediction of Aggregation Rate and Aggregation-Prone Segments in Polypeptide Sequences, *Protein Sci. 14*, 2723 - 2734.

224. Marshall, K. E., Hicks, M. R., Williams, T. L., Hoffman, S. V., Rodger, A., Dafforn, T. R., and Serpell, L. C. (2010) Characterizing the Assembly of the Sup35 Yeast Prion Fragment, GNNQQNY: Structural Changes Accompany a Fiber-to-Crystal Switch, *Biophys. J. 98*, 330 - 338.

225. Gazit, E. (2002) A Possible Role for π-Stacking in the Self-Assembly of Amyloid Fibrils, *FASEB J. 16*, 77 - 83.

226. Azriel, R., and Gazit, E. (2001) Analysis of the Minimal Amyloid-Forming Fragment of the Islet Amyloid Polypeptide, *J. Biol. Chem. 276*, 34156 - 34161.

227. Tartaglia, G. G., Cavalli, A., Pellarin, R., and Caflisch, A. (2004) The Role of Aromaticity, Exposed Surface, and Dipole Moment in Determining Protein Aggregation Rates, *Protein Sci. 13*, 1939 - 1941.

228. Gazit, E. (2007) Self-Assembly of Short Aromatic Peptides into Amyloid Fibrils and Related Nanostructures, *Prion 1*, 32 - 35.

229. Ma, M., Kuang, Y., Gao, Y., Zhang, Y., Gao, P., and Xu, B. (2010) Aromatic-Aromatic Interactions Induce the Self-Assembly of Pentapeptidic Derivatives in Water To Form Nanofibers and Supramolecular Hydrogels, *J. Am. Chem. Soc. 132*, 2719-2728.

230. Smith, A. M., Williams, R. J., Tang, C., Coppo, P., Collins, R. F., Turner, M. L., Saiani, A., and Ulijn, R. V. (2008) Fmoc-Diphenylalanine Self Assembles to a Hydrogel via a Novel Architecture Based on π-π Interlocked β-Sheets, *Adv. Mater. 20*, 37 - 41.

231. Gsponer, J., Haberthür, U., and Caflisch, A. (2003) The Role of Side-Chain Interactions in the Early Steps of Aggregation: Molecular Dynamics Simulations of an Amyliod-Forming Peptide From the Yeast Prion Sup35, *Proc. Natl. Acad. Sci. USA 100*, 5154 - 5159.

232. Marek, P., Abedini, A., Song, B., Kanuungo, M., Johnson, M. E., Gupta, R., Zaman, W., Wong, S. S., and Raleigh, D. P. (2007) Aromatic Interactions Are Not Required for Amyloid Fibril Formation by Islet Amyloid Polypeptide but Do Influence the Rate of Fibril Formation and Fibril Morphology, *Biochemistry 46*, 3255 - 3261.

233. Bemporad, F., Taddei, N., Stefani, M., and Chiti, F. (2006) Assessing the Role of Aromatic Residues in the Amyloid Aggregation of Human Muscle Acylphosphatase, *Protein Sci. 15*, 862 - 870.

234. Yokoi, H., Kinoshita, T., and Zhang, S. (2005) Dynamic Reassembly of Peptide RADA16 Nanofiber Scaffold, *Proc. Natl. Acad. Sci. USA 102*, 8414 - 8419.

235. Zhang, S., and Altman, M. (1999) Peptide Self-Assembly in Functional Polymer Science and Engineering, *Rect. Funct. Polym. 41*, 91 - 102.

236. Kim, W., and Hecht, M. H. (2006) Generic Hydrophobic Residues are Sufficient to Promote Aggregation of the Alzheimer's Aβ42 Peptide, *Proc. Natl. Acad. Sci. USA 103*, 15824 - 15829.

237. Rochet, J.-C., and P. T. Lansbury, J. (2000) Amyloid Fibrillogenesis: Themes and Variations, *Curr. Opin. Struct. Biol. 10*, 60 - 68.

238. Schweitzer-Stenner, R., and Measey, T. J. (2007) The Alanine-Rich XAO Peptide Adopts a Heterogeneous Population, Including Turn-Like and Polyproline II Conformations, *Proc. Natl. Acad. Sci. USA 104*, 6649 - 6654.

239. Klunk, W. E., Jacob, R. F., and Mason, R. P. (1999) Quantifying Amyloid by Congo Red Spectral Shift Assay, *Method. Enzymol. 309*, 285 - 305.

240. Lee, C., and Cho, M. (2004) Local Amide I Mode Frequencies and Coupling Constants in Multiple-Stranded Antiparallel β-Sheet Polypeptides, *J. Phys. Chem. B 108*, 20397 - 20407.

241. Caughey, B., and P. T. Lansbury, J. (2003) Protofibrils, Pores, Fibrils, and Neurodegeneration: Separating the Responsible Protein Aggregates from The Innocent Bystanders, *Ann. Rev. Neurosci. 26*, 267 - 298.

242. Perutz, M. F., and Windle, A. H. (2001) Cause of Neural Death in Neurodegenerative Diseases Attributable to Expansion of Glutamine Repeats, *Nature 412*, 143 - 144.

243. Bitan, G., Kirkitadze, M. D., Lomakin, A., Vollers, S. S., Benedek, G. B., and Teplow, D. B. (2002) Amyloid β-Protein (Aβ) Assembly: Aβ40 and Aβ42 Oligomerize Through Distinct Pathways, *Proc. Natl. Acad. Sci. USA 100*, 330 - 335.

244. Bucciantini, M., Giannoni, E., Chiti, F., Baroni, F., Formigli, L., Zurdo, J., Taddei, N., Ramponi, G., Dobson, C. M., and Stefani, M. (2002) Inherent Toxicity of Aggregates Implies a Common Mechanism for Protein Misfolding Diseases, *Nature 416*, 507 - 511.

245. Johnson Jr., W. C. (1990) Protein Secondary Structure and Circular Dichroism: A Practicle Guide, *Proteins 7*, 205 - 214.

246. Ma, J. C., and Dougherty, D. A. (1997) The Cation - π Interaction, *Chem. Rev. 97*, 1303 - 1324.

247. Measey, T. J., Smith, K. B., Decatur, S. M., Zhao, L., Yang, G., and Schweitzer-Stenner, R. (2009) Self-Aggregation of a Polyalanine Octamer Promoted by Its C-Terminal Tyrosine and Probed by a Strongly Enhanced Vibrational Circular Dichroism Signal, *J. Am. Chem. Soc. 131*, 18218 - 18219.

248. Klimov, D. K., and Thirumalai, D. (2003) Dissecting the Assembly of $A\beta_{16-22}$ Amyloid Peptides into Antiparallel β Sheets, *Structure 11*, 295 - 307.

249. Balbirnie, M., Grothe, R., and Eisenberg, D. S. (2001) An Amyloid-Forming Peptide from the Yeast Prion Sup35 Reveals a Dehydrated b-Sheet Structure for Amyloid, *Proc. Natl. Acad. Sci. USA 98*, 2375 - 2380.

250. Sawaya, M. R., Sambashivan, S., Nelson, R., Ivanova, M. I., Sievers, S. A., Apostol, M. I., Thompson, M. J., Balbirnie, M., Wiltzius, J. J. W., McFarlane, H. T., Madsen, A. Ø., Riefel, C., and Eisenberg, D. (2007) Atomic Structures of Amyloid Cross-β Spines Reveal Varied Steric Zippers, *Nature 447*, 453 - 457.

251. van der Wel, P. C. A., Lewandowski, J. R., and Griffin, R. G. (2007) Solid-State NMR Study of Amyloid Nanocrystals and Fibrils Formed by the Peptide GNNQQNY from Yeaset Prion Protein Sup35p, *J. Am. Chem. Soc.*, 5117 - 5130.

252. Whitesides, G. M., Mathias, J. P., and Seto, C. T. (1991) Molecular Self-Assembly and Nanochemistry: A Cheical Strategy for the Synthesis of Nanostructures, *Science 254*, 1312 - 1319.

253. Whitesides, G. M. (2002) Self-Assembly at All Scales, *Science 295*, 2418 - 2421.

254. Ma, P. X. (2008) Biomimetic Materials for Tissue Engineering, *Adv. Drug Delivery Rev. 60*, 184 - 198.

255. Rosenbloom, J., Abrams, W. R., and Mecham, R. (1993) Extracellular Matrix 4: The Elastic Fiber, *FASEB J. 7*, 1208 - 1218.

256. Peppas, N. A., Huang, Y., Torres-Lugo, M., Ward, J. H., and Zhang, J. (2000) Physicochemical Foundations and Structural Design of Hydrogels in Medicine and Biology, *Annu. Rev. Biomed. Eng. 2*, 9 - 29.

257. Satish, C. S., Satish, K. P., and Shivakumar, H. G. (2006) Hydrogels as Controlled Drug Delivery Systems: Synthesis, Crosslinking, Water and Drug Transport Mechanism, *Indian J. Pharm. Sci. 68*, 133 - 140.

258. Wichterle, O., and Lim, D. (1960) Hydrophilic Gels for Biological Use, *Nature 185*, 117 - 118.

259. Zhang, S. (2004) Hydrogels: Wet or Let Die, *Nature Mat. 3*, 7 - 8.

260. Trevors, J. T., and Pollack, G. H. (2005) Hypothesis: The Origin of Life in a Hydrogel Environment, *Prog. Biophys. Mol. Biol. 89*, 1 - 8.

261. Koutsopoulos, S., Unsworth, L. D., Nagai, Y., and Zhang, S. (2009) Controlled Release of Functional Proteins Through Designer Self-Assembling Peptide Nanofiber Hydrogel Scaffold, *Proc. Natl. Acad. Sci. USA 106*, 4623-4628.

262. Ruan, L., Zhang, H., Luo, H., Liu, J., Tang, F., Shi, Y.-K., and Zhao, X. (2009) Designed Amphiphilic Peptide Forms Stable Nanoweb, Slowly Releases Encapsulated Hydrophobic Drug, and Accelerates Animal Hemostasis, *Proc. Natl. Acad. Sci. USA 106*, 5105 - 5110.

263. Holmes, T. C., Lacalle, S. d., Su, X., Liu, G., Rich, A., and Zhang, S. (2000) Extensive Neurite Outgrowth and Active Synapse Formation on Self-Assembling Peptide Scaffolds, *Proc. Natl. Acad. Sci. USA 97*, 6728 - 6733.

264. Nagai, Y., Unsworth, L. D., Koutsopoulos, S., and Zhang, S. (2006) Slow Release of Molecules in Self-Assembling Peptide Nanofiber Scaffold, *J. Cont. Rel. 115*, 18 - 25.

265. Zhang, S., Holmes, T., Lockshin, C., and Rich, A. (1993) Spontaneous Assembly of A Self-Complementary Oligopeptide To Form A Stable Macroscopic Membrane, *Proc. Natl. Acad. Sci. USA 90*, 3334 - 3338.

266. Semino, C. E., Merok, J. R., Crane, G. G., Panagiotakos, G., and Zhang, S. (2003) Functional Differentiation of Hepatocyte-Like Spheroid Structures From Putative Live Propenitor Cells In Three-Dimensional Peptide Scaffolds, *Differentiation 71*, 262 - 270.

267. Sun, J., and Zheng, Q. (2009) Experimental Study on Self-Assembly of KLD-12 Peptide Hydrogel and 3-D Culture of MSC Encapsulated with Hydrogel *In Vitro*, *J. Huazhong Univ. Sci. Tech. 29*, 512 - 516.

268. Kisiday, J., Jin, M., Kurz, B., Hung, H., Semino, C., Zhang, S., and Grodzinsky, A. J. (2002) Self-Assemling Peptide Hydrogel Fosters Chondrocyte Extracellular Matrix Production and Cell Division: Implications for Cartilage Tisue Repair, *Proc. Natl. Acad. Sci. USA 99*, 9996 - 10001.

269. Zhang, S., Holmes, T. C., DiPersio, C. M., Hynes, R. O., Su, X., and Rich, A. (1995) Self-Complementary Oligopeptide Matrices Support Mammalian Cell Attachment, *Biomaterials 16*, 1385 - 1393.

270. Zou, D., Tie, Z., Lu, C., Qin, M., Lu, X., Wang, M., Wang, W., and Chen, P. (2010) Effects of Hydrophobicity and Anions on Self-Assembly of the Peptide EMK16-II, *Biopolymers 93*, 318 - 329.

271. Schneider, J. P., Pochan, D. J., Ozbas, B., Rajagopal, K., Pakstis, L., and Kretsinger, J. (2002) Responsive Hydrogels from the Intramolecular Folding and Self-Assembly of a Designed Peptide, *J. Am. Chem. Soc. 124*, 15030 - 15037.

272. Connelly, S., Choi, S., Johnson, S. M., Kelly, J. W., and Wilson, I. A. (2010) Structure-Based Design of Kinetic Stabilizers that Ameliorate the Transthyretin Amyloidoses, *Curr. Opin. Struct. Biol. 20*, 54 - 62.

273. Measey, T. J., and Schweitzer-Stenner, R. (2006) Aggregation of the Amphipathic Peptides $(AAKA)_n$ into Antiparallel β-Sheets, *J. Am. Chem. Soc. 128*, 13324 - 13325.

274. Scholtz, J. M., Marqusee, S., Baldwin, R. L., Work, E. J., Stewart, J. M., Santoro, M., and Bolen, D. W. (1991) Calorimetric Determination of the Enthalpy Change for the a-Helix to Coil Transition of an Alanine Peptide in Water, *Proc. Natl. Acad. Sci. USA 88*, 2854 - 2858.

275. Woody, R. W. (2009) Circular Dichroism Spectrum of Peptides in the Poly(Pro) II Conformation, *J. Am. Chem. Soc. 131*, 8234 - 8245.

276. Jang, S., Yuan, J.-M., Shin, J., Measey, T. J., Schweitzer-Stenner, R., and Li, F.-Y. (2009) Energy Landscapes Associated with the Self-Aggregation of an Alanine-Based Oligopeptide $(AAKA)_4$, *J. Phys. Chem. B 113*, 6054 - 6061.

277. Woodward, C. (1993) Is the Slow-Exchange Core the Protein Folding Core?, *Trends Biochem. Sci. 18*, 359 - 360.

278. Hsu, I.-J., Shiu, Y.-J., Jeng, U.-S., Chen, T.-H., Huang, Y.-S., Lai, Y.-H., Tsai, L.-N., Jang, L.-Y., Lee, J.-F., Lin, L.-J., Lin, S.-H., and Wang, Y. (2007) A Solution Study on the Local and Global Structure Changes of Cytochrome c: An Unfolding Process Induced by Urea, *J. Phys. Chem. A 111*, 9286 - 9290.

279. Chakrabartty, A., Kortemme, T., and Baldwin, R. L. (1994) Helix Propensities of the Amino Acids Measured in Alanine-Based Peptides Without Helix-Stabilizing Side-Chain Interactions, *Protein Sci. 3*, 843 - 852.

280. Auer, S., Dobson, C. M., Vendruscolo, M., and Maritan, A. (2008) Self-Templated Nucleation in Peptide and Protein Aggregation, *Phys. Rev. Lett. 101*, 258101.

APPENDIX A: LIST OF PUBLICATIONS

Publications which contributed to the writing of this thesis:

1. **T. J. Measey**, R. Schweitzer-Stenner, V. Sa, and K. Kornev. Defying the Rules of Peptide Self-Assembly: A Spectroscopic and Rheological Study of an Alanine-Rich Oligopeptide. (2010) *Submitted to Macromolecules.*

2. **T. J. Measey**, K. B. Smith, S. M. Decatur, L. Zhao, G. Yang, R. Schweitzer-Stenner. The Self-Aggregation of A Polyalanine Octomer Promoted by its C-Terminal Tyrosine and Probed by A Strongly Enhanced VCD Signal. ***J. Am. Chem. Soc.*** (2009) 131, 18218-18219.

3. S. Jang, J-M. Yuan, J. Shin, **T. J. Measey**, R. Schweitzer-Stenner. Energy Landscapes Associated with the Self-Aggregation of an Alanine-Based Oligopeptide $(AAKA)_4$. ***J. Phys. Chem. B*** (2009) 113, 6054-6061.

4. R. Schweitzer-Stenner and **T.J. Measey**. The Alanine-Rich XAO Peptide Adopts a Heterogeneous Population, Including Turn-Like and PPII Conformations. ***Proc. Natl. Acad. Sci.*** (2007) 104, 6649-6654.

5. R. Schweitzer-Stenner, **T. Measey**, L. Kakalis, F. Jordan, S. Pizzanelli, C. Forte, and K. Griebenow. Alanine-Based Peptides in Water Probed by FTIR, Raman, Vibrational Circular Dichroism, Electronic Circular Dichroism, and NMR Spectroscopy. ***Biochemistry*** (2007) 46, 1587-1596.

6. **T. J. Measey**, and R. Schweitzer-Stenner. Aggregation of the Amphipathic Peptides $(AAKA)_n$ into Antiparellel β-Sheets. ***J. Am. Chem. Soc.*** (2006) 128, 13324-13325.

Publications which did not directly contribute to this thesis:

7. R. Schweitzer-Stenner, **T.J. Measey**. Simulation of IR, Raman, and VCD Amide I Band Profiles of Self-Assembled Peptides. ***Spectroscopy*** (2010) 24, 25 – 36.

8. L. Tooke, L. Duitch, **T. J. Measey**, R. Schweitzer-Stenner. Kinetics of the Self-Aggregation and Film Formation of Poly-L-Proline at High Temperatures Explored by Circular Dichroism Spectroscopy. ***Biopolymers*** (2010) 93, 451-457.

9. S. Pizzaneli, C. Forte, S. Monti, G. Zandomeneghi, B.H. Meier, A. Hagarman, **T. J. Measey**, R. Schweitzer-Stenner. Violation of the Flory isolated-pair hypothesis for the tripeptides AFA and GFG. A combined NMR, FTIR, Polarized Raman, and VCD Spectroscopy Study. **J Phys. Chem. B** (2010) 114, 3965 - 3978.

10. A. Hagarman, T.J. Measey, D. Mathieu, H. Schwalbe, R. Schweitzer-Stenner. Intrinsic Propensities of Amino Acid Residues in GXG peptides Inferred from Amide I' Band Profiles and NMR Scalar Coupling Constants. **J. Am. Chem. Soc.** (2010) 132, 540–551.

11. R. Schweitzer-Stenner, **T. J. Measey**, A. Hagarman, I. C. Dragomir. The Structure of Unfolded Peptides and Proteins explored by Vibrational Spectroscopy. In Instrumental Analysis of Intrinsically Disordered Proteins. 171 – 224. Eds. S. Longhi and V. N. Uversky. Wiley. 2010.

12. I.C. Dragomir, **T.J. Measey**, A.M. Hagarman, and R. Schweitzer-Stenner. Environment-Controlled Interchromophore Charge Transfer Transitions in Dipeptides Probed by UV Absorption and Electronic Circular Dichroism Spectroscopy. ***J. Phys. Chem. B*** (2006) 110(26), 13235-13241.

13. R. Schweitzer-Stenner, **T. Measey**, A. Hagarman, F. Eker, and K. Griebenow. Salmon Calcitonin and Amyloid β: Two Peptides with Amyloidogenic Capacity Adopt Different Conformational Manifolds in Their Unfolded States. ***Biochemistry*** (2006) 45, 2810-2819.

14. A. Hagarman, **T. Measey**, R. Doddasomayajula, I. Dragomir, F. Eker, K. Griebenow, and R. Schweitzer-Stenner. Conformational Analysis of XA and AX Dipeptides in Water by Electronic Circular Dichroism and ^{1}H NMR Spectroscopy. ***J. Phys. Chem. B*** (2006) 110, 6979-6986.

15. **T. Measey** and R. Schweitzer-Stenner. The Conformations Adopted by The Octamer Peptide $(AAKA)_2$ in Aqueous Solution Probed By FTIR and Polarized Raman Spectroscopy. ***J. Raman Spectrosc.*** **(invited article for a special issue celebrating the 65th birthday of Prof. Wolfgang Kiefer)** (2006) 37, 248-254.

16. **T. Measey** and R. Schweitzer-Stenner. Simulation of amide I band profiles of trans-polyproline based on an excitonic coupling model. ***Chem. Phys. Letts***. (2005) 408, 123-127.

17. **T. Measey**, A. Hagarman, F. Eker, K. Griebenow and R. Schweitzer-Stenner. Side chain dependence of intensity and wavenumber position of amide I' in IR and visible Raman spectra of XA and AX dipeptides. ***J. Phys. Chem. B*** (2005) 109, 8195-8205.

VITA

Thomas Jeffrey Measey was born at 7:47 am on May 15th, 1982 to Thomas Mann and Cassandra Marie Measey, their first child. Thomas grew up in Burlington City, NJ, where he had a typical childhood. He played sports, including soccer, baseball, and basketball, and loved to draw. He attended St. Paul's Catholic School in Burlington, NJ from grades K – 8. He developed an interest in music at a young age, and took up various instruments, including piano and guitar. Thomas received his first guitar as a Christmas gift in 1993, at the age of 11. Thomas attended Burlington Township High School in 1996. He continued to pursue music, fronting a band with some friends, and performing at various local events. During his high school years, he enjoyed science and art classes, in particular, Chemistry. Thomas entered The College of Arts and Sciences at Drexel University as a science major with an undecided concentration in the fall of 2000. He soon changed his major from "Undecided" to "Chemistry." During his undergraduate career, Thomas began doing research in the laboratory of Prof. Reinhard Schweitzer-Stenner in the Chemistry Department, who, at the time, was a new faculty member at Drexel. Thomas was pivotal (along with Andrew Hagarman) in the developmental phase of Prof. Schweitzer-Stenner's new laboratory at Drexel. Thomas' undergraduate research concerned the measurement of various spectral parameters of a series of short, alanine-based dipeptides using vibrational spectroscopic methods such as FTIR, Raman, and VCD. After completing his undergraduate studies, Thomas continued at Drexel in pursuit of advanced degrees. All the while, Thomas continued conducting research in the field of biophysical chemistry under the supervision of Prof. Schweitzer-Stenner. During his graduate student career, Thomas' various projects included (1) determining the intrinsic conformational propensities of amino acids in short peptides, (2) the use of vibrational circular dichroism (VCD) to probe the fibril formation of short alanine-rich peptides, and (3) investigating the unusual conformational instability and hydrogel formation of

alanine-based oligopeptides for biotechnological applications. Thomas obtained his Masters in Physical Chemistry in 2008, and his Ph.D in Physical Chemistry in 2010, both from Drexel University.