SPECTROSCOPIC CHARACTERIZATION OF CONFORMATIONS OF SHORT ALANINE PEPTIDES, POLY-L-LYSINE AND INSULIN DURING FORMATION OF AMYLOID FIBRILS

Andreja Mirtič

Doctoral Dissertation
Jožef Stefan International Postgraduate School
Ljubljana, Slovenia, May 2013

Evaluation Board:

Prof. Dr. Eva Žerovnik, Jožef Stefan Institute, Jamova cesta 39, 1000 Ljubljana
Prof. Dr. Dušan Turk, Jožef Stefan Institute, Jamova cesta 39, 1000 Ljubljana
Asst. Prof. Dr. Marjana Novič, National Institute of Chemistry, Hajdrihova 19, 1000 Ljubljana

MEDNARODNA PODIPLOMSKA ŠOLA JOŽEFA STEFANA
JOŽEF STEFAN INTERNATIONAL POSTGRADUATE SCHOOL

Andreja Mirtič

SPECTROSCOPIC CHARACTERIZATION OF CONFORMATIONS OF SHORT ALANINE PEPTIDES, POLY-L-LYSINE AND INSULIN DURING FORMATION OF AMYLOID FIBRILS

Doctoral Dissertation

SPEKTROSKOPSKO DOLOČEVANJE KONFORMACIJ KRATKIH PEPTIDOV ALANINA, POLI-L-LIZINA IN INZULINA PRI TVORBI AMILOIDNIH FIBRIL

Doktorska disertacija

Supervisor: Asst. Prof. Dr. Jože Grdadolnik

Ljubljana, Slovenia, May 2013

Index

Abstract

The process of protein folding describes the transition by which an unordered polypeptide chain attains its functional native three-dimensional structure. A detailed understanding of the principles which govern the protein folding, such as conformational preferences of amino acid residues, the neighbour residue effect and the solvent effects, remains an important unsolved challenge that encodes the folding energy landscape of a protein. Knowing the factors that contribute to the conformational stability of amino acid residues would provide the insight not only into the molecular basis of unfolded states, but also into the earliest events that occur during the protein folding and misfolding. To study the conformational preferences of peptide backbone in unfolded state, the short alanine peptides were examined. By changing the peptide's environment, the competition between the intra- and intermolecular hydrogen bonding was challenged. We applied IR, Raman, VCD, NMR and UV-CD spectroscopy to determine the distribution of conformations in aqueous and non-aqueous solution. We have observed high sensitivity of amide III spectral region where each conformation has the characteristic band frequency and shape in the case of VCD. In water, the P_{II} conformation of alanine dipeptide was found to be stabilized by formation of directionally oriented water bridges between C=O and NH peptide groups using DFT calculations. The agreement between the experimental and calculated spectra improved significantly upon the addition of explicit solvent molecules to the computational model in the case of polar solvents. Additionally, we found a good correlation between the conformational stabilization of short alanine peptides by the solvent features and the electrostatic screening model, where conformations strongly depend on the local electrostatic energy and its screening by the solvent. With increasing the peptide chain the neighbour effect of alanine residue was observed that resulted in decreasing of P_{II} population and strengthening the favourable backbone electrostatic interactions in short alanine peptides.

Protein misfolding can lead to the protein self-assembly and furthermore to the formation of amyloid fibrils that give rise to serious lethal diseases. The mechanism of misfolding and aggregation of proteins composing the amyloid fibrils are poorly understood. Spectroscopic investigations of the conformational changes during fibrillation require the elucidation of the

relationship between spectral features and the protein folding reaction coordinate(s). By applying infrared spectroscopy, we have gained the insights into peptide secondary structures through assignation of the amide bands for different conformations of the model peptide poly-L-lysine (PLL). In water at low pH values, PLL mainly possessed the P_{II} and β structures, while at higher pH values and low temperatures, characteristic band for the α-helical conformation was found. The increase in temperature induced the formation of β structures that are components of amyloid fibrils. To determine the assignment of the infrared bands of individual conformations, different solvents were used that selectively stabilize one type of conformation. Among all the solvents, only ethylene glycol promoted the formation of a uniform P_{II}-helix, suggesting that the P_{II} structure is not limited to the presence of water molecules or charged side chains, as previously assumed. Knowing the spectral assignment, we could follow the fibrillation mechanism of PLL by applying the difference spectroscopy that provide us with information about the structural changes during temperature induced PLL fibrillation. Spectral changes below transition temperature were ascribed mainly to the melting of α-helces. The band assigned to P_{II}-helix starts to lose the intensity just before the fibrillation of PLL started, thus indicating its role as intermediate structure in fibrillation process. In order to affect the kinetic of PLL fibrillation, we performed a series of measurements of PLL with promotor of PLL fibrillation, such as salt $NaClO_4$, or inhibitor of PLL fibrillation, such as DPPA+DPPC vesicles. The salt $NaClO_4$ lowered the transition temperature from the α-helical structure to the β-sheet fibrils of PLL for 10 °C, where the mechanism of its action was accounted to the stabilization of the P_{II}-helix. With the addition of the DPPA+DPPC vesicles to the PLL, the stabilization of α- and P_{II}-helices was observed and, simultaneously, the destabilization of those vesicles by this interaction. However, above the melting temperature of the vesicles, the band characteristic for the β-strand appeared, followed by the increase of intensity of the band that is characteristic for the aggregated β-sheets. We proposed a mechanism for PLL amyloid fibril formation in which the α-helical PLL melts into P_{II}-helix, followed by the formation of β-strands that stack into β-sheets. We showed that difference infrared spectroscopy is a very convenient method to obtain structural properties of the intermediate in the fibrillation process.

The therapeutic importance of gaining a detailed knowledge on insulin fibrillation in relation to type I diabetes has led to intensive studies focusing on its fibrillation kinetics and structural characteristics. Insulin fibrils feature the characteristics that are common to all amyloid fibrils, such as an elongated, unbranched morphology, characteristic cross-β diffraction pattern and

Thioflavin T fluorescence. A full understanding of the fibrillation process requires structural elucidation of every species and determination of the kinetics of the interconversion between species on the reaction pathway. Therefore, we focused on the kinetic of insulin fibrillation that was measured by analysing the intensities of the bands in amide I region of IR spectra that are characteristic of β-sheets. The kinetic parameters were comparable to those obtained with Thioflavin T fluorescent measurements. Additionally, the concentrational measurements suggest that the rate of elongation of fibrils is consistent with a first-order reaction. Infrared spectroscopy allows monitoring the structural changes that occur during fibrillation and consequently, the individual structural component present in evolving equilibrium can be quantified. We observed the melting of α-helix and P_{II} conformation of native insulin with the formation of β-sheets. Two different low-frequency bands in amide I region characteristic of β-sheet provide the insight into the nature of the intermolecular contacts in the insulin fibrils and consequently into the morphology of the fibrils.

Povzetek

Proteinsko zvijanje opisuje prehod med neurejeno in urejeno strukturo, v kateri polipeptidna veriga doseže nativno, tri-dimenzionalno obliko s pripadajočo funkcijo. Podrobno razumevanje gonilnih sil, ki sodelujejo pri proteinskem zvijanju, kot so konformacijske preference aminokislinskega ostanka, vpliv sosednjega aminokislinskega ostanka in vpliv topila, ostajajo pomemben izziv, s katerim bi dobili vpogled v energijske stopnje proteinskega zvijanja. Poznavanje dejavnikov, ki prispevajo h konformacijski stabilnosti aminokislinskih ostankov, bi omogočilo vpogled v molekularne osnove nezvitega stanja, kot tudi v začetne dogodke, ki se zgodijo na poti proteinskega zvijanja oziroma napačnega zvijanja proteinov. Za študijo konformacijskih preferenc glavne peptidne verige nezvitega peptida smo uporabili kratke peptide alaninov. Sprememba okolja peptida je vplivala na težnjo tvorbe intra- ali intermolekularnie vodikove vezi. Z uporabo IR, Ramanske, VCD, NMR in UV-CD spektroskopije smo določili porazdelitev konformacij v vodnem in nevodnem okolju. Opazili smo visoko občutljivost amidne III regije, kjer imajo posamezne konformacije karakteristične frekvence trakov in karakteristične oblike vrhov v primeru VCD spektra. Pri alanin dipeptidu v vodi je bila z DFT računanjem ugotovljena stabilizacija P_{II} konformacije z usmerjenimi vodnimi mostički med C=O in NH peptidnimi skupinami. V primeru polarnih topil se je ujemanje med eksperimentalnimi in izračunanimi spektri močno izboljšalo z dodatkom eksplicitnih molekul topila. Pri kratkih peptidih alanina je bilo ugotovljeno dobro ujemanje med stabilizacijo konformacije s topilom in modelom elektrostatskega senčenja, kjer je konformacija močno odvisna od lokalne elektrostatske energije in senčenjem le-te s topilom. Z daljšanjem peptidne verige je bil opažen efekt sosednjega aminokislinskega ostanka, ki je zmanjšal P_{II} populacijo in ojačal ugodne elektrostatske interakcije glavne verige kratih peptidov alanina.

Napačno zvijanje proteina lahko vodi do agregacije proteinov, kar lahko posledično privede do nastanka amiloidnih fibril, ki povzročajo smrtonosne bolezni. Mehanizem napačnega

zvijanja proteinov in njihova agregacija v fibrile še ni povsem razjasnjena. Spektroskopske raziskave konformacijskih sprememb, ki se zgodijo tekom fibrilacije, pojasnjujejo odnos med spektralnimi lastnostmi in reakcijskimi koordinatami proteinskega zvijanja. Z uporabo infrardeče spektroskopije smo razvili vpogled v sekundarno strukturo peptida tako, da smo določili asignacijo amidnih trakov različnih konformacij modelnega peptida poli-L-lizina (PLL). PLL v vodi pri nizkih pH vrednostih v večini zavzame P_{II} in β strukture, medtem ko se pri višjih pH vrednostih in nizkih temperaturah pojavi karakteristični vrh za α-helično konformacijo. Višje temperature inducirajo tvorbo β struktur, ki pa so hkrati tudi komponente amiloidnih fibril. Za določevanje asignacije vrhov v infrardečih spektrih so bila uporabljena različna topila, ki selektivno stabilizirajo določeno konformacijo. Med vsemi topili je le etilen glikol spodbudil nastanek P_{II}-heliksa, kar nakazuje na drugačno razlago P_{II} stabilizacije, kot je bila predvidena do sedaj, in ni vezana na prisotnost vode ali nabitih stranskih skupin. S pridobljeno asignacijo spektralnih vrhov smo sledili mehanizmu fibrilacije PLL-ja. Uporabili smo diferenčno spektroskopijo, s katero smo dobili informacije o strukturnih spremembah med toplotno inducirano PLL fibrilacijo. Spektralne spremembe pod temperaturo prehoda so bile pripisane taljenju α-heliksa. Vrh, ki predstavlja P_{II}-heliks, je začel izgubljati intenziteto malo pred začetkom PLL fibrilacije. To nakazuje na njegovo vlogo intermediata v procesu fibrilacije. Na kinetiko PLL fibrilacije smo vplivali z dodanim promotorjem PLL fibrilacije, kot je sol $NaClO_4$, ali inhibitorjem PLL fibrilacije, kot so DPPA+DPPC vezikli. Sol $NaClO_4$ zniža temperaturo prehoda iz α-helične strukture do β-ploskev fibril za 10 °C. Mehanizem delovanja soli smo razložili s stabilizacijo P_{II}-heliksa. Dodatek DPPA+DPPC veziklov je stabiliziral α-heliks in P_{II}-heliks PLL-ja. Ta interakcija je hkrati destabilizirala vezikle. Po temperaturi taljenja veziklov se je pojavil spektralni vrh, karakterističen za β-trak, kateremu je sledil vrh, karakterističen za agregirane β-ploskve. Predlagali smo mehanizem nastanka PLL amiloidnih fibril, kjer se α-heliks stali v P_{II}-heliks, kateremu sledi tvorba β-traku, ki je gradnik β-ploskve. Pokazali smo, da je diferenčna infrardeča spektroskopija primerna metoda za pridobivanje strukturnih lastnosti intermediatov tekom procesa fibrilacije.

Terapevtski pomen za pridobivanje poglobljenega znanja o fibrilaciji inzulina v povezavi s sladkorno boleznijo tipa I je pripeljal do študij, ki se osredotočajo na kinetiko fibrilacije in strukturne lastnosti. Fibrile inzulina kažejo na lastnosti, ki so skupne vsem amiloidnim vlaknom, to so podolgovata in nerazvejana morfologija, značilen difrakcijski vzorec prečnih β-ploskev ter Tioflavin T fluorescenca. Za boljše razumevanje fibrilacijskega procesa je potreben opis strukturnih podrobnosti vseh oblik in določitev kinetike interkonverzije med

posameznimi oblikami tekom reakcijske poti. Zato smo študijo osredotočili na kinetiko fibrilacije inzulina, ki je bila merjena z analizo intenzitet spektralnih vrhov v amidni I regiji infrardečih spektrov, ki so značilni za β-ploskve. Kinetične parametre smo lahko primerjali s tistimi, ki smo jih pridobili s Tioflavin T fluorescenco. Koncentracijske meritve so nakazale, da je hitrost elongacije fibril v skladu z reakcijo prvega reda. Infrardeča spektroskopija omogoča spremljanje strukturnih sprememb, ki se zgodijo med fibrilacijo, in tako omogoča kvantifikacijo individualne strukture znotraj razvijajočega ravnotežja. Opazili smo taljenje α-heliksa in P_{II} konformacije nativnega inzulina in tvorbo β-ploskev. Nizko-frekvenčna spektralna vrhova v amidni I regiji, ki opisujeta β-ploskve, omogočata vpogled v naravo intermolekularnih stikov v fibrilah inzulina in s tem v morfologijo fibril.

Abbreviations

α_R	=	Right Handed Alpha Conformation
β	=	Beta Conformation
ϕ	=	Dihedral Angle Phi
ψ	=	Dihedral Angle Psi
3D	=	Three-dimensional
A	=	Alanine
AFM	=	Atomic Force Microscopy
ATR	=	Attenuated Total Reflectance
D_2O	=	Deuterium Dioxide
Da	=	Dalton
DFT	=	Density Functional Theory
DMSO	=	Dimethyl Sulfoxide
ESM	=	Electrostatic Screening Model
FTIR	=	Fourier Transform Infrared Spectroscopy
IPH	=	Isolated Pair Hypothesis
MD	=	Molecular Dynamic
$NaClO_4$	=	Sodium Perchlorate
NMR	=	Nuclear Magnetic Resonance
NOE	=	Nuclear Overhauser Enhancement
NOESY	=	Nuclear Overhauser Enhancement Spectroscopy
PDB	=	Protein Data Bank
P_{II}	=	Polyproline II
PLL	=	Poly-L-lysine
PLP	=	Poly-L-proline
PLQ	=	Poly-L-glumatine
R_g	=	Radius of Gyration
ROA	=	Raman Optical Activity
SFE	=	Solvation Free Energy
TK hypothesis	=	Tiffany and Krimm hypothesis
UV-CD	=	Ultraviolet Circular Dichroism
VCD	=	Vibrational Circular Dichroism

1 Introduction

1.1 Protein and Protein Folding

A protein is a polypeptide chain consisting of a sequence of amino acids that are covalently linked together with peptide bonds. There are 20 naturally occurred α-L-amino acids which serve as building blocks for all proteins found in living organisms. Proteins are involved in almost every biological process in a cellular life where their variety of functions relates to their diverse properties and structures.

Protein folding is a dynamic process described by formation of a native, uniquely ordered, three-dimensional (3D) structure of a polypeptide chain. The principles which govern the folding process, including a detailed understanding of how the amino acid sequence codes for the 3D native structure, has been a major objective of experimental and theoretical studies for more than half a century. Understanding the protein folding mechanism and forces that govern this process will enable the detailed insight into the origin of many diseases connected to protein misfolding and aggregation.

According to Anfinsen thermodynamic hypothesis, the protein folding is described by the simple two-state reaction between the unfolded or denaturated (U) and folded or native (N) state in such a way that ultimately yields to a structure with a minimum Gibbs energy (ΔG) [1, 2]. Moreover, in Anfinsen's thermodynamic hypothesis of protein folding, the native conformation is encoded by the protein's amino acid sequence [3].

$$\mathrm{U} \leftrightharpoons \mathrm{N} \tag{1}$$

1.1.1 Levinthal Paradox and Classical View of Protein Folding

In contrary to the thermodynamic hypothesis of protein folding, Levinthal paradox illustrates the protein folding problem when the protein folding process is described as a random search towards the most stable (native) conformation. Levinthal assumed that if each residue in protein of 100 amino acids can sample only three distinct conformations, the total sum of conformations is $3^{100} \cong 10^{47}$, and if it is assumed that each conformation can be sampled in 1 picosecond, then it would take for protein to fold after a random walk through the energy landscape more than the time of the universe [4]. In reality, proteins typically fold in the microsecond to second time scale. This apparent paradox originally described by Levinthal clearly indicates that the folding process does not occur through a random search of conformations, rather it proceeds along a defined pathway, which dominates the classical view of protein folding [5]. In this view, proteins fold kinetically, that is via a series of sequential steps, to ultimately attain their native 3D structure with a minimum ΔG.

1.1.2 The New View of Protein Folding: Energy Landscape and Folding Funnel

To elucidate the protein folding mechanisms, understanding of energetic states of protein folding motifs and the sequence, specific phenomena need to be resolved. In this regard, the free energy of folding landscape and folding funnel has been introduced to help visualize the protein folding process. In this new view of protein folding, proteins on top of the funnel possess unfolded state with high entropy. Following a gradient down the slope of the energy landscape, in which myriads local minima of partially folded conformations exist, a protein finally reaches the native state at the global free energy minimum of the funnel (Figure 1). During folding, an unfolded polypeptide chain accumulates favourable interactions that lower its energy, which promote a more compact structure and reduce the conformational entropy.
Major forces that govern the process of protein folding and reduce a conformational search in folding funnel are still a matter of debate. Hydrophobic free energy [6, 7], electrostatic interactions [8-10], van de Waals interactions [11, 12], hydrogen bonds (H-bonds) [13, 14] and solvation are the energetic factors that predominate, although the extent of which each contributes to the folding process is unknown. Much of the argument in favour of H-bonding contributing to the folding comes from the results that state that intrapeptide H-bonding is favoured over the peptide-water H-bonds [15-17].

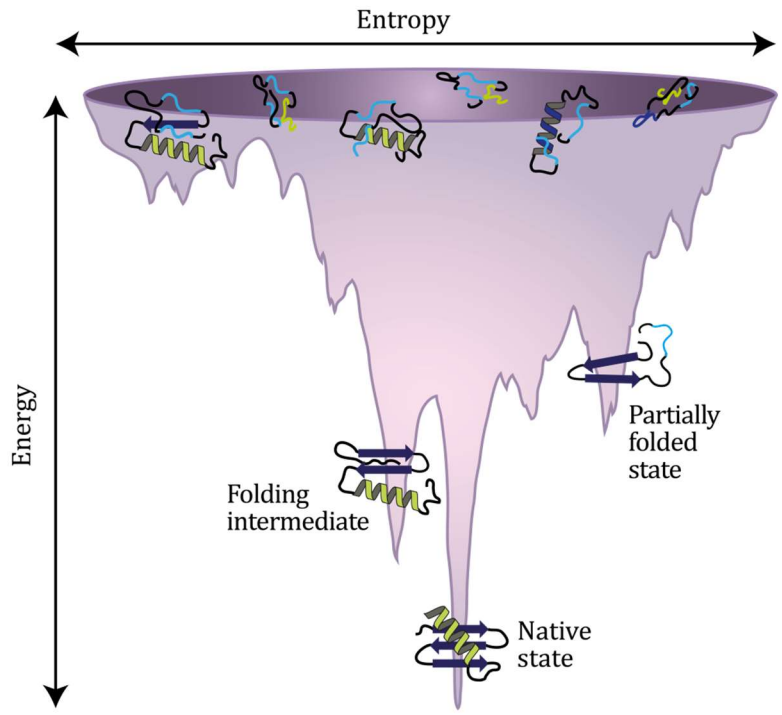


Figure 1: The schematic representation of an energy landscape of a protein folding represented as folding funnel, where the native state is located at the free energy minimum of the funnel.

Recently, the funnel concept of protein folding has been modified to include the protein aggregation, which is now considered as a generic feature of all proteins and peptides. The fact that the native protein structure is thermodynamically stable does not necessarily imply that there is no other conformation, which exhibits a lower free energy than the native state [18, 19]. In many cases, only small variations in environment are required to disrupt the natural stabilizing intrachain contacts. As a consequence the protein misfolds, which leads to the formation of the aggregates. The aggregation may be driven thermodynamically by new H-bonding possibilities or by the hydrophobic effect, the need to shield hydrophobic parts of proteins from the aqueous environment. Furthermore, the formation of proteins in aggregated state, A, is today considered as a process that competes with the natural folding reaction [20].

$$\begin{array}{ccc} \mathrm{N} & \leftrightharpoons & \mathrm{U} \\ & \searrow \quad \swarrow & \\ & \mathrm{A} & \end{array} \tag{2}$$

The fact that natively folded proteins can misfold and interact with one another into well-ordered aggregates is well established. It was even suggested by Gazit [21] that the aggregated state may be the most stable proteins structure, consistent with the energy funnel

depicted in Figure 2. Failure to fold correctly gives rise to the serious malfunction of living system and therefore to lethal diseases.

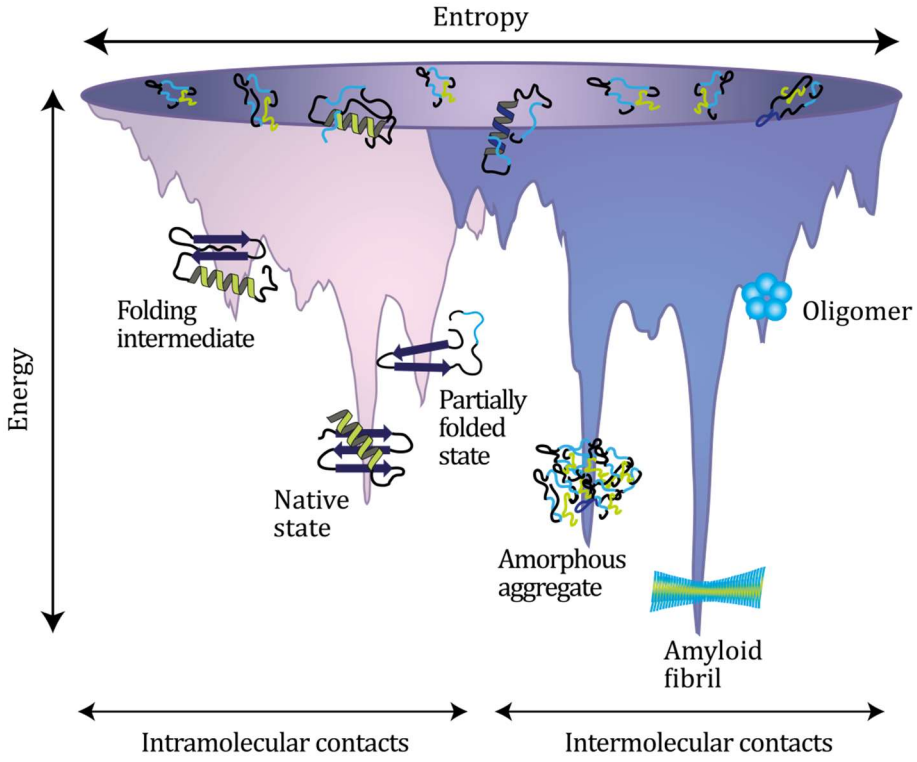


Figure 2: The schematic representation of energy landscape of a protein folding and misfolding followed by the aggregation with a lower ΔG as a native state.

1.2 Unfolded State

The extent of interest in studying the propensities of backbone conformation in unfolded states is growing due to the discovery of naturally disordered, yet biologically functioning proteins and peptides [22], as well as to achieve a complete picture of the folding process. Therefore, it is important to understand the nature of the local structural order in unfolded peptides and proteins that is represented as a collection of conformations populated under extreme non-native conditions, including high and low temperatures, extremes of pH, high pressure and high concentrations of denaturant [23]. The structure of unfolded proteins was first described by Tanford [24] who related the radius of gyration (R_g) of polymeric random coils to its length in monomer units and interpolate this theory to the proteins in strong denaturant by the following equation:

$$R_g = R_o N^{\nu} \tag{3}$$

where N is the number of monomer in the chain, R_o is a measure of the excluded volume of the polymer chain, and ν is an exponential scaling factor which depends on the quality of the solvent (ν = 0.33 for a poor solvent that promotes chain contraction by favouring intrachain interactions, ν = 0.5 for an ideal solvent and ν = 0.6 for a good solvent that promotes chain expansion by favouring chain-solvent interactions) [25, 26]. Typically, ν = 0.6 and R_o = 1.93 are used in the treatment of unfolded and denatured polypeptides [27].

1.2.1 Random-coil or Statistical-coil Model

The properties of the polypeptide backbone in water revealed that the unfolded protein chains should correspond to an extended random coil due to the combined effects of dipolar interactions, steric repulsion and excluded volume [28]. This pioneering work led Paul J. Flory to develop a statistical model for unfolded proteins, now known as the random- or statistical-coil model [28], where all the accessible backbone conformations of a random coil are separated by barriers of the order of thermal energy. The random coil model further assumes the absence of correlations between the local interactions of two monomers of the polymer chain at any length scale, thus leading to the inference that proteins under denaturing conditions, on an average, are devoid of any structure. Important evidence supporting the random coil model for the denaturated state of proteins came from measurements in high concentration of denaturants demonstrating the chain length dependence of R_g yields ν near 0.6 [25, 29-31].

The statistical-coil theory of unfolded polypeptides is based on the assumption that each pair of the dihedral angles phi and psi (ϕ, ψ) is independent of the neighbouring 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 IPH may be formulated mathematically as [32]:

$$S_n \leq n\, S_{n,\mathrm{I}} \tag{4}$$

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.

1.2.2 Tiffany and Krimm Hypothesis

There is now increasing evidence for the existence of significant conformational biases, mostly at the local level, in denaturated proteins and models for unfolded proteins under folding conditions. It has been suggested that the local structural order may guide a polypeptide chain from the unfolded to the native state [33, 34]. In this context, the possible existence of local residue structure would certainly affect the initial phase of the folding process [35]. The existence of such locally ordered segments was first proposed in 1968 by Tiffany and Krimm based on UV-circular dichroism (UV-CD) spectrum of poly-L-lysine (PLL) and poly-L-glutamic acid, at pH values where the side chains were charged, that resemble the UV-CD spectrum of poly-L-proline, which was known to adopt an extended left-handed helix, known as poly-L-proline II helix (P_{II}) [36]. Thus, the Tiffany and Krimm hypothesis (TK hypothesis) claims that unfolded peptides exhibit a large amount of P_{II} conformation. In the early 1990's TK hypothesis has been rekindled by Dukor and Keiderling [37] with the vibrational circular dichroism (VCD) studies of proline oligomers. Since then, various experimental studies have shown that small peptides, which are too short to form any structure that is stabilized by long-range peptide H-bonds, sample the P_{II} conformation in water [38-45].

1.2.2.1 Polyproline II Conformation

The P_{II} conformation is characterized by the lack of intramolecular hydrogen bonds. The factors that stabilize P_{II} conformation are not well established [46]. It was proposed that P_{II}-helical conformation arises from combination of minimization of unfavourable intrachain steric interactions and favourable backbone solvation [47, 48], even though the role of water remains debated. The resulting opened and rather flexible structure is becoming increasingly recognized as a general structural characteristics of natively unfolded proteins [41, 44] and as a candidate for the intermediate structure in amyloid fibril formation [49, 50]. Moreover, it has been shown that the P_{II} structure is one of the most preferential conformations found in dipeptides in water [45, 51]. However, the doubts of several consecutive residues forming the P_{II}-helix as secondary structure element have emerged from the suggestions that formation of P_{II}-helix is noncooperative [52].

1.2.3 Models for Unfolded State

In addition to the charged homopolypeptides, unfolded and partially folded states of peptides and proteins can be studied with short peptides that lack of secondary structure [45, 53], denatured proteins [24, 54], natively unfolded proteins [55], or coil libraries of dihedral angles (ϕ, ψ) sampled within residues that are not part of the regular secondary structures in proteins from the Protein Data Bank (PDB) [56, 57]. The coil libraries constructed and investigated by Swindells et al. [58], Serrano [57] and Avbelj and Baldwin [56] suggest that different amino acid residues have very different structural propensities in the unfolded states. The coil libraries represent the intrinsic conformational propensities of all of the natural amino acid residues that are similar to those found in short peptides. Grdadolnik et al. [45] showed relative populations of the three major backbone conformations in 19 dipeptides in water: polyproline II (P_{II}), beta (β) and right handed alpha (α_R) conformation obtained by infrared (IR) and Raman spectroscopy.

Moreover, a growing number of experimental studies have shown that unfolded proteins display a certain degree of local structural order even under most severe denaturing conditions [54], although some suggest that denatured proteins exhibit random coil properties [25]. However, Whittington et al. have shown that chemical denaturants are promoting P_{II}-helix formation in peptides and proteins and thus the denaturated proteins cannot be used as a models for unfolded proteins under folding condition [59]. Characterizing the unfolded state under native conditions is a major challenge in protein folding research.

1.2.3.1 Short Alanine Peptides

When alanine peptides are too short to form α-helices or any other stable elements of secondary structure and are, in essence, intrinsically unfolded, they can be used as convenient model systems to study the unfolded state under physiological conditions [40, 60]. Moreover, the interest of the protein/peptide folding community has focused on the unfolded state of alanine peptides, after the experimental and theoretical studies suggest that the unfolded state of short alanine peptides cannot be described by the classical statistical coil model of Flory [28]. In particular, di- [45], tri- [38, 61] and tetra-alanine [62] have recently been suggested to adopt predominant fractions of P_{II}-like conformations in aqueous solution. Given that alanine is essentially just backbone, this implies that the polypeptide backbone has some propensity towards adopting the P_{II} helical conformation, at least in the context of short peptides. This propensity for alanine, and the backbone, to adopt the P_{II} helical conformation arises from

combination of minimization of unfavourable intrachain steric interactions [47] and favourable backbone solvation [48]. However, many controversial reports have emerged regarding the conformational sampling of alanine residues in unfolded peptides and its stabilization with solvent molecules. The combination of various techniques like NMR and vibrational spectroscopy as well as theoretical methods have been introduced to determine the alanine backbone preference, especially the alanine propensity to adopt P_{II} conformation in several model systems:

a) Hepta-alanine polypeptide Ac-$X_2A_7O_2$-NH_2 (XAO, X and O denote diaminobutyric acid and ornithine, respectively): XAO model system was used in studies of Kallenbach and co-workers [40], Barron and co-workers [60] and Schweitzer-Stenner and co-workers [63] who all showed that the alanine residues in a XAO predominantly populate a P_{II} conformation whereas Pande and co-workers [64] and Scheraga and co-workers [65] showed conflicting results concerning the sampling of P_{II} in the XAO. They concluded that the P_{II} is only one of many conformations sampled by the alanine in XAO.

b) Host-guest model system: Proline-based host peptide system P_3AP_3 was used by Creamer and co-workers [66], who showed that central alanine residue possesses 61 % of P_{II} conformation. Hagarman et al. used GXG tripeptide host-guest system [67] and reported 79 % of P_{II} for the central alanine. Similar population of the P_{II} conformation for the alanine was proposed with the model system pentapeptide, GGAGG, that was introduced by Kallenbach and co-workers [68].

c) Short all-alanine peptides with blocked or charged terminus: Nearly ten years ago Wouterson et al. estimated, using nonlinear time-resolved vibrational spectroscopy, that trialanine populates P_{II} with a mole fraction 0.8 [69]. Additionally, Graf et al. showed that the central residue in trialanine populates mostly a P_{II}-like conformation with a mole fraction of 0.90 [70]. They extended the study to longer alanine oligomers peptides and confirmed that alanine preferentially adopts dihedral angles in the P_{II} region of the Ramachandran space. However, the authors observed that P_{II} mole fractions for alanine are independent of the number of residues. In contrary to these results, Eker et al. showed that trialanine exhibits a 50:50 mixture of P_{II} and β, while the alanines in the tetraalanine peptide prefer P_{II} [61, 62, 71, 72] indicating P_{II} propensity of alanine increases with an increasing number of residues. An even shorter unit, namely the alanine dipeptide, showed high P_{II} population (61 %), followed by β (29 %) and right handed α_R conformation (10 %) determined by IR and Raman spectroscopy [45].

d) The coil libraries constructed and investigated by Avbelj and Baldwin [56], Serrano [57] and Jha et al. [73] show high P_{II} propensity for the alanine residue. However, if an unrestricted protein database is used, alanine residue exhibits a clear propensity for right handed helical conformation [74]. Distributions derived from coil libraries seem to support a P_{II} preference of the alanine residue, but this depends on how the data set is chosen [57].
e) Theoretical simulations: Different force fields applied in molecular dynamics (MD) simulations of the alanine dipeptide and polyalanines of varying lengths lead to striking differences in the predictions of preferential conformations. Some of these simulations reproduce the high P_{II} propensity of alanine [75, 76], while others do not even yield a substantial population of this conformation [64, 76]. Moreover, nearly all MD simulations of alanine peptides indicate an additional sampling of α_R conformations [74, 77, 78], which is generally not considered in the analysis of experimental data [62, 79]. Han et al. [80] preformed DFT calculations on the alanine dipeptide in explicit water and found nearly isoenergetic energies for P_{II} and right handed α-helical conformation. Pappu and co-workers [47, 48] have shown that the propensity for alanine to adopt the P_{II}-helical conformation arises from combination of minimization of unfavourable intrachain steric interactions and favourable backbone solvation, even though the role of water remains debated. Han et al. [80] showed that water molecules can form a H-bonding bridge between adjacent C=O and NH groups, thus compensating for the missing intramolecular H-bonding present in helices. However, Drozdov et al. [48] performed a Monte Carlo simulation of alanine dipeptide in explicit water and suggested that water molecules connecting C=O and NH groups of adjacent peptides via H-bonding do not contribute significantly to the solvation free energy of the peptide. These simulations suggested that solvation itself does not yield to a preference of P_{II} over α_R conformation, i.e. the effect of solvation to the stabilization of P_{II} conformation was reduced to the neutralization of electrostatic interactions between non-bonded pairs of atoms. As a consequence, steric interactions were thought to be a dominant factor for the P_{II} conformation.

1.2.3.2 Poly-L-Lysine

Homopolypeptides of charged amino acid residues, such as lysine and glutamic acid, are good models for unfolded state under physiological conditions. When the charge of their side chains is neutralized, they are able to form an amyloid structure and are good model representatives to follow conformational transitions from natively folded peptide to the formation of amyloid fibrils. Because of their non-specific sequence and inability to form

globular structures, they demonstrate a generic property of polypeptide main chain interactions leading to the common structure of amyloid fibrils [81]. The secondary structure of PLL has been characterized by numerous spectroscopic methods such as CD [36, 82], IR [83], NMR [84-86], Raman optical activity [87-89], VCD [90] and most recently by UV resonance Raman spectroscopy [49, 91, 92]. At the neutral and low pH values PLL has positively charged side chain amino groups resulting in electrostatic repulsion between them that allows only extended conformation of PLL. By using UV resonance Raman spectroscopy (UVRS), Mikhonin et al. [93] revealed the conformation of PLL at pH 2 as an equilibrium between P_{II} and the extended β-strand conformations in the ratio 60:40, respectively. By increasing the pH value above pKa of PLL, i.e. 10.5, an α-helix become the predominant conformation of PLL. This conformation can be easily transformed to β-sheet fibrils by heating the peptide. In general amyloid forming proteins tend to interact with acidic phospholipids [94] and there have been many studies of phosphatidic acid membrane interacting with PLL [95-98] where interaction is governed by the negatively charged phospholipid and positively charged PLL.

1.3 Conformational Propensities and Secondary Structure

1.3.1 Ramachandran Plot

A fundamental descriptor of a polypeptide's conformation is determined by its side chain steric constrains on rotation around allowed backbone dihedral ϕ and ψ angles. Such representation got the name Ramachandran plot developed by G. N. Ramachandran in the early 1960's [99]. These angles exhibit a smaller variation between amino acids and mainly populate the same three regions (α_R, β and P_{II}) except for proline and glycine that are quite distinct due to their unusual side chains. Hence, amino acids do not adopt all accessible backbone conformations with equal bias. The conformational preferences of a residue are influenced by the property of the side chain and conformation of the neighbouring residues [100]. These observations of neighbour dependence contradict the Flory isolated pair hypothesis [19], which states that the conformations adopted by any residue are independent of the chemical identity and conformation of its neighbours.

However, the nearest-neighbour effect is absent in dipeptides, leading the conformational preferences of dipeptides represent true intrinsic backbone propensities of residues [45]. The

Ramachandran plot for alanine dipeptide, which shows the sterically allowed ϕ and ψ angles, is shown in Figure 3.

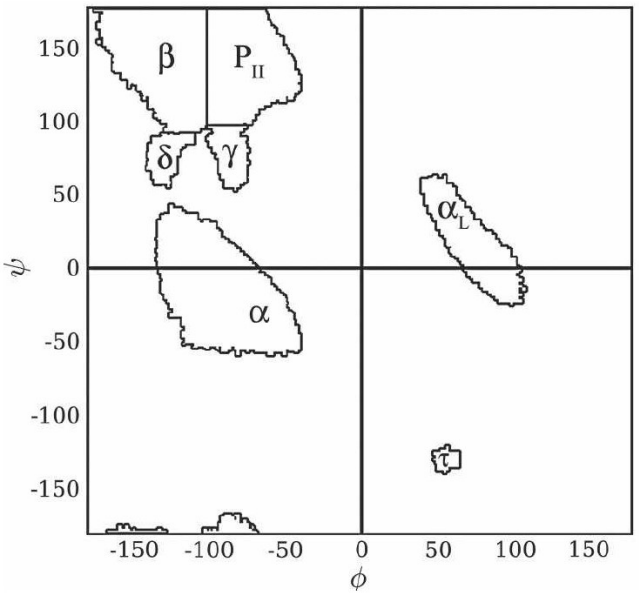


Figure 3: The Ramachandran plot showing the sterically allowed ϕ and ψ angles for alanine dipeptide. The locations of various conformations are indicated. Adopted from Ref. [101].

The polypeptide backbone consists of three degrees of freedom resulting from the three dihedral angles, ω, ϕ and ψ, defined by rotation about the Cα-C'-N-Cα, C'-N-Cα-C' and N-Cα-C'-N backbone atoms, respectively (Figure 4). Peptide bond (H-N–C'-O) exhibits partial double bond character, resulting in very rigid, nearly co-planar arrangement of atoms. Aside from proline, naturally L-amino acids prefer the formation of peptide bonds with a *trans*-conformation, so that ω is constrained to 180°.

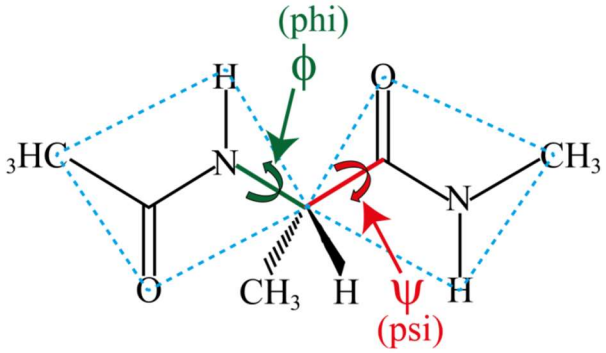


Figure 4: ϕ, ψ and ω angle definitions of N-acetyl-L-alanine-N-methylamide (alanine dipeptide). The ω angle is essentially constant (~180°) for most abundant secondary structures. Thus, the peptide bonds on either side of the central alpha carbon can be treated as rigid plates, which rotate about the dihedral backbone angles ϕ and ψ, indicated by arrows.

1.3.2 Secondary Structures

In order to better understand the process of folding, protein's folded states need to be characterized in greater detail. Polypeptide chain can adopt the specific conformation stabilized by non-covalent interactions – H-bonding between amino and carbonyl groups of the peptide backbone and/or H-bonding between the solvent molecules and the backbone. This is a protein's secondary structure. Most proteins are mainly composed of regular secondary structural elements like α-helices and β-sheets. Each amino acid residue in particular conformation has characteristic pair of dihedral angles (ϕ,ψ) that define its orientation (Table 1).

The α-helix is far the most abundant secondary structure motif in the native proteins. Specifically, the analysis of modern protein databases indicates that 32–38 % of all amino acid residues in globular proteins occur in α-helical conformation. α-helix is stabilized by H-bonding between amide proton of the residue *i* and carbonyl oxygen of residue *i*+4.

β-sheet consists of several strands of extended polypeptide chains which stick to one another via interstrand H-bonding. There are two types of β-sheet structures depending on the mutual orientation of the strands. In the parallel β-sheet the neighbouring strands propagate into the same direction (i.e. direction from N terminus to C terminus are the same), while in the antiparallel β-sheet adjacent chains are aligned in opposite directions.

Polyproline II (P_{II}) helix is an extended left-handed helix, also called 3_1-helix, containing three amino acid residues per helical turn, and therefore exhibits three-fold rotational symmetry. P_{II} conformation is supposed to be stabilized by water H-bonding to backbone amide nitrogen and carbonyl oxygen [61, 102, 103]. P_{II} helix has no intramolecular H-bonds while charge in the peptide dramatically changes its stabilization [104].

2.5_1-helix is stabilized by electrostatic repulsions between ionized side chains and also has no intramolecular H-bonds.

γ-turns occur as one of the two possible enantiomers, called classic and invers, where the inverse ones are more common. Stabilization of these conformations occur with a weak intramolecular H-bonding between C=O group of residue *i* and the NH group of residue *i + 2*, thus connecting 7 atoms (also referred to as a C_7 conformation). A significant number of inverse γ-turns exist as consecutive turns in a structural feature, called the 2.2_7-helix [105].

Table 1: ϕ and ψ Ramachandran angles for common protein and peptide secondary structure motifs.

Secondary Structure	ϕ	ψ
α-helix	-57°	-47°
Antiparallel β-sheet	-139°	135°
Parallel β-sheet	-119°	113°
P_{II}-helix	-75°	145°
2.5_1-helix	-130°	170°
γ-turn	70°	-60°
Inverse γ-turn	-70°	60°

Both intramolecular interactions in protein and intermolecular protein-solvent interactions determine native protein secondary and tertiary structures. For example, intramolecular protein H-bonding stabilizes α-helical and β-sheet conformations while protein-water interactions stabilize peptide backbone in P_{II} conformations [61, 106]. However, even though intramolecular H-bonding is expected to most stabilize α-helix conformation, carbonyl oxygen, engaged in intramolecular bonded N-H--O=C linkages, is capable of forming two H-bonds, also one with a molecule of water [60]. Moreover, water molecule could interact with peptide-peptide H-bond forming a three centred H-bond [107, 108]. All these H-bonding interactions of water molecules additionally stabilize the α-helix conformation and are expected to depend on the capability of peptide side chain to shield the peptide backbone from water. Additionally, Avbelj and Baldwin noticed that solvation strongly affects the preferences for different backbone conformations [109].

1.3.3 Electrostatic Screening Model

Regarding the electrostatic screening model (ESM) the total free energy of an amino acid residue is determined predominately by the local electrostatic energy of the backbone dipole moments (NH, C=O) caused by the interaction with neighbouring peptide groups and by the solvation free energy of the backbone dipole moments, which both depend on the conformations of at least two neighbour residues [8, 110, 111]. In the extended β-strand conformation the C=O and NH dipoles of adjacent peptide units are aligned antiparallel,

whereas in the α-helical conformation the parallel arrangement of dipole moments results in favourable interactions with polar solvent leading to larger backbone solvation for α than for β conformers. Furthermore, screening of backbone electrostatic interactions by side chains or by changing local environment and inter- and/or intramolecular interactions causes different conformational preferences of amino acid residues in aqueous solution. Therefore, already small variations in pH, temperature and ionic strength may have sufficient potential to induce changes in conformational propensities of amino acids and consequently in secondary structure and further in the ability of protein to aggregate.

1.4 Misfolding and Amyloid Disease

Amyloid fibrils are pathogenic agents found as initiators of diseases like diabetes type II, Alzheimer's, Parkinson's disease, prion-related diseases and many others pathological states of organisms described as conformational diseases [112-115]. A defining feature of neurodegenerative diseases, systemic and localized amyloidosis is the accumulation of intra- or extra-cellular deposits, consisting primarily of amyloid fibrils. 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 2.

Table 2: Human diseases associated with formation of extracellular amyloid deposits or intracellular inclusions with amyloid-like characteristics. Adapted from Ref. [116].

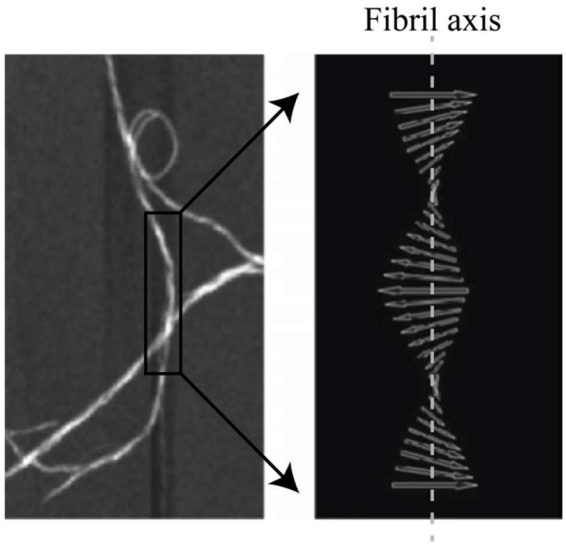
Disease	**Aggregating protein or peptide**	**Number of residues**	**Native structure of protein or peptide**
Neurodegenerative diseases			
Alzheimer's disease	Amyloid β peptide	40 or 42	Natively unfolded
Spongiform encephalopathies	Prion protein or fragments thereof	253	Natively unfolded (residues 1-120), α-helical (residues 121-230)
Parkinson's disease	α-Synuclein	140	Natively unfolded
Frontotemporal 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)
Familial British dementia	ABri	23	Natively unfolded
Familial Danish dementia	ADan	23	Natively unfolded
Non-neuropathic systemic amyloidosis			
AL amyloidosis	Immunoglobulin light chains or fragments	»90	All-β, Ig like
AA amyloidosis	Fragments of serum amyloid A protein	76–104	All-α, unknown fold
Senile systemic amyloidosis	Wild-type transthyretin	127	All-β, prealbumin like
Hemodialysis-related amyloidosis	β2-microglobulin	99	All-β, Ig like
ApoAI amyloidosis	N-terminal fragments of apolipoprotein AI	80–93	Natively unfolded
Finnish hereditary amyloidosis	Fragments of gelsolin mutants	71	Natively unfolded
Lysozyme amyloidosis	Mutants of lysozyme	130	α+β, lysozyme fold
Icelandic hereditary cerebral amyloid angiopathy	Mutant of cystatin C	120	α+β, cystatin like
Non-neuropathic localized diseases			
Type II diabetes	Amylin, or islet amyloid polypeptide (IAPP)	37	Natively unfolded
Medullary carinoma of the thyroid	Calcitonin	32	Natively unfolded
Atrial amyloidosis	Atrial natriuretic factor	28	Natively unfolded
Injection-localized amyloidosis	Insulin	21 + 30	All-α, insulin like
Aortic medial amyloidosis	Medin	50	Unknown
Cataract	γ-Crystallins	Variable	All-β, g-crystallin like
Inclusion-body myositis	Amyloid β peptide	40 or 42	Natively unfolded

Despite extensive research the amyloidosis still remains poorly understood mainly in the view of protein misfolding and fibrillation mechanism. Accumulating evidences suggest that unfolded or misfolded proteins are prerequisite for the initiation of fibrillation, where the initial misfolding of a protein may be driven by specific mutations, misprocessing phenomena, aberrant interactions with metal ions, changes in environmental conditions, such as pH or temperature, or chemical modification (oxidation, proteolysis) [117-119]. The presence of membrane may also play a key role in this process, as lipid bilayers have been shown to influence both the conformation of amyloid peptides and the propensity of amyloid peptides to aggregate. The increased equilibrium populations of molecules in the partially or completely unfolded structures are in the cells usually degraded by the proteasome; when this clearance mechanism is impaired such species often form disordered aggregates or shift equilibrium towards the nucleation of prefibrillar assemblies that eventually grow into mature fibrils.

1.4.1 Structural Features of Amyloid Fibrils

The various peptides and proteins associated with amyloid diseases have no obvious similarities in size, amino acid composition, sequence or structure. However, the amyloid fibrils from different pathologies display many common properties including a core cross-β-sheet structure in which continuous β-sheets are formed with β-strands aligned perpendicular to the fibrils axis [120]. All fibrils have similar morphologies, being long, straight, unbranched assemblies where the thickness of the fibrils is govern by the number of protofilaments. Typically, amyloid fibrils are 6–12 nm in diameter and usually consist of 2–6 protofilaments arranged in a helical manner [121]. Amyloid deposits show specific optical properties on binding dyes molecules such as Congo red and Thioflavin T. Amyloid fibrils have been formed *in vitro* form disease-associated as well as from disease-unrelated proteins. Moreover, there is an increasing belief that the ability to fibrillate is a generic feature of a polypeptide chain, and all proteins are potentially able to form amyloid fibrils under appropriate conditions [81, 122, 123]. Factors including length and sequence may be important in determining details of the structures. Tendency of a polypeptide chain to aggregate rather than fold correctly vary substantially with polypeptide sequence [124] and is determined by the properties like an enhanced β-strand propensity and hydrophobicity and a decreased overall charge.



(a) (b)

Figure 5: Atomic force microscopy image of amyloid fibrils (a) and outline of the ordered core structure (b) of one of the interwoven protofilaments constituting the fibril. In (b), each arrow represents a couple of parallel or antiparallel β-strands. The stacking of β-strand couples generates a double twisted β-sheet that propagates along the fibril axis and whose strands run perpendicular to the latter. Adopted from Ref. [121].

1.4.2 Mechanism of Fibrillation

The formation of aggregates with similar structural features by proteins and peptides of unrelated sequence, suggest a generic mechanism governing the process [125, 126]. In general, the amyloid fibrils do not developed directly from the native conformation, but from precursors that appear to be only partly folded [127]. Thus the process of amyloid formation starts with the conformational change of the native structure leading to favours reciprocal and specific intermolecular interactions, including electrostatic attraction, H-bonding and hydrophobic contacts that are necessary for oligomerization and fibrillation [122, 128-131].

Several models have been proposed to quantitatively describe protein fibrillation, classified into three groups: (a) templating and nucleation; (b) linear, colloid-like assembly of spherical oligomers; and (c) domain-swapping [126]. Generally, the nucleation dependent polymerization model can describe the experimentally observed kinetics of fibrillogenesis of various amyloidogenic proteins [132]. This process is characterized by three phases (see Figure 6) [133]:

(1) *Slow nucleation phase or lag phase*. During this time, the protein undergoes a series of conformational changes and unfavourable association steps to form a partially ordered

oligomeric species, which act as nuclei for fibril formation. The slow nucleation step can be bypassed by introduction of exogenous nuclei or seeds.

(2) *Growth phase*, in which the nucleus rapidly grows to form larger prefibrillar species that eventually fold into the mature fibril. The conversion of protofibrils to fibrils involves inter- and intrafibrillar changes [134, 135]. However, the growth phase does not occur until the amyloid prone species is at concentrations above a certain level (i.e. the critical concentration). Furthermore, increasing protein concentration can reduce the time span before amyloid formation commences.

(3) *A steady state phase*, in which the ordered aggregate and the monomer appear to be in dynamic equilibrium. Additionally, end-to-end annealing and lateral association mechanisms are thought to take place during the maturation of fibrils [136].

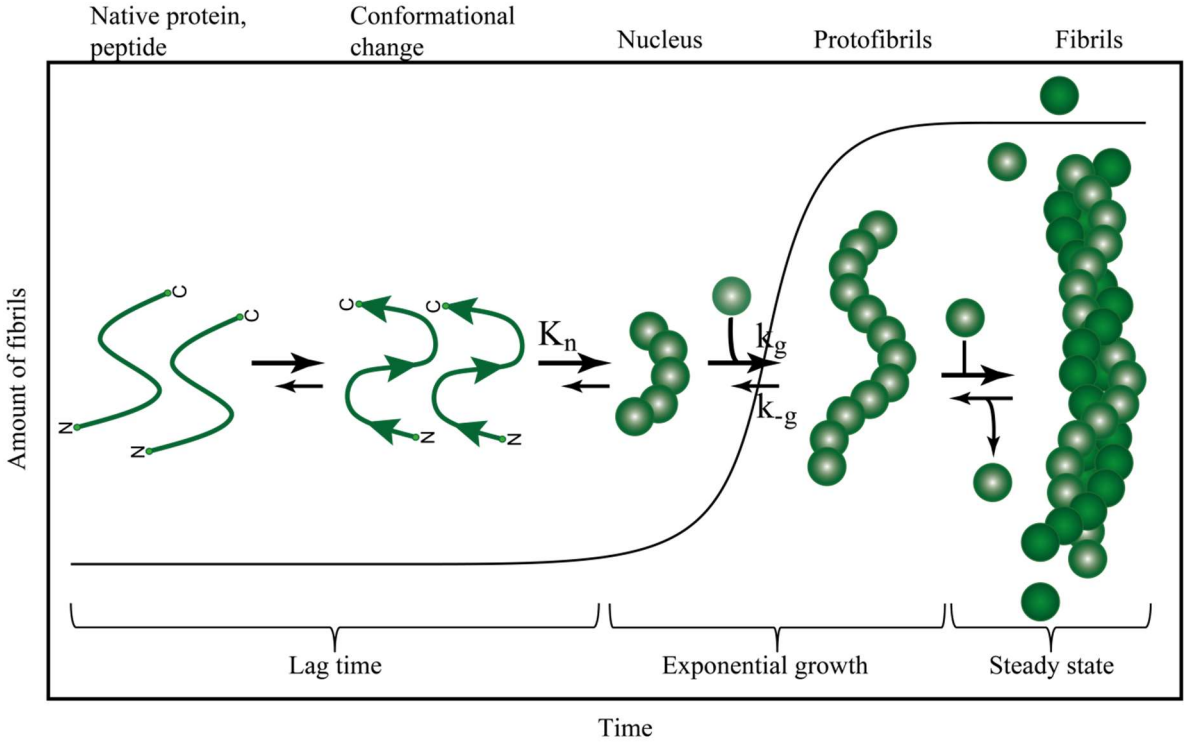


Figure 6: The nucleation-dependent mechanism of fibril formation shows an unfavourable peptide association equilibriums (K_n) accompanied by a structural transition of the peptide monomers, followed by a favourable equilibriums (k_g) that lead to fibril formation.

Even though a simple nucleation dependent polymerization model can describe the overall process of fibril assembling, most of the sub-processes are not understood at the molecular level. The key for a better understanding of fibril formation is the identification and characterization of the slowest or rate-limiting step in the overall process. The nucleus

formation has to overcome two great entropy barriers, namely the protein conformation and association.

1.4.3 Amyloid Oligomers

It is believed that the β-sheet amyloid fibrils are not the cytotoxic species in amyloidogenic diseases, rather the small oligomeric species on the cell membrane were identified as a critical step in amyloid-induced cell death [137]. Oligomers were found to exhibit high levels of cytotoxicity in cell cultures and also have been found to localize in human tissue [138, 139]. Furthermore, the presence of oligomers correlates better with the pathological changes than the insoluble fibrillar deposits do [140]. This indicates a possible origin for the relation between amyloid formation and cellular toxicity, namely the disruption of membrane integrity by oligomeric species. Several possible modes of membrane perturbations offer an explanation for cellular stress, e.g. through loss of chemical potential and compartmentalization [140-142]. In addition, exposure or association with membranes was also found to be correlated with an increased rate of amyloid formation, leading to highly structured fibrillar states [143]. Molecular weights of the oligomeric precursor state are reported to span a range from a few ten to hundreds of kDa, corresponding from dimeric to multimeric oligomers, respectively. In addition, polyclonal antibodies recognize common structural features of oligomers formed from different amyloidogenic proteins [138].

1.4.4 The Role of P_{II} 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 the interactions of exposed hydrophobic regions, which can favour aggregation via non-native interactions. The P_{II} conformation, which is known to be prevalent structure in the short peptides lacking of secondary structure, has been named as the ‘killer’ conformation by Blanch et al. [50]. These authors demonstrated that the prefibrillar species of lysozyme at low pH and a temperature of 57 °C, exhibits characteristics of a P_{II} conformation in its Raman optical activity (ROA) spectrum [50]. The P_{II} 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 intramolecular H-bonds. Another study illustrated the population of P_{II}-like conformations by PLL prior to β-sheet formation [49]. Moreover, many disease-related peptides and proteins, such as tau (Alzheimer’s) [55], Aβ (Alzheimer’s) [144] and α-synuclein

(Parkinson's) [55] have been shown to adopt elements of P_{II} structure in their monomeric states.

In contrast to the suggested precursor role for the P_{II} conformation in fibril formation, other studies suggest that the presence of P_{II} conformations inhibits fibril formation. In the huntingtin 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 poly-L-proline (PLP) residues. Darnell et al. have provided conclusive evidence which suggest that the flanking PLP residues promote the PLQ tracts to form P_{II} helices, which prevent β-sheet and thus fibril formation [145]. 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 P_{II} helices of the central glutamine residues, which then form β-sheet conformations, ultimately leading to the formation of amyloid 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 P_{II} structure [146].

1.4.5 Insulin

Insulin is a 51-residue hormone that regulates the uptake and storage of glucose in the liver and muscles as well as the biosynthesis of triglycerides by fat cells. It is produced in the β-cells of the islets of Langerhans in the pancreas where is stored as a zinc-ligated hexamer microcrystal in acidified secretory granules [147, 148]. Upon release into the blood serum, and triggered by the accompanying pH change, the hexamer dissociates into three dimers and then subsequently into monomers, which is its physiologically active form [149]. Human insulin consists of two polypeptide chains, namely chain A and chain B, which are linked together by disulphide bonds. Whittingham et al. [150] determined the secondary structure of insulin, represented in Figure 7, as 44 % of α-helix, 13 % of β-sheet and remaining amino acids that populate turns and coil.

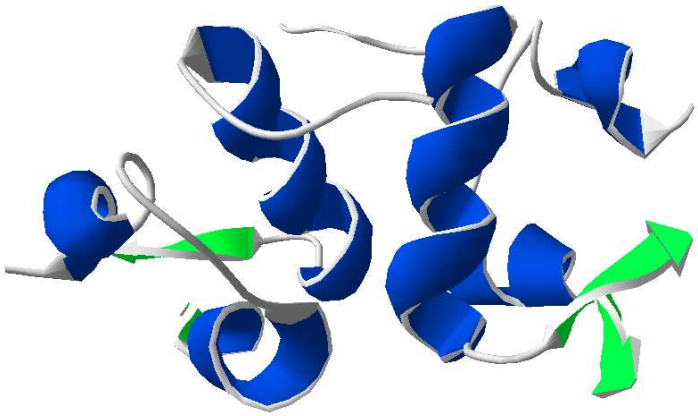


Figure 7: X-ray crystal structure of insulin dimer (PDB 2CEU [150]). Ribbon structure showing the secondary structural motifs in colour: α-helix: blue, β-sheet: green and turns and coils: silver.

Studies on insulin aggregation started in the early years of 20th century, soon after pure insulin was purified by Abel [151]. First it was reported that insulin forms precipitates at high temperature and low pH [152, 153], following the study of Waugh who showed that the precipitates were fibrillar in nature [154] and that the process of fibrillation could be initiated at room temperature with seeding [155]. Since then, a number of investigations have been undertaken to elucidate the mechanism of insulin fibrillation [156-158]. It has been proposed that insulin fibrillation occurs, under conditions of low pH and high temperature, through dissociation of hexamer into monomers which undergoes a structural change to a conformation that have a strong propensity to fibrillate [155, 156, 159]. Fink and co-workers [160] found that both chain-A and chain-B of insulin are capable of forming fibrils individually. The exact mechanism of insulin amyloid fibril formation, however, remains unresolved [161-163]. Despite the vast number of studies about fibrils and prefibrillar structures, only little is known about the primary steps of the aggregation [164, 165]. The main reason for this slow progress in most experimental as well as theoretical approaches is related to the fact that the aggregation mechanism consists of multiple steps and that the rate-limiting transition state is still unknown.

Insulin fibrils are associated with the clinical syndrome injection-localized amyloidosis and are known to deposit in arterial walls and on membrane surfaces. Additionally, amyloid deposits have been observed in patients with type II diabetes after continuous treatment with subcutaneous insulin infusion with micro-pumps and repeated insulin injections [166-169]. Aggregation and fibrillation of insulin represents a major challenge for pharmaceutical production, storage and clinical formulation [170, 171]. The studies of insulin aggregation

therefore have immense application in improving insulin therapy prescribed to diabetic patients. This is important, as the number of people suffering from this disease is immense. In 2010, 285 million people were estimated to be suffering from diabetes [172] and the number is expected to increase for a double till 2030 [173].

2 Aims and Hypothesis

A full understanding of the protein aggregation mechanism requires the knowledge of the kinetic and conformational changes occurring during the protein transition from the native to the aggregated state. To study the thermodynamic driving forces leading to misfolding, oligomerization, aggregation and structural changes of the fibrillation process is the main goal of this thesis.

To follow the structural changes during fibrillation process we applied IR spectroscopy that is a powerful method in structural biology capable of monitoring detailed structural changes in the dynamic equilibrium. To determine the distribution of various types of secondary structure elements, a uniform and detailed assignment of the spectral bands needs to be resolved first. Short alanine peptides were chosen for model molecules because of their simple vibrational spectra that make it easy to compare them with those computed from the models. Literature provide us with conflicting data obtained from the different experimental model systems, different coil library data sets and theoretical simulations that needs further clarification and unification in the view of the conformational population of the alanine residue in the short peptides. Investigation of the conformation of short alanine peptides is difficult due to their rapidly interconverting conformers. To obtain information about such conformational sampling, we employed a various spectroscopic techniques like IR, Raman, VCD, UV-CD, and NMR. These methods are complementary to each other but still each contributes important part toward understanding conformational properties of alanine peptides. We further tried to correlate our experimental results with theoretical simulations using *ab initio* calculations. Our aim was to study the impact of the local electrostatic energy of the peptide backbone and the electrostatic solvation free energy at that site, which depends on the screening interaction of water (solvent) dipoles with the polar peptide group. Additionally, short alanine peptides provide a good model system of unfolded state of protein. The aim of this thesis was to obtain the intrinsic conformational preferences of alanine and how this preference changes with the prolonged peptide chain in the view of neighbour residue effect. The critical role of water has emerged as a factor to condition protein conformation. To test this idea, we calculated the total ground state energy of alanine dipeptide with explicit water

where the optimal bridging of water with the peptide backbone groups determined the conformation preferences. The knowledge of intrinsic propensities will allow computational biochemists to calibrate molecular mechanics force fields so that they can be used for the simulations of unfolded, intrinsically disordered peptides and the formation of secondary structures.

Understanding the mechanism of peptide folding and misfolding at the molecular level is a subject currently under intensive study. We chose the model peptide poly-L-lysine (PLL) because of its non-specific sequence and inability to form globular structures. PLL rapidly forms amyloid fibrils and demonstrates a generic property of polypeptide main chain interactions leading to the common structure of amyloid fibrils. Our goal was to follow the fibrillation pathway of PLL and to stabilize the individual structures with the specific environment change. The mechanism of the self-assembly involves several intermediate states that follow a progressive decrease of the free energy. These states have a very short lifetime, so it is very difficult to study them. In most cases, only the unfolded and the native state are well populated. Our goal was to stabilize the structural intermediate in order to study its conformation.

We studied the effect of different factors (solvent, salt, lipid vesicles) on amyloid fibril formation of PLL, and looked at how these additives cause the specific changes favouring amyloid over native fold. Additionally, we seek to provide further insight into the mechanism of PLL aggregation in the presence of membranes and how this interaction affects on the membrane integrity. It is believed that the β-sheet amyloid fibrils are not the cytotoxic species in amyloidogenic diseases, rather the small oligomeric species on the cell membrane has been identified as a critical step in amyloid-induced cell death. However, although intensively studied, little is known about intermediate states and the formation of early oligomers. So the purpose of this study is also to characterize the interaction of intermediate state of amyloidogenic PLL with the membrane mimicking vesicles.

To follow the fibrillation kinetics we in turn used insulin peptide that forms fibrils under acidic conditions and elevated temperatures and represents a general threat for patients with type I diabetes. Our aim was to probe different spectroscopic methods to determine the kinetic parameters of fibrillation reaction. We studied the effect of peptide concentration on the kinetics of insulin fibrillation. Like PLL, insulin transforms from predominantly α-helical native structure to amyloid fibril, thus we wanted to see if the mechanism of the self-assembly is the same as in PLL.

Amyloid fibrils' formation with similar final structures by proteins of different and unrelated primary sequence leads one to suggest a generic mechanism that governs the process. One further hypothesis suggests that peptide backbone preferences in the unfolded state lead to precursor conformations of the amyloid state. Here, we studied model short alanine peptides, model polypeptide PLL and 51 amino acids long insulin peptide. From the current study we cannot confirm the general hypothesis of a generic mechanism of fibrillation as we did not include longer proteins of different initial conformation. We, however, can gain insight about the role of the backbone preferences.

Our first hypothesis is that, at least in most cases, fibrillation proceeds from misfolded monomers/oligomers that possess the increased population of P_{II} conformation.

Our second hypothesis is that the short alanine peptides have intrinsic backbone preferences that can be influenced by changing the peptide's environment with the effect on the formation of inter- and intramolecular H-bonds.

3 Materials and Methods

3.1 Sample Preparation

3.1.1 Short Alanine Peptides

Blocked alanine peptides (e.g., acetyl-$(Ala)_n$-N-methylamide, n = 1, 2, 3, 4) were purchased from Bachem or were re-blocked from corresponding methylester by using methylamine. In Table 3 the alanine peptides used in our studies are described.

Table 3: The description of short alanine peptides used in our studies with their molecular weights and solubility in different solvents.

Full Name	Abbreviation	Molecular weight	Maximal solubility [M]			
			water	DMSO	TFE	$CDCl_3$
Alanine dipeptide	Ala	144	0.3	0.3	0.3	0.3
Alanine tripeptide	Ala_2	215	0.2	0.14	0.3	/
Alanine tetrapeptide	Ala_3	286	0.1	0.06	0.15	/
Alanine pentapeptide	Ala_4	357	0.002	0.002	0.002	/

When describing particular proton or any other atom in alanine peptide the numbering of individual atom starts with 0 at acetyl group, following the C-terminal end, like represented in the case of Ala_2 in Figure 8.

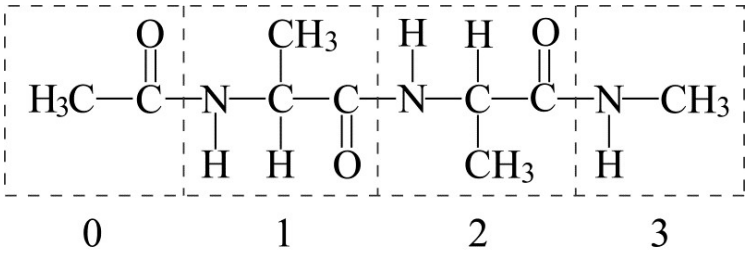


Figure 8: The schematic representation of proton numbering for Ala_2.

For spectroscopic measurements the maximum solubility of each peptide was determined (Table 3) and freshly prepared in H_2O and D_2O, DMSO-d_6, TFE-d_2 and $CDCl_3$ (all Cambridge Isotope Laboratories, Inc.).

3.1.2 Poly-L-Lysine

Poly-L-Lysine (PLL) with HCl (MW_{MALLS} = 31185 Da) was purchased from Sigma and used without further purification. The samples were freshly prepared in H_2O or D_2O (Cambridge Isotope Laboratories, Inc.) as a 4 wt% solution. The pH was adjusted to pH 4 and pH 11.6 using appropriate amounts of HCl (DCl) and NaOH (NaOD). The solutions in DMSO (Sigma-Aldrich), DMSO-d_6 (Cambridge Isotope Laboratories, Inc.), TFE (Sigma-Aldrich) and ethylene glycol (Sigma-Aldrich) were prepared as 4 wt% concentrations. The phospholipid vesicles were prepared by dissolving dipalmitoyl phosphatidic acid (DPPA) and dipalmitoyl phosphatidyl choline (DPPC) (Sigma) in equivalent ratio in chloroform to obtain the homogeneous solution followed by evaporation of solvent under a stream of nitrogen. To remove any residual chloroform the sample was placed in a vacuum desiccator for at least 90 minutes. Afterwards, the lipids were dissolved into water with pH adjusted to 11.6 and sonicated in water bath sonicator at 80 °C for 1 hour to obtain the vesicles. The PLL was added to the vesicles at low temperature 1 hour before measurements and pH value was again adjusted to 11.6.

3.1.3 Insulin

Human insulin (MW = 5700 Da) was purchased from Sigma and used without further purification. The samples were freshly dissolved in H_2O or D_2O (Cambridge Isotope Laboratories, Inc.) to a final concentration of 20 mg/mL in 0.1 M hydrochloric acid (HCl) and adjusted to pH 1.6. Before measurements the sample was filtrated through 0.2 μm filter to remove any aggregates formed during the sample preparation. Fibrillation was induced by heating the insulin solution at 60 °C.

Exposure of proteins to acidic pH and high temperature may result in significant chemical degradation. Deamidation at Asn^{A21} has been shown to be the main degradation for insulin at pH 1.6 [174]. Even though some deamidation is occurring under the conditions studied, it is considered to be of minor importance because the deamidated form of insulin has been shown to have the same fibrillation kinetic properties and structural properties as the native insulin [156].

3.2 Vibrational Spectroscopy

Vibrational spectroscopy is the main spectroscopic method used in the present thesis. It gives information about the vibrational modes of the molecules. Vibrational spectra are the result of transitions between vibrational energy levels that occur when light with a frequency that corresponds to a characteristic frequency of the molecule is absorbed (Figure 9). The Raman effect occurs when molecule is excited by an incident photon form a ground state to a virtual energy state. The molecule then relax and returns to the ground state by emission of a photon whose energy is the same as that of incident one and results in elastic Rayleigh scattering (Figure 9, b). A very small part of the scattered light have the frequency that is smaller than those of the elastically scattered light because a part of the energy of the incoming photons was employed to excite molecules to a higher vibrational state $\nu = 1$. As a result, the emitted photons will be shifted to lower energy (Stokes shift) (Figure 9, c). If the process starts from a vibrationally excited state and relaxes to the ground state, then the emitted photons will be shifted to higher frequency (anti-Stokes shift) [175] (Figure 9, d). Usually radiation occurs in the mid-infrared region which spans from roughly 2.5 to 50 μm or 4000 to 200 cm^{-1}. Energy is transferred from the photon to the molecular vibration, such as a bond stretch or angle bend, changing the amplitude of the vibration. This can only occur if a change in the dipole moment takes place in the case of IR spectroscopy or the change in polarizability occurs for the Raman spectroscopy of the molecule during the vibration.

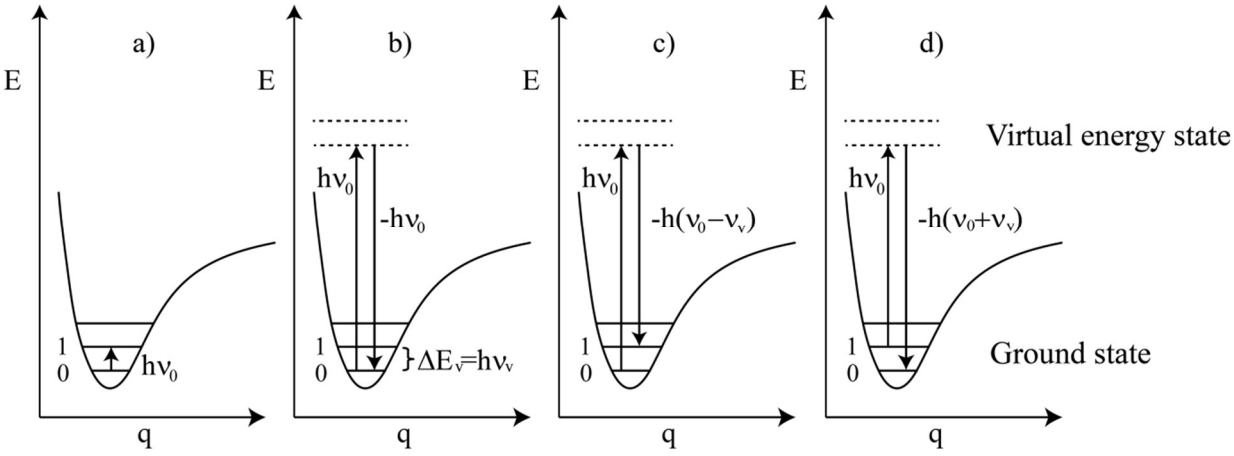


Figure 9: The comparison of a) IR absorption, b) elastic Rayleigh scattering, c) inelastic Stokes scattering and d) inelastic anti-Stokes scattering.

In aqueous solutions the strong absorption of solvent overlaps this region, consequently use of

D_2O instead of H_2O is needed to obtain IR or VCD spectrum in amide I region, being only minimally shifted in frequency.

3.2.1 Vibrational Circular Dichroism

The vibrational circular dichroism (VCD) provides the band shape variability of UV-CD with the frequency resolution of IR where the band shape and its frequency position predict the dominant secondary structural type in peptides and proteins [90]. The technique has been successfully applied to determine the structure and absolute configuration of molecules by absorbing right and left-circularly polarized incident light of chiral molecules like amino acids. Typical VCD band patterns of various types of secondary structure were reported [176]. More recently, VCD spectroscopy has shown sensitivity to amyloid fibrils, and was used to probe fibrillogenesis [177].

3.2.2 Infrared Difference Spectroscopy

The subtraction of two spectra, or difference spectroscopy, is a straightforward way of simplifying an experimental spectrum and/or extracting overlapping bands. Difference spectroscopy is a spectroscopic method, which is sensitive to all types of band changes and is an ideal tool for spectral interpretation [178]. First, the solvent is subtracted from the original spectrum followed by the subtraction of two similar spectra. The subtraction factor is determined by eliminating the bands that corresponds to the CH_2 or CH_3 vibrations that are not affected by a change caused by the particular perturbation that is observed.

3.2.3 Amide Modes

The conformational sensitive regions in the vibrational spectrum of peptides and proteins are amide I (1600–1690 cm^{-1}), amide II (1450–1580 cm^{-1}) and amide III (1200–1350 cm^{-1}). The amide I mode has a large contribution of the backbone C=O stretching and its frequency is very sensitive to changes in conformation, depending on the environment of C=O group, the H-bonding and inter-residue coupling [92, 179]. The amide II mode is less sensitive to the various secondary structure elements. However, its frequency reflects the participation of backbone NH groups in the various types of H-bonds with proton acceptor groups from the protein or molecules from the protein environment. The amide III region is the most sensitive indicator for the peptide conformations. The spectral band results chiefly from N-H in-plane

bending, mixing with C-C and C-N stretching, C=O in-plane bending and C_α-H bending.
IR and Raman spectra of short alanine peptides have substantial different model band intensities in amide I region, therefore the amide III region was used to determine the conformational population. The IR spectroscopy has much higher sensitivity, enabling the measurements of lower peptide concentrations.

3.2.4 Experimental Setup for Vibrational Spectroscopy

The vibrational spectra were measured using PerkinElmer System 2000, Bruker Vertex 80 IR spectrometer, Bruker PMA 50 VCD spectrometer and Bruker Raman spectrometer. Spectra were recorded in the range between 7000 and 450 cm^{-1}. The spectral resolution was 4 cm^{-1} for all measurements. The spectra obtained from the ATR-infrared measurements were first recalculated to get the pure absorption spectra [180]. The subtraction of bulk solvent of the recorded spectrum was done by using Grams (Thermo Electron, San Jose, CA)) or Opus software. The band overlapped regions were analysed using the Grams band fitting procedure by optimising the sum of the bands with the mixed Lorentzian and Gaussian shapes. The IR spectra were measured using diamond ATR cell (Specac) equipped with a heated top plate or transmission cell with a path length of 12 μm. Typically, 256 interferograms were collected and apodized with a triangular function and Blackman-Harris apodization function on PerkinElmer and Bruker spectrometer, respectively. To reduce the strong bands due to the ATR absorption of diamond, backgrounds were collected for each recorded temperature. The Raman spectra were obtained with the 1064 nm excitation from a Nd YAG laser by averaging between 5000 to 20000 scans. For VCD experiment 10 spectra were averaged, each scanned 20 min, to obtained higher signal-to-noise ratio. Samples for VCD measurements were placed into a cell with a path length of 25 μm. The temperature was controlled by thermostating system.

3.3 *Ab Initio* Calculations

The density functional theory (DFT) and Hartree-Fock (HF) calculations of Ala were performed in gas phase, in water, in dimethyl sulfoxide (DMSO) and in chloroform ($CHCl_3$) using Gaussian 09 [181]. The hybrid B3LYP and HF functionals have been used together with the basis set 6-31G(d,p) [182-184]. This level of theory turned out to be a good compromise between the computational cost and accuracy for complexes which exhibit H-bonding. Apart

from the selected explicit molecules of solvent the system environment was modelled by the polarizable continuum model (PCM) [185].

DFT calculations were used to evaluate vibrational properties for each geometry-optimized structure found during conformational analysis, such as the dipole and rotational strengths. With these data, IR and VCD spectra can be obtained for each conformer. Spectra from theoretical models of each conformer were obtained by the Boltzmann averaging over the representative structures of each configuration. Calculated standard ground state energies (ΔG^0) were used to determine the Boltzmann weights:

$$w_A = \frac{e^{-\Delta G_A^0/RT}}{\sum_i e^{-\Delta G_i^0/RT}} \quad (5)$$

Spectra from theoretical models of each conformer were obtained by the Boltzmann averaging over ten representative structures of each conformation. Initial structures were extracted from all-atoms MD simulations performed by Charmm [186] in which dipeptides were constrained to a given conformation.

In practice, the calculated harmonic vibrational frequencies, depending on the computational method, are overestimated according to the fundamentals observed experimentally. To compare the experimental and calculated spectra the standard spectral shift of the calculated spectra was introduced.

Including the explicit solvation molecules to the model reveals detailed information about the hydrogen bonding between the peptide and solvent molecules and furthermore indicate, how these molecules are vibrationally coupled with each other [80]. The number of explicit solvent molecules, introduced in DFT calculations, was capped as low as possible due to the reduction of the computational complexity.

3.3.1 *Ab Initio* Calculations of the pKa

The *ab initio* pKa calculations of propylamine in the different solvents were performed for the gas phase and in solution using Gaussian 09 [181]. The solvation energies have been calculated using the Polarisable Continuum Model (PCM) [187] at the B3LYP/6-31++G(d,p) [182-184] level.

3.4 UV-CD Spectroscopy

In proteins and peptides, the far UV-CD signals arise with absorption of radiation of peptide

bond below 240 nm, where weak but broad n → π* transition centred around 220 nm occurs and more intense π → π* transition around 190 nm takes place. The different types of regular secondary structure found in proteins give rise to characteristic UV-CD spectra. In spite of its convenience and commercial availability, UV-CD spectroscopy allows only rough estimation of protein and peptide secondary structure composition.

UV circular dichroism studies were performed using Aviv model 62 ADS CD spectropolarimeter equipped with a thermoelectric sample holder for temperature control in the cell. The spectra were obtained using a 0.1 mm, 1 mm or 1 cm path length cell. The spectra were measured in the range from 180–260 nm with a scan speed of 500 nm/min and a data pitch of 1 nm. The spectra were obtained by averaging ten scans and were collected as ellipticity as a function of wavelength. Far data analysis, the ellipticity was converted to molar absorptivity per residue using the following relationship:

$$\Delta\varepsilon = \frac{\theta}{32980 \cdot c \cdot l(n-1)} \tag{6}$$

where θ is the ellipticity in [mdeg], c is the concentration in [M], l is the path length in [cm] and n is the number of peptide bonds.

3.5 NMR Spectroscopy

High resolution NMR is a powerful spectroscopic method for determination of conformational and dynamic properties of molecules in solution.

3.5.1 $^3J(H_\alpha,H_N)$ Coupling Constants

To obtained the conformational preferences of short alanine peptides the $^3J(H_\alpha,H_N)$ coupling constants (vicinal coupling constant between $C_\alpha H$ and NH protons) were measured that are directly connected to dihedral angle ϕ by using the Karplus relation [188]:

$$^3J(H_\alpha,H_N) = a\cos^2(\phi - 60) - b\cos(\phi - 60) + c \tag{7}$$

The parameters *a, b* and *c* were determined experimentally by Vuister and Bax [189] with the values: a = 6.51 Hz, b = -1.76 Hz and c = 1.6 Hz. The experimental coupling constant ($^3J_{exp}$) of alanine peptide is used to determine the mole fraction of the β conformation (χ_β) as follows:

$$^{3}J_{exp} = {}^{3}J_{\beta} \cdot \chi_{\beta} + \left(1 - \chi_{\beta}\right) \cdot {}^{3}J_{\alpha/P_{II}/c_7} \quad (8)$$

where $^{3}J_{\beta}$ is the coil library coupling constants for β conformation and $^{3}J_{\alpha/P_{II}/c_7}$ is the average coil library coupling constants of α_R, P_{II} and C_7 conformation. The small difference in dihedral angle ϕ, and therefore in $^{3}J(H_\alpha,H_N)$ coupling constant, of α_R, β and C_7 conformations makes it impossible to distinguish between them with Karplus equation, therefore an average coupling constant of those conformations ($^{3}J_{\alpha/P_{II}/c_7}$) is used (5.14 Hz) to determine their population. Furthermore, the individual ϕ-values are obtained for residues from the coil library of the Protein Data Bank, while the corresponding values of $^{3}J(H_\alpha,H_N)$ are found from the different parameterizations of the Karplus relation for flexible molecules [51]: β(ϕ = -120°, J = 8.5 Hz), P_{II}(ϕ = -75°, J = 5.24 Hz), α_R(ϕ = -60°, J = 5.07 Hz) and C_7(ϕ = -78°, J = 5.1 Hz). Using these values a reasonably good linear correlation (R = 0.82) of the coil library coupling constant versus the fraction of coil library residues in the β-basin confirms that coupling constant significantly depend on the fraction of residues in the β conformation [51]. NMR spectroscopy gives an average value of $^{3}J(H_\alpha,H_N)$ over all coupling constants of different conformations for one amino acid. A NMR study of 19 dipeptides in water showed the $^{3}J(H_\alpha,H_N)$ coupling constants of alanine dipeptide is 6.02 Hz [51].

3.5.2 NOE Distances

The Nuclear Overhauser Effect (NOE) is among the most important NMR parameters for structure determination. The NOE depends on the rate of molecular tumbling (correlation time), which is in large part determined by the molecular weight, and the distance between protons. To obtain the intramolecular distances between two protons, the cross-peaks intensities are typically referenced to a cross-peak with the known reference distance and can be translated into proton-proton distances by the following equation:

$$r_{ij} = r_{ref} \sqrt[6]{\frac{I_{ref}}{I_{ij}}} \quad (9)$$

r_{ref} is fixed distance between two hydrogens that is characteristic for all conformations while I_{ref} represents its measured NOE volume. This equation is valid only at appropriate mixing time. Usually we see the proton-proton distances within 5 Å. The model distances between hydrogens of Ala, Ala_2 and Ala_3 in particular conformation were calculated using Avogadro

program. The restrains were dihedral angles: α_R (ϕ = -60°, ψ = -45°), β (ϕ = -120°, ψ = 120°), P_{II} (ϕ = -75°, ψ = 145°), C_7 (ϕ = -78°, ψ = 65°) and optimized geometry of the peptide. For the comparison of distances obtained for model peptides from Avogadro program and those obtained from NOE values of the measured NMR spectra of alanine peptides we chose those distances that change with conformation. In Table 4 are listed the model distances for Ala_2 that are conformational sensitive.

Table 4: The distances between two hydrogens in model peptide Ala_2 in particular conformation, measured with Avogadro program [190].

Assignment	α_R	C_7	β	P_{II}
$H_N(2) \leftrightarrow C_\alpha H(1)$	3.6	2.5	2.1	2.2
$H_N(3) \leftrightarrow C_\alpha H(2)$	3.6	2.5	2.1	2.1
$H_N(2) \leftrightarrow H_N(3)$	3.0	3.7	4.4	4.6

3.5.3 NMR Experimental Setup

For NMR experiments alanine peptides were dissolved in a mixture of 90 % H_2O and 10 % D_2O at pH 3 with concentration of 20 mM for Ala_2 and Ala_3 and 2 mM for Ala_4. The temperature measurements were performed from 10 °C to 70 °C. One- (1D) and two-dimensional (2D) proton NMR spectra were measured using a Varian INOVA NMR spectrometer (Varian Inc., Palo Alto, CA) operating at a proton frequency of 600 MHz and equipped with a 5 mm HCN triple resonance probe with z-gradients. The coupling constants $^3J(H_\alpha,H_N)$ were measured from 1D-^{1}H spectra from H_N resonances, using band fitting algorithm (GRAMS program, Thermo Electron, San Jose, CA) with Lorenzian line shape to calculate the peak frequency of the H_N doublet. The accuracy of the coupling constants is believed to be ± 0.05 Hz. 2D-^{1}H NOE spectroscopy, NOESY, experiments were recorded in the phase sensitive mode. The mixing times of NOESY spectra were set at 150 ms. Typically, 256 t_1 increments of 2K complex data points over 6 kHz spectral width were collected with 32 scans per t_1 increments and a relaxation delay 1.5 s. All spectra were acquired with the carrier offset placed on the solvent resonance, which was reduced using either solvent presaturation or excitation sculpting. The data sets were processed using the VNMR software package (Varian Inc., Palo Alto, CA). The sequential assignment was made using the $H_{\alpha(i)}$–$H_{N(i+1)}$ cross-peaks of the NOESY spectra.

3.6 Thioflavin T Fluorescence Spectroscopy

Fibril formation was detected with the histological dye Thioflavin T (ThT) that has been widely used for the detection of amyloid fibrils [191, 192]. In the presence of fibrils, ThT displays a dramatic shift of excitation maximum (from 385 nm to 450 nm) and the emission maximum (from 445 nm to 482 nm), leading to easily differentiation between bounded and unbounded ThT [156]. ThT is believed to interact relatively specifically and rapidly with amyloid fibrils, and the binding is independent of the primary structure of the protein [193]. Only the multimeric fibrillar forms, not multiple β-sheet domains in native proteins, fluoresce with ThT. Measuring the level of binding of ThT to insulin fibrils is thus a potential powerful tool for studying the kinetics of insulin fibril formation [156].

Insulin fibril formation using *in situ* ThT fluorescence was performed on Perkin Elmer LS 50B spectrofluorimeter using quartz cuvettes with a 1 cm path length. The excitation wavelength was 482 nm and the emission slits were maintained at 2.5 nm. Insulin was dissolved at the concentration of 1, 5 and 10 mg/mL in 0.1 M HCL, pH 1.6 and incubated at 60 °C. Typically, a portion of 1 μL of amyloid sample was suspended in 999 mL of water (pH= 1.6) with 10 μL of 2 mM ThT (Sigma).

3.7 Transmission Electron Microscopy

The samples of PLL amyloid fibrils were freshly prepared in H_2O (pH 11.6) with heating of 2 wt% PLL for 15 minutes at 70 °C. 10-fold diluted sample was applied to Formvar and carbon-coated grid and left to adsorb for 3 min. The excess of sample was soaked away and stained with 1 % (w/v) water solution of uranyl acetate. The excess of stain was removed immediately. The fibrils were observed with a Philips (Amsterdam, the Netherlands) CM 100 transmission electron microscope. Images were recorded by Bio-scan CCD camera and Digital Micrograph software, (Gatan Inc., Washington, DC, USA).

4 Conformational Properties of Short Alanine Peptides

4.1 Results of Short Alanine Peptides

4.1.1 IR and Raman Spectroscopy of Short Alanine Peptides

4.1.1.1 IR and Raman Spectroscopy of Short Alanine Peptides in Water

The analysis of the amide III region of Ala_2 in water showed the presence of characteristic bands for conformations already found in alanine dipeptide [45]. These band components belong to P_{II} (1305 cm^{-1}), β (1278 cm^{-1}) and α_R (1292 cm^{-1}) conformations, i.e., conformations that all form the intermolecular H-bonding with solvent molecules. In Figure 10 the spectrum of the amide III region of Ala_2 displays an additional band at 1261 cm^{-1} that is not present in the amide III region of Ala in water. This component of amide III vibrational mode is dominant in the IR spectra of Ala in argon matrix and in CCl_4 for which the computational simulations predict the C_7 conformation. The measurement of Ala_2 in deuterium water indicated that bands at 1343 cm^{-1} and 1283 cm^{-1} correspond to $C_\alpha H$ and CH_3 deformation modes. Additionally, the low frequency band at 1245 cm^{-1} in Figure 10 shows an aggregated form of peptide.

In the amide I region of the Raman spectrum of Ala_2 in water several bands can be found (Figure 10, a). The band at 1682 cm^{-1} was assigned to carbonyls that are involved in weak intermolecular H-bonds with the water molecules. The three bands located at 1664 cm^{-1}, 1648 cm^{-1} and 1635 cm^{-1}represent the C=O groups that are involved in stronger interactions with the water molecules, while the band at 1618 cm^{-1} indicates the presence of a strong intramolecular H-bond between the carbonyl group and the proton donor amide group. The candidate for such H-bond is the C_7 conformation that has characteristic vibrational mode in the region between 1625 and 1615 cm^{-1} [194, 195].

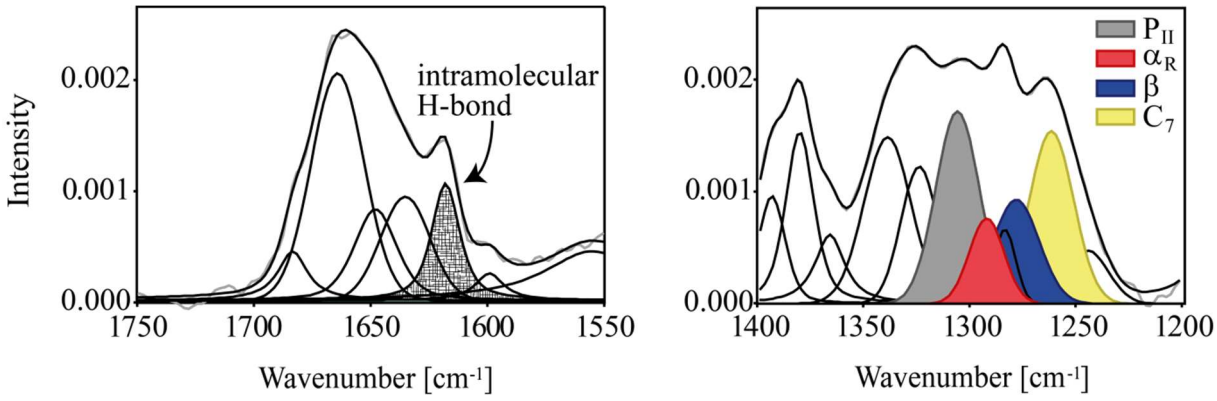


Figure 10: The fitted Raman spectrum in the amide I (a) and amide III (b) region of Ala_2 in water at 0.2 M concentration at room temperature.

Ala_4 has very low solubility in the water (2 mM); therefore only the transmission IR spectroscopy provided a spectrum with sufficient signal-to-noise ratio. From the comparison of the amide III region of IR spectra of all short alanine peptides in water (Figure 11) the increase of low-frequency bands at 1261 cm^{-1} and 1245 cm^{-1} was observed with the longer peptide chain. The first band can be assigned to C_7 conformation, although this assignment is based on computational methods and cannot be finally confirmed. The second band at 1245 cm^{-1}, which is also gaining in the population with the longer peptide chain, represents the aggregated form of peptides. The parallel increase of the band occurs in amide I region of Raman spectra at 1664 cm^{-1} suggesting this band can be assign to the same conformation as one at 1261 cm^{-1}. The band at 1618 cm^{-1} (Figure 10, a) that represents intramolecular H-bonded carbonyl group is the strongest in Ala_2.

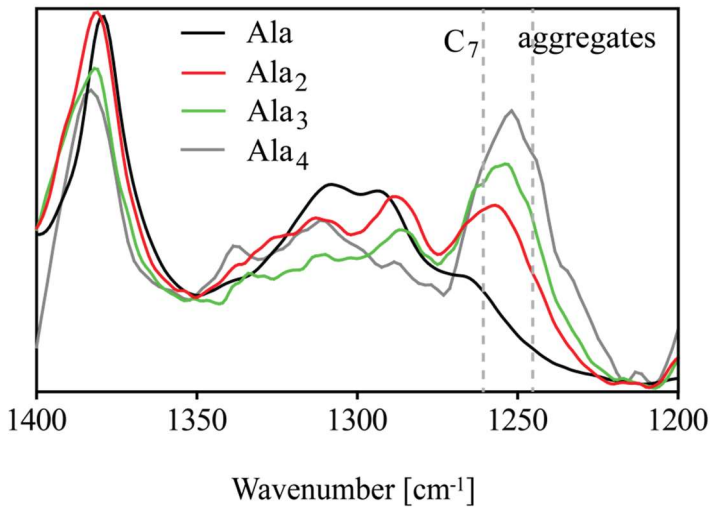


Figure 11: The IR spectra of Ala_n (n = 1, 2, 3, 4) in water at room temperature.

Using the spectral fitting method of the amide III region of IR and Raman spectra the population of each conformation was obtained for all short alanine peptides (Table 5). From the Table 5 it is evident that with increasing of the chain length of the alanine peptide the proportions of P_{II} and β conformation decrease while the increase of the band at 1261 cm^{-1} was observed that was assigned to C_7 conformation.

Table 5: Population of the conformations obtained by fitting Raman and IR spectra in the amide III region of Ala, Ala_2, Ala_3 and Ala_4 in water.

	P_{II}	α_R	β	C_7
Ala[a]	0.60	0.11	0.29	/
Ala_2	0.36	0.12	0.20	0.32
Ala_3	0.31	0.13	0.18	0.38
Ala_4	0.27	0.13	0.15	0.45

[a] Grdadolnik et al. [45]

4.1.1.2 IR Spectroscopy of Ala in $CDCl_3$

The concentrational measurements of Ala in chloroform showed a strong structural dependence of the peptide. The IR spectra of Ala at 5 and 100 mM concentrations displayed low-frequency amide III bands at 1222 and 1235 cm^{-1}, suggesting the formation of oligomer structures of the peptide that mainly comprise of Ala in β aggregated form. With the increasing concentration of Ala the IR spectral bands at 1656 cm^{-1}, 1567 cm^{-1}, 1278 cm^{-1} and 1235 cm^{-1} increase their intensities, while the bands at 1676 cm^{-1}, 1508 cm^{-1} and 1222 cm^{-1} decrease their population (violet bands in Figure 12). The two sets of spectral bands that change their intensity with concentration indicate the presence of two types of small oligomers, one dominated at low concentrations and the other at higher concentrations. Furthermore, analysis of amide III region revealed the difference in conformational assembly of monomer structures at low and high concentration that are in equilibrium with β oligomer structures. At 5 mM concentration the band at 1259 cm^{-1} represents the conformation of monomer that is the most populated at low concentrations. This band was observed in water at longer alanine peptides, in CCl_4 and argon matrix and can be assigned to C_7 conformation [196]. At 100 mM concentration the increase of band at 1278 cm^{-1} was observed, suggesting

the 30 % increase of β conformation of Ala monomer relative to the monomer structure at low concentration. Additionally, the low-frequency band at 1235 cm^{-1} that is assigned for β oligomer at 100 mM concentration has higher wavenumber than the band at 1222 cm^{-1} of β oligomer at 5 mM suggesting weaker H-bonds in oligomer structures at higher concentration.

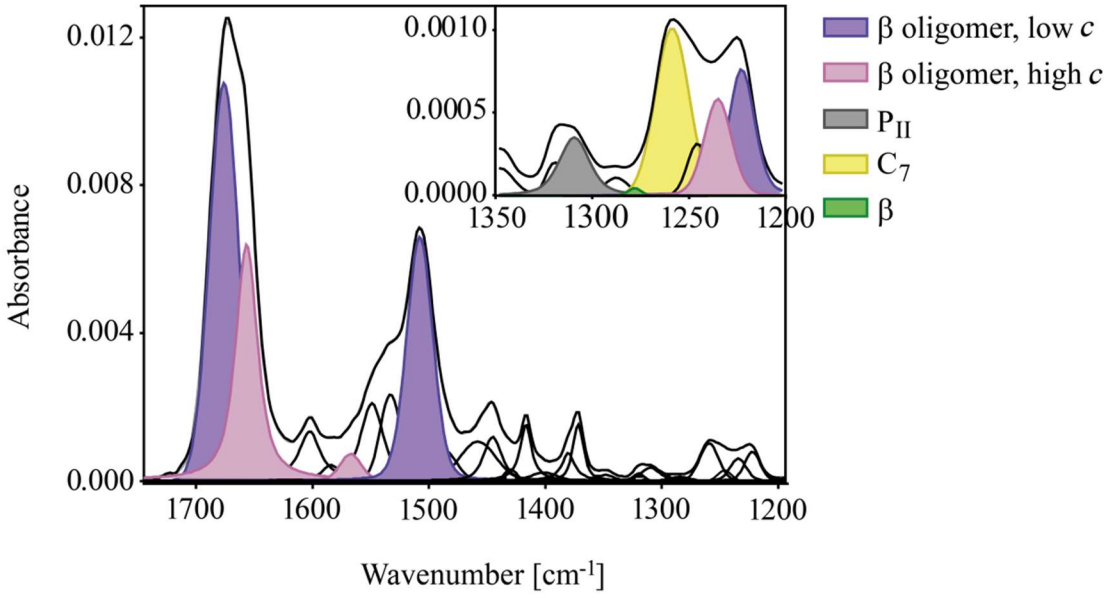


Figure 12: Amide I, amide II and amide III (zoomed in upper left square) region of IR spectrum of Ala in $CDCl_3$ at room temperature at 5 mM concentration.

4.1.1.3 IR Spectroscopy of Short Alanine Peptides in DMSO

Ala was measured in DMSO where mainly β conformation was observed from amide III region in IR spectrum (Figure 13) that is in accordance with the literature [196]. The corresponding band has a frequency 1277 cm^{-1}. The band at 1296 cm^{-1} represents small portion of α_R population, while the bands at 1307 is characteristic of P_{II} conformation [45]. There is an additional low-frequency band in amide III region at 1245 cm^{-1} that is not related to any conformation, rather it represents the aggregated form of peptide.

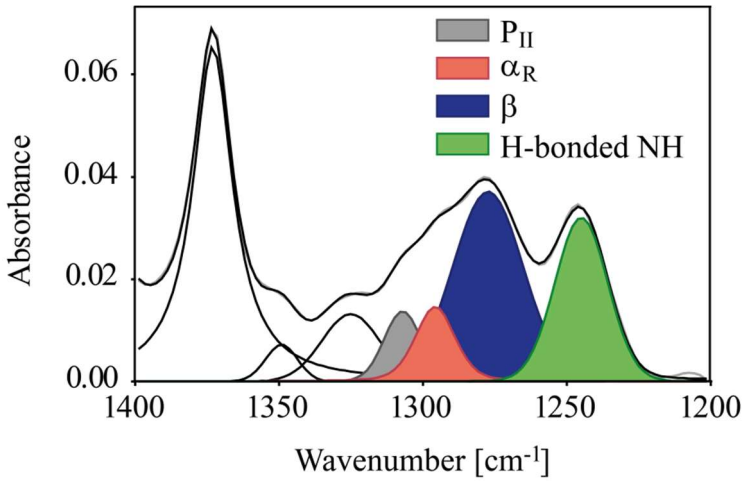


Figure 13: Amide III region of IR spectrum of Ala in DMSO at room temperature.

Due to the low dielectric constant of DMSO the amide I band of alanine peptides is blue-shifted compared to the amide I position in water. In the IR spectrum in Figure 14 the amide I frequency has the maximum at 1665 cm^{-1} for all studied peptides. However, amide II band is downshifted and broader indicating that the solvent molecules are coordinated around the NH amide groups and thus forming the intermolecular H-bonds. The NH groups involved in aggregation of short alanine peptides have the low-frequency band at 1503 cm^{-1} in amide II region that is more pronounced in longer alanine peptides. The other band that gains in intensity with the longer peptide chain is the band at 1245 cm^{-1} in amide III region that also downshifts with the length of peptide. In the amide III region of short alanine peptides IR spectrum shows similar band composition with wavenumbers 1307 (P_{II} conformation), 1296 (α_R conformation), 1277 (β conformation) and 1245–1236 cm^{-1} (aggregated peptides).

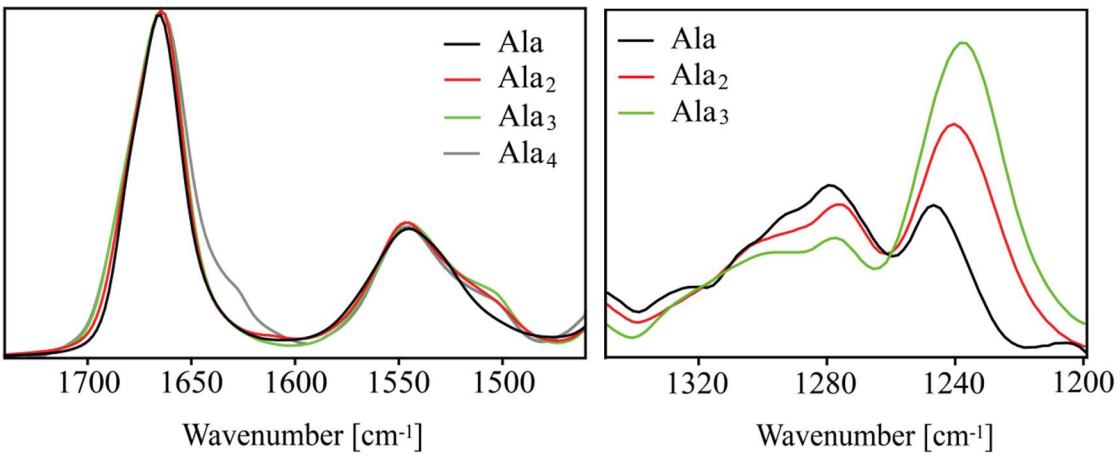


Figure 14: a) Amide I and II region; b) amide III region of IR spectra of Ala_n (n= 1, 2, 3, 4) in DMSO-d_6 at room temperature.

4.1.1.4 Raman Spectroscopy of Short Alanine Peptides in TFE

The Raman spectrum of Ala_2 in TFE shows the upshift of amide I frequency that corresponds to the intensity increase of the band at 1677 cm^{-1} due to the weakly shielded carbonyls. The band at 1658 cm^{-1} represents the solvent exposed amide carbonyls. The third component at 1638 cm^{-1} is assigned to a vibration of carbonyls involved in inter- or intramolecular H-bonds (Figure 15, a). In amide III region, Raman spectrum shows strong vibrational band at 1262 cm^{-1} which could be assigned to C_7 conformation. The band at 1304 cm^{-1} is assigned to P_{II} conformation while the band at 1237 cm^{-1} represents small population of aggregated peptides.

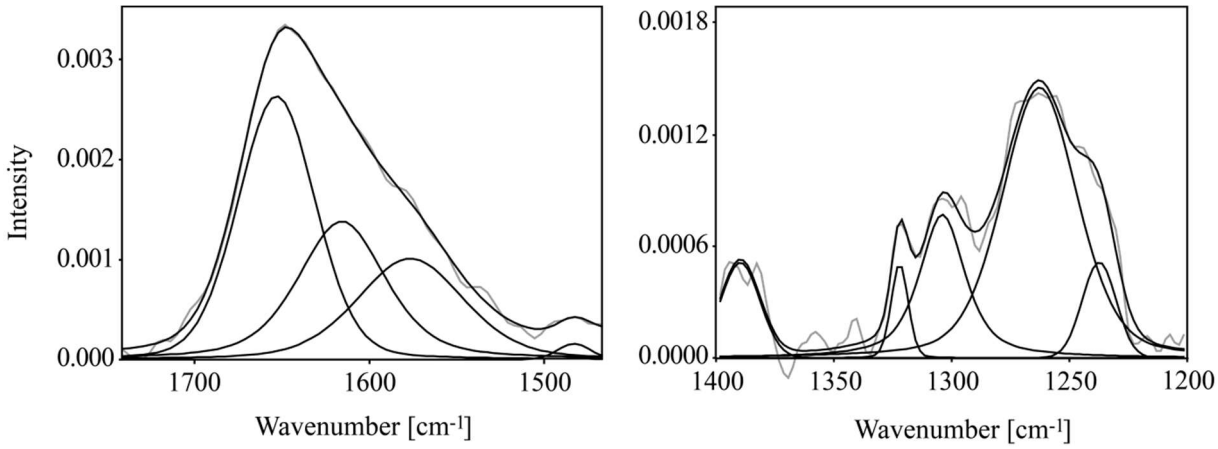


Figure 15: a) Amide I and b) amide III region of Raman spectrum of Ala_2 in TFE-d_2 at room temperature. Spectrum in amide III region was obtained after the subtraction with the spectrum of Ala_2 in TFE-d_3 and represents NH vibrational modes without CH_3 and $C_\alpha H$ deformation modes.

Raman spectra of Ala, Ala_2 and Ala_3 in TFE-d_2 show similar band shapes in amide I and amide III region. The solubility of Ala_4 in TFE-d_2 is too low to measure Raman spectrum. The corresponding IR spectrum shows the amide I and amide II bands that are similar to the one of the shorter alanine peptides.

4.1.2 VCD of Short Alanine Peptides

4.1.2.1 VCD of Short Alanine Peptides in Water

The VCD spectra of Ala were measured in H_2O and D_2O (Figure 16, a) to obtain the amide II and amide I region, respectively. From Figure 16 it is evident that both vibrational bands are negative and have the minimum at 1632 cm^{-1} in amide I and at 1543 cm^{-1} in amide II region. IR spectrum of Ala shows the predominant P_{II} conformation obtained from the fitted model

bands in amide III region with corresponding positive VCD band at 1308 cm^{-1} while β conformation has small negative VCD band in amide III region (Figure 16, a).

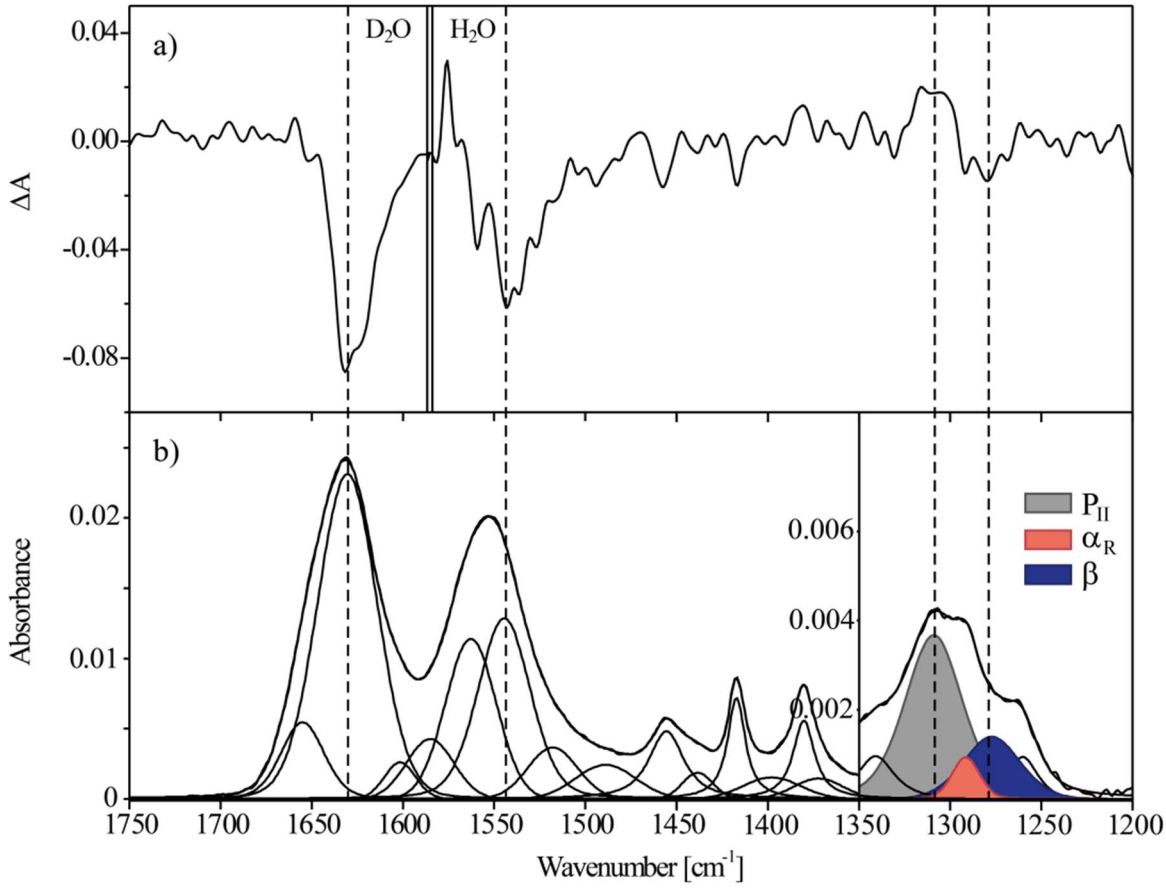


Figure 16: a) The experimental VCD spectrum of Ala in D_2O (left part) and in H_2O (right part) and b) the IR spectrum of Ala in water with the concentration of 0.2 M.

Ala_2 in water has a symmetric negative coupled amide I band with the minimum at 1628 cm^{-1} and the maximum at 1658 cm^{-1} (Figure 17). The low frequency band is indicative of a strong intramolecular H-bond. The similar band shape and band frequencies were reported for the C_7 H-bonded inverse γ-turn of cyclic tetrapeptides [195]. Amide II band of Ala_2 is negative with minimum at 1539 cm^{-1}. Furthermore, Ala_3 in water displays the negative coupled amide I band with small positive band at 1661 cm^{-1} and large negative band at 1637 cm^{-1}. Amide II band is similar to the one found in spectra of Ala and Ala_2.

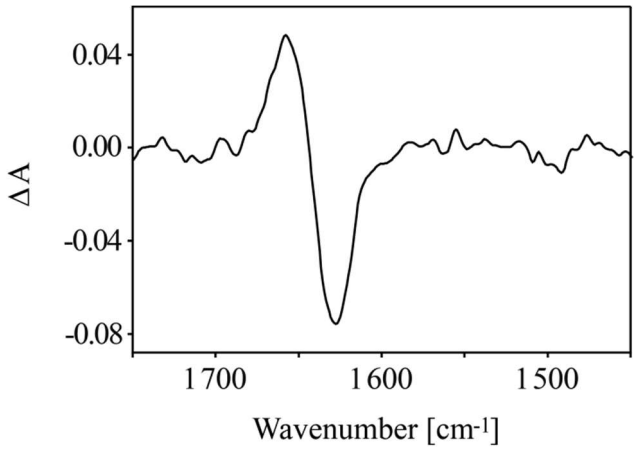


Figure 17: The amide I band of VCD spectrum of Ala_2 in D_2O at a concentration of 0.2 M at room temperature.

Amide III VCD band for particular conformation have not been jet documented in details. In water the short alanine peptides have the VCD bands positioned at those frequencies as found in IR and Raman spectra. From Figure 18 it is evident that the P_{II} and β conformations decrease with the longer peptide chain while the band characteristic of aggregated peptides and the band of C_7 conformation increase that is consistent with the results obtained from IR and Raman spectroscopy.

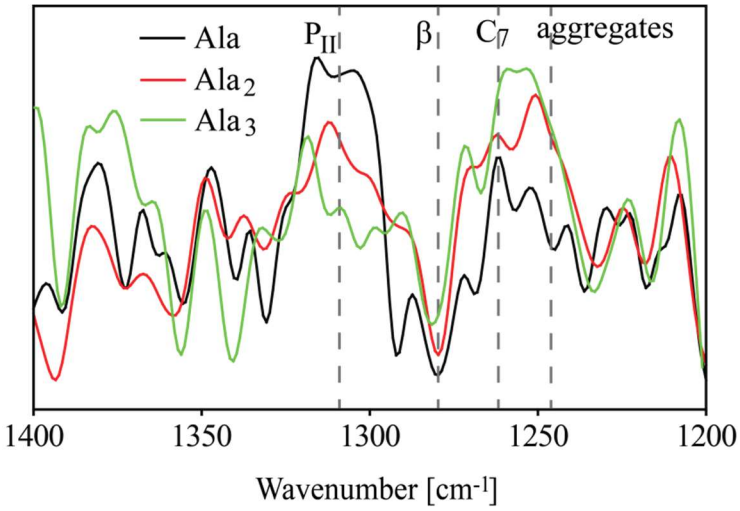


Figure 18: The amide III region of the VCD spectra of Ala, Ala_2 and Ala_3 in water.

4.1.2.2 VCD of Ala in $CDCl_3$

From the VCD spectrum of Ala in $CDCl_3$ (Figure 19) an asymmetric negative couplet amide I band is observed with a large negative signal at the position of the lowest wavenumber at 1647 cm^{-1}, while the positive signal is centred at 1680 cm^{-1}. The amide II VCD band is small and negative at 1509 cm^{-1}. Amide III band is negative at frequencies characteristic of C_7

conformation while positive band stands for intermolecular H-bond of oligomeric Ala. The band for β conformation in amide III region has to low intensity to assign it.

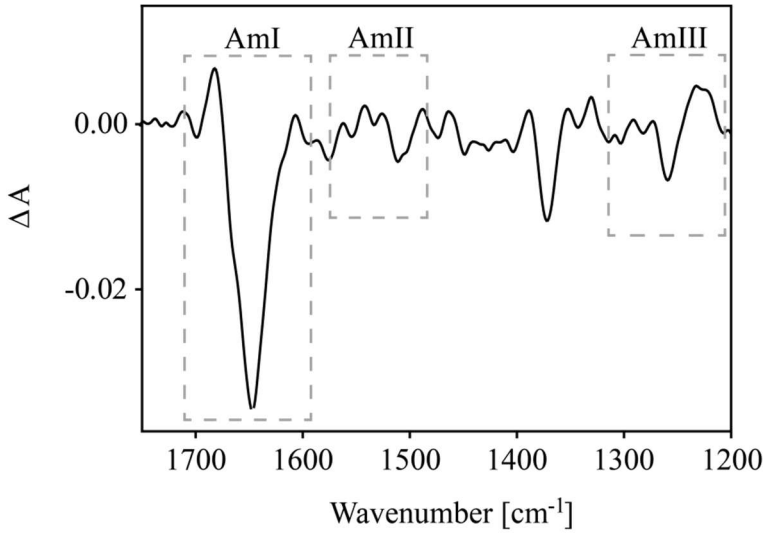


Figure 19: The VCD spectrum of Ala in $CDCl_3$ at 0.2 M concentration.

4.1.2.3 VCD of Short Alanine Peptides in DMSO

The VCD spectrum of Ala in DMSO has an asymmetric negative coupled amide I band with large negative signal at the position 1667 cm^{-1} (Figure 20). The amide II band looks similar although its intensity is three times smaller than that of amide I and has the minimum at 1529 cm^{-1}. The amide III VCD region has positive band at 1295 cm^{-1} (α_R), negative band at 1276 cm^{-1} (β conformation) and again positive band at 1245 cm^{-1} assigned for amino groups of aggregated peptides.

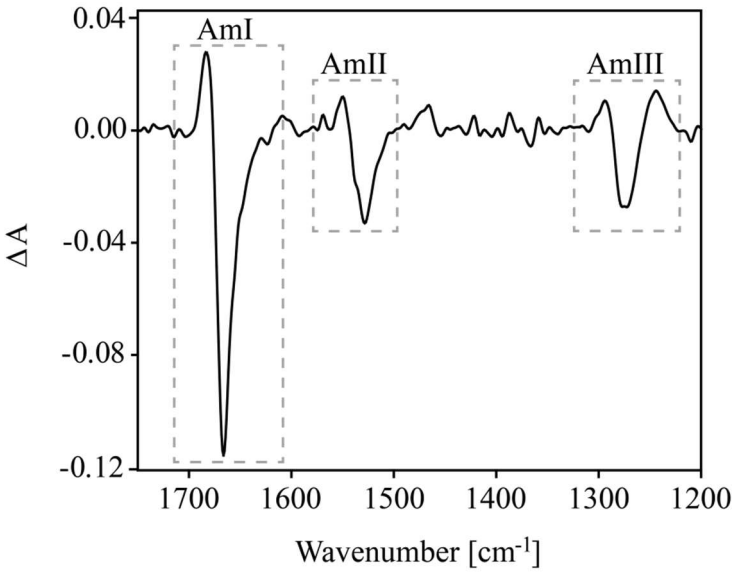


Figure 20: The VCD spectrum of Ala in DMSO-d_6 at 0.2 M concentration.

VCD spectra of Ala_2 and Ala_3 show similar spectral characteristics as VCD spectrum of Ala. The difference between those spectra can be observed only in band at 1245 cm^{-1} that increases with the longer peptide chain. This is consistent with the results found in IR and Raman measurements.

4.1.2.4 VCD of Short Alanine Peptides in TFE

Both, the amide I and the amide II VCD band of Ala in TFE-d_2 are broad and negative with the frequencies at 1644 cm^{-1} and 1537 cm^{-1}, respectively. Ala_2 and Ala_3 both have similar VCD spectrum as Ala although the amide I is even broader while amide II shifts to lower wavenumbers as peptide chain increases. Amide III region of VCD spectra of short alanine peptides in TFE-d_2 is overlapped with the solvent bands and thus cannot be assigned.

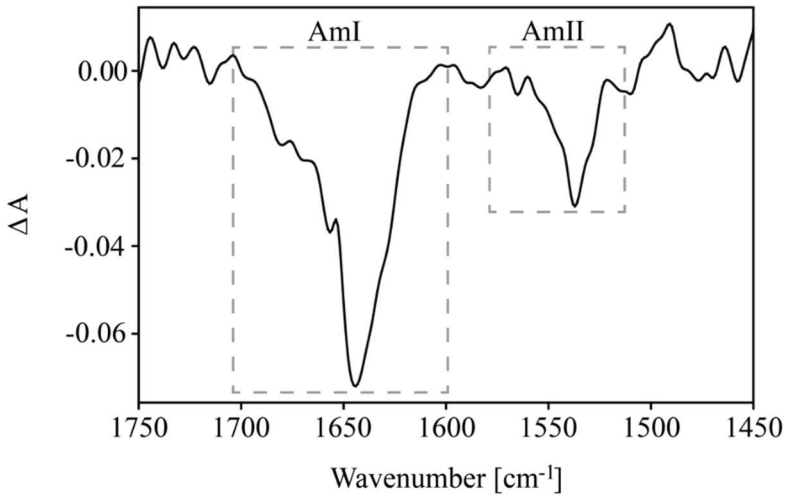


Figure 21: The VCD spectrum of Ala in TFE-d_2 at 0.3 M concentration.

4.1.3 *Ab Initio* Calculations of Ala

4.1.3.1 *Ab Initio* Calculations of Ala in Gas Phase

Structural optimization of individual conformation in the gas phase was evaluated for Ala using *ab initio* methods DFT and HF with the 6-31G(d,p) basic set, where frequencies, dipole and rotational strengths for each conformational minimum on the potential energy surface were calculated. The initial values of dihedral angles (ϕ,ψ) used in calculations for different conformations were (-75°, 145°) for P_{II}, (-65°, -45°) for α_R, (-120°, 120°) for β and (-78°, 65°) for C_7 conformation. These values were obtained from the maxima of the coil library distribution of alanine as reported by Avbelj and Baldwin [56]. Free energies and dihedral angles of individual conformation obtained from DFT and HF were derived and compared in

Table 6. Based on the results in gas phase we concluded that both the HF and DFT yield similar descriptions of dihedral angles while thermodynamics of conformations were significantly different. The DFT calculations gave optimized conformations with favourable free energy compared to HF. Furthermore, Boltzmann distribution of conformations (w_A) obtained by DFT can be directly correlated to those obtained from the IR spectroscopy of Ala in argon matrix (Unpublished data) where C_7 is the prevailing conformation. Therefore, only DFT calculations were used in further work. However, P_{II} and α_R conformations were not stable in gas phase and resulted in their relaxation to the structure that were reminiscent of C_7 and conformation from the leftmost part of Ramachandran plot, respectively. Therefore, P_{II} and α_R are not included in Table 6 as conformations that are populated in gas phase.

Table 6: The optimized conformations of Ala in gas phase calculated by B3LYP/6-31G(d,p) and HF/6-31G(d,p) method without P_{II} and α_R conformations that are not persistent in the gas phase potential surface. Each conformation is represented by calculated dihedral angles, the differences in total ground state energy and Boltzmann weight (w_A).

Conformation in gas	ϕ [°]		ψ [°]		ΔE [kcal/mol]		w_A	
	B3LYP	*HF*	*B3LYP*	*HF*	*B3LYP*	*HF*	*B3LYP*	*HF*
β	-158.6	-157.3	164.7	159.3	0	0	0.081	0.347
C_7	-83.2	-86.1	72.6	78.8	-1.45	-0.38	0.919	0.653

From Table 6 it is evident that in gas phase energetically the most favourable conformation of Ala is C_7 that is stabilized with the formation of the intramolecular H-bond between NH and C=O terminal groups while the dihedral angles yield the value (ϕ,ψ) = (-83.2°,72.6°) (Figure 22). Beside C_7 conformation there is small population of β conformation (8 %) with 1.45 kcal/mol less favourable energy than C_7 conformation.

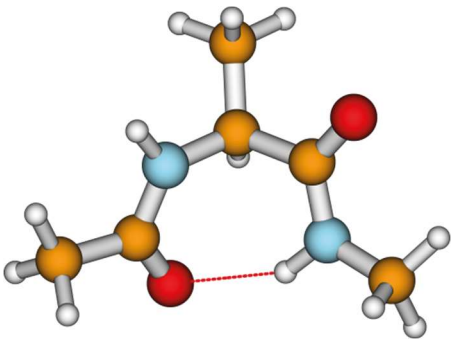


Figure 22: The schematic presentation of Ala in C_7 conformation after optimization at the B3LYP/6-31G(d,p) level. Dashed line shows the H-bond.

4.1.3.2 DFT Calculations of Ala in Water

When studying the alanine peptide in the liquids the solvent effect must be incorporated. DFT calculations of Ala were first performed in water modelled as dielectric continuum. The Boltzmann populations of β conformation increased from 8 % in gas phase to 24 % in water while C_7 conformation still populated 76 % of total conformation in water (Table 7). In contrary to the gas phase, P_{II} conformation was stable in water with calculated dihedral angles $(\phi,\psi) = (-57.6°,140.3°)$ whereas α_R conformation is still unstable. Additionally, like in gas phase, stabilization of C_7 conformation arises with formation of intramolecular H-bond. However, these results contrast the experimental predictions of Ala in water from Grdadolnik et al.[45].

Table 7: The optimized conformations of Ala in water calculated by B3LYP/6-31G(d,p) method. Each conformation is represented by calculated dihedral angles, the differences in total ground state energy and Boltzmann weight (w_A).

Conformation in water	ϕ [°]	ψ [°]	ΔE [kcal/mol]	w_A
P_{II}	-69.7	143.6	0	0.016
β	-156.8	159.8	-1.60	0.236
C_7	-84.4	70.1	-2.30	0.764

Explicit water molecules were appended to the peptide systems to inspect weather the formation of hydrogen bonding between water and peptide backbone C=O and NH groups could account for the missing accuracy of the previous model and lead to potentially better agreement in the prediction of conformational distribution of Ala in water. The simulations were performed for Ala in P_{II}, β, α_R and C_7 conformation, respectively, in the presence of four explicit water molecules occupying different positions and orientations forming characteristic hydrogen bonding patterns. The initial structures were sampled by MD simulations which were further geometry optimized and used for IR and VCD calculations by the DFT approach. The average Boltzmann weight of each conformation was determined by grouping the structures according to their occupied conformation and taking into account their ground state energies (Table 8).

Table 8: The optimized conformations of Ala in water with explicit four molecules of water calculated by B3LYP/6-31G(d,p) method. Each conformation is represented by calculated dihedral angles, the differences in total ground state energy and Boltzmann weight (w_A).

Conformation + 4 H_2O	ϕ [°]	ψ [°]	ΔE [kcal/mol]	w_A
P_{II}	-73.1	125.9	-1.49	0.750
α_R	-97.0	-55.0	-1.21	0.069
β	-127.3	135.7	-0.47	0.157
C_7	-84.3	76.4	0	0.024

From Table 8 it is evident that hydrogen bonds formed by water molecules have a significant impact on overall conformational populations. P_{II} conformation turns out to be dominant one with the population of 75 %, whereas β conformation populates 16 % of Ala. The addition of four water molecules indeed leads to the change of conformation populations as it stabilizes P_{II} conformation which was predicted to be non-populated when water was modelled only as dielectric continuum. The most favourable structure of Ala in P_{II} conformation with 4 water molecules, shown in Figure 23, has the C=O and NH backbone groups connected with water bridges that stabilize this conformation.

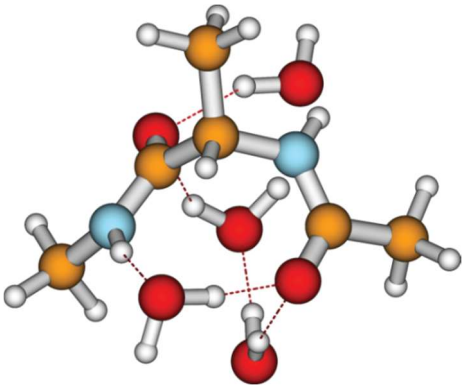


Figure 23: The schematic presentation of most favourable structure of Ala in P_{II} conformation surrounded by 4 molecules of water after optimization at the B3LYP/6-31G(d,p) level. Dashed lines show the H-bonds.

Rotational strengths of vibrational modes in Ala for each representative structure in the ensample were calculated by B3LYP/6-31G(d,p) method. Overall spectrum was obtained by the averaging of the spectra of the individual structures taking into account the proper Boltzmann weight of each contribution giving rise to 75 % of P_{II}, 16 % of β, 7 % of α_R and 2

% of C_7 conformation (Figure 24, a: grey curve). The resulting spectrum has positive coupled amide I and amide II VCD band.

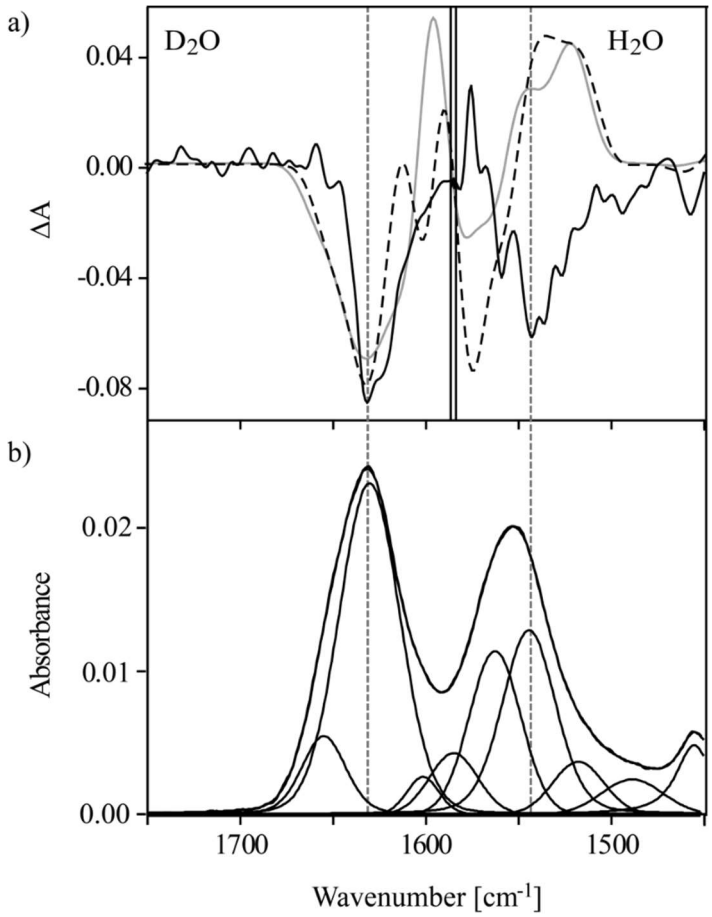


Figure 24: a) Amide I and II regions of the VCD spectra of Ala: the experimental spectrum in H_2O and D_2O (black curve); the calculated spectrum with 91 % of P_{II} and 9 % of β (grey curve); the calculated spectrum with 60 % of P_{II}, 29 % of β and 11 % α_R (dotted curve). b) Amide I and II regions of experimental IR spectrum of Ala in water.

To show the computational spectrum with experimentally obtained populations of individual conformation we averaged calculated spectra in the following proportion: 60 % of P_{II}, 29 % of β and 11 % of α_R conformation as reported by Grdadolnik et al. [45]. The resulting spectrum (Figure 24, a: dotted curve) shows a negative amide I band while the amide II band is positively coupled. Such calculated spectrum has better agreement with experimental one as observed from Figure 24.

Both, the experimental IR and VCD spectrum show the predominant P_{II} conformation obtained from the fitted model bands in amide III region. The calibrated VCD spectra show the positive signal in amide III region that is in accordance with experimental spectrum (Figure 25).

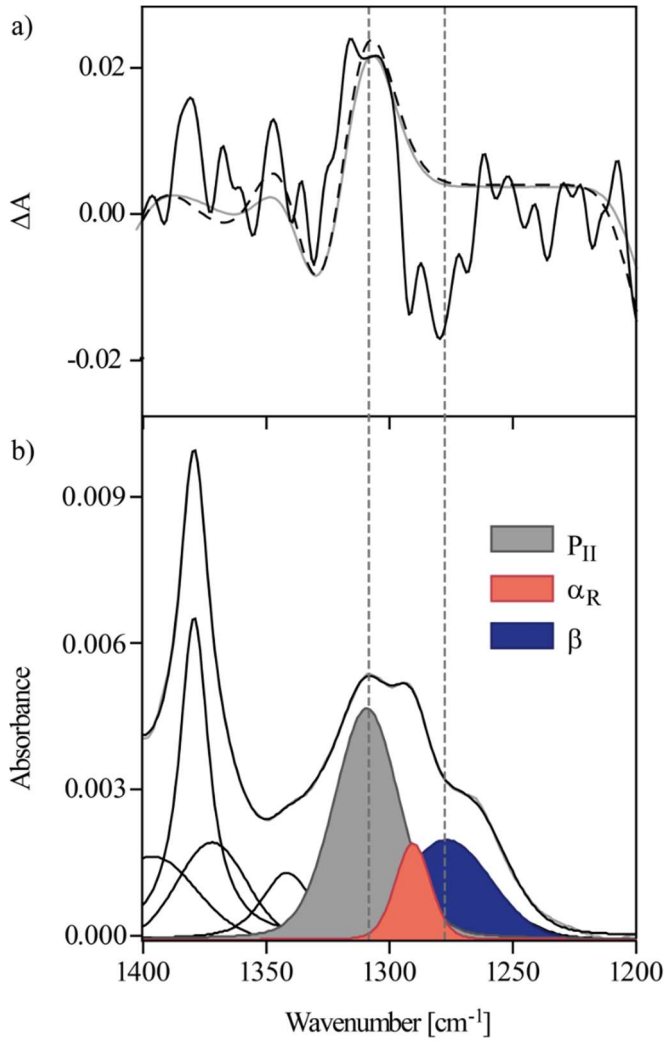


Figure 25: Amide III region of a) the experimental VCD spectrum (black curve), the calculated spectrum with 75 % of P_{II}, 16 % of β, 7 % of α_R and 2 % of C_7 (grey curve) and the calculated spectrum with 60 % of P_{II}, 29 % of β and 11 % α_R (dotted curve); b) the IR spectrum of Ala in water with the concentration of 0.2 M.

4.1.3.3 DFT Calculations of Ala in DMSO

When simulating the Ala with the bulk solvent DMSO using PCM [185] with the dielectric constant ε = 46.826, the VCD spectrum shows positively coupled amide I band, strong positive amide II band and strong negatively coupled amide III band that is not in agreement with experimental results. Therefore, two explicit DMSO molecules were included in simulation that after DFT optimization bind to peptide NH backbone groups. The corresponding VCD spectrum has similar band shape of amide I and amide III as experimental one whereas amide II band shape shows small positively coupled band that is inconsistent with experimentally observed spectrum (Figure 26, a).

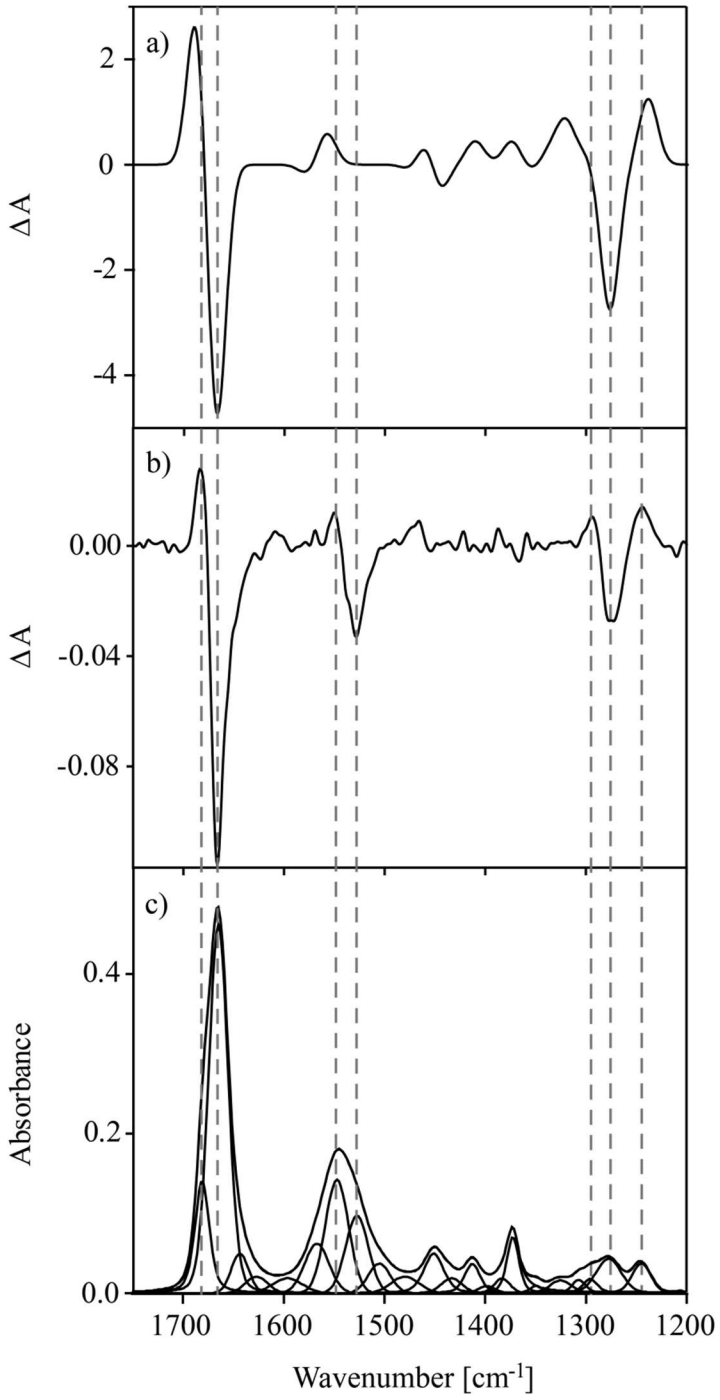


Figure 26: a) The calculated VCD (B3LYP(d,p)); b) the experimental VCD and c) the experimental IR spectra of Ala in DMSO-d_6.

From the DFT calculations we can observe that two DMSO molecules bind to NH peptide backbone groups as shown in Figure 27 where the optimized dihedral angles (ϕ, ψ) = (-129.8°, 133.3°) of Ala reveals characteristics of β conformation that is in agreement with literature [196].

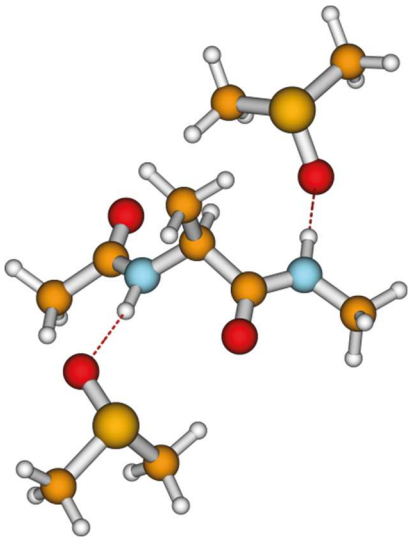


Figure 27: Schematic presentation of Ala in β conformation surrounded by 2 molecules of DMSO after optimization at the B3LYP/6-31G(d,p) level. Dashed lines show the H-bonds between S=O group of DMSO and NH group of peptide.

4.1.3.4 DFT Calculations of Ala in Chloroform

DFT calculations of Ala in $CHCl_3$ showed that only C_7 and β conformation are stable in chloroform. The results obtained from IR spectroscopy indicate that Ala in chloroform at 100 mM forms a β oligomer structure that is in equilibrium with monomer structure with 0.7 of C_7 population and 0.3 of β population. The DFT calculation of VCD spectrum of Ala in chloroform was performed using PCM [185] with dielectric constant, ε = 4.711. The calculated spectra were averaged in 0.7:0.3 ratio for C_7:β conformation, resulting in negative coupled amine I band and positive amine II and positive coupled amide III band (Figure 28, a). When Ala was simulated only as C_7 monomer the amide I and II had a similar shape while negative amide III band was obtained indicating that calculated spectrum of averaged C_7 and β monomers brings higher degree of similarity with experimental results. However, it would be necessary to take into account the effect of oligomerization of Ala in chloroform in ordered to compare calculated VCD spectrum with the experimental one.

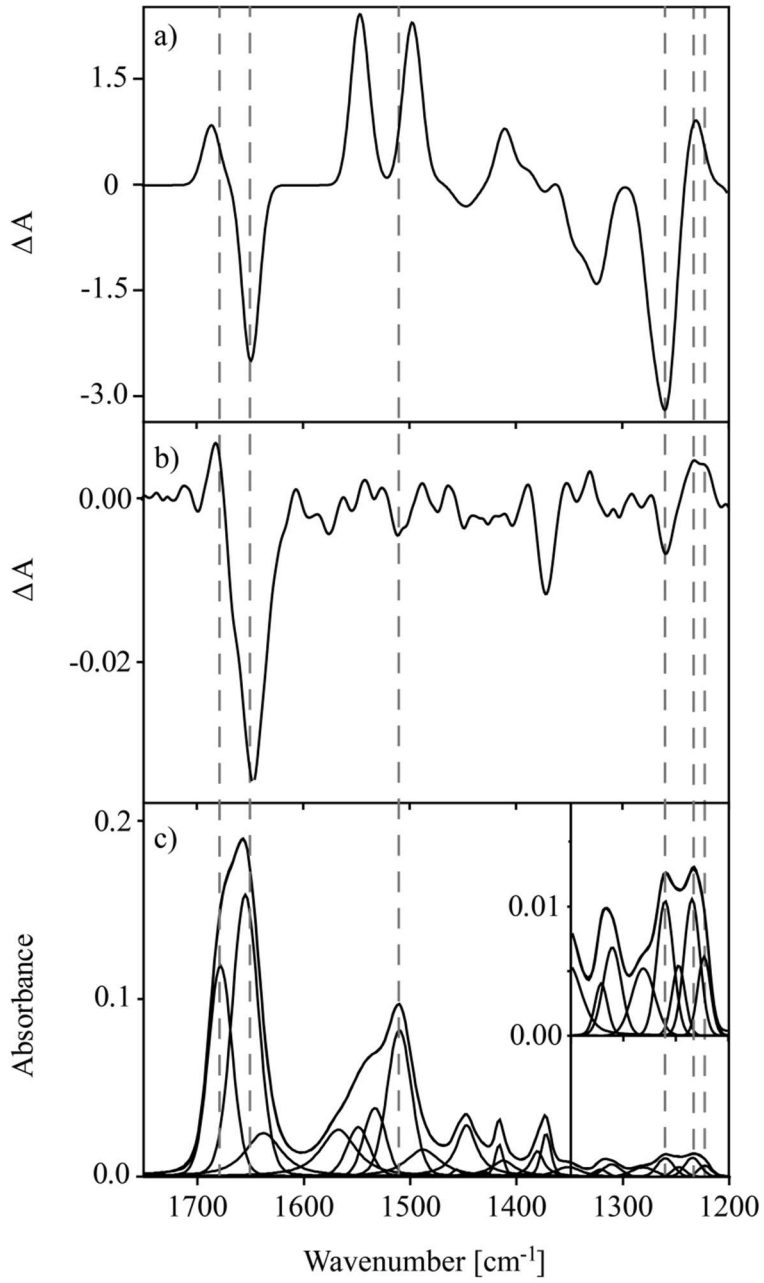


Figure 28: a) The calculated VCD (B3LYP(d,p)); b) the experimental VCD; and c) the experimental IR spectra of Ala in $CDCl_3$. Peptide concentration in experimental measurement was 100 mM.

The oligomer structures were not included in calculations of VCD spectrum because the number of Ala peptides that comprise the oligomer structure is not known. However, we modulated the parallel dimer of Ala in β conformation, as shown in Figure 29, and simulated the structural optimization. After DFT calculations the β dimer formed stable association of two Ala peptides that are H-bonded with each other.

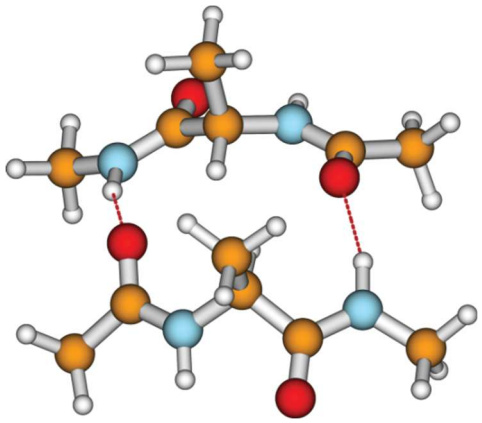


Figure 29: Schematic presentation of parallel dimer of Ala in β conformation in chloroform after optimization at the B3LYP/6-31G(d,p) level. Dashed lines show the H-bonds.

4.1.4 UV-CD Spectroscopy of Short Alanine Peptides

4.1.4.1 UV-CD of Short Alanine Peptides in Water

The UV-CD spectrum of Ala in water is invariant over the concentration range from 0.01 mM to 200 mM; therefore, it is assumed that intermolecular interactions have no influence on the measurements. Additionally, the upper concentration limit for the UV-CD studies approaches that used in the NMR, IR, RAMAN and VCD experiments, enabling the direct correlations between those experiments.

The UV-CD spectrum of Ala is characteristic for P_{II} conformation with the strong negative $\pi \rightarrow \pi^*$ band that is similar to the one found in the literature [40]. However, the lack of positive band at around 215 nm suggests that P_{II} is not the only conformation in Ala (Figure 30). With the longer peptide chain the $\pi \rightarrow \pi^*$ band becomes less pronounced indicating the decreasing of the P_{II} conformation while small negative band at 228 nm gains in the intensity (inset in Figure 30) that is characteristic of C_7 conformation and β-turns [106]. Compared to the Ala at 0.05 mM concentration the Ala_2, Ala_3 and Ala_4 all have weak band at around 218 nm. However, much stronger band at 218 nm is expected for pure P_{II} conformation [63].

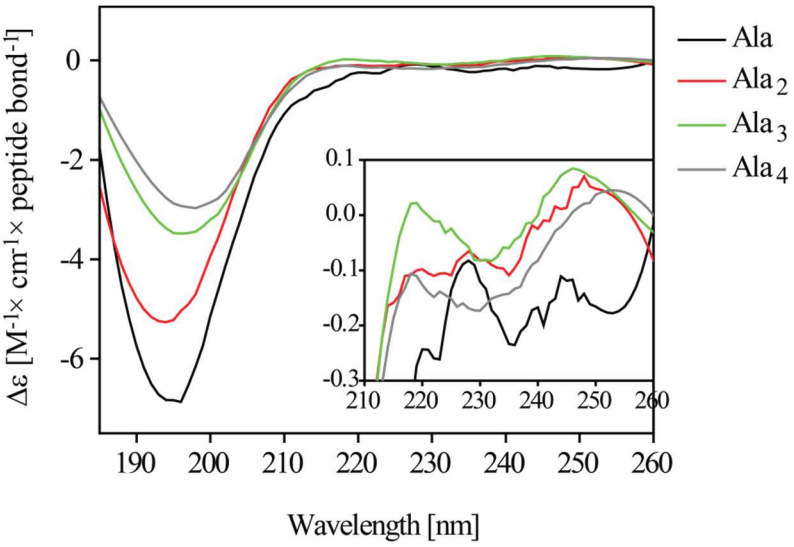


Figure 30: UV-CD spectra of Ala, Ala_2, Ala_3 and Ala_4 in water at 0.05 mM concentration and 30 °C.

The temperature measurements of short alanine peptides showed a decrease of P_{II} conformation with increasing temperature (Figure 31). The maximum dichroism at approximately 215 nm decreases with increasing temperature, reflecting the change of the conformational distribution from P_{II}-like to more extended β conformation (difference spectrum in Figure 31) that is in agreement with the literature [46].

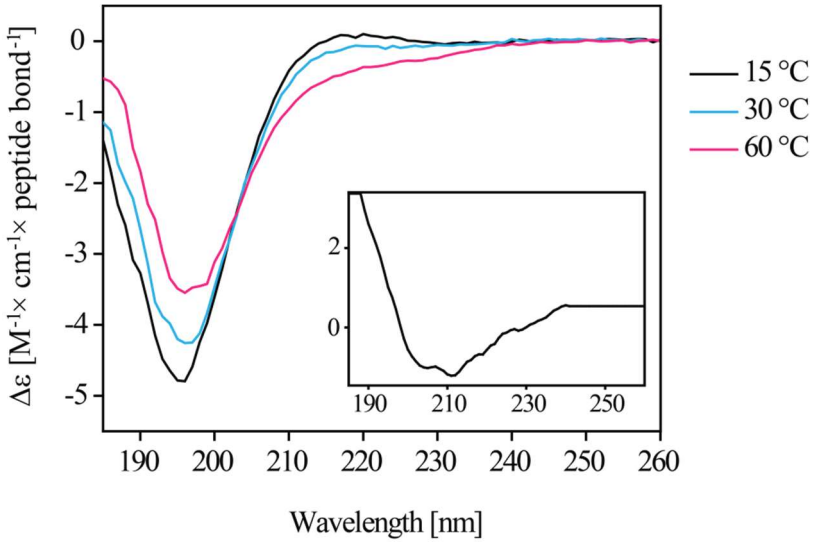


Figure 31: The UV-CD spectra of Ala_2 in water at concentration 1.4 mM at 15 °C, 30 °C and 60 °C. Inset: The UV-CD difference spectrum obtained by subtracting the spectrum recorded at 15 °C from the spectrum recorded at 60 °C.

4.1.4.2 UV-CD of Short Alanine Peptides in TFE

When Ala_2 is dissolved in TFE the UV-CD spectrum shows a negative n→π* band near or somewhat below 230 nm and a positive band around 200 nm that are characteristic UV-CD band positions of α-helices [197]. Generally, the π→π* band intensity of α-helix is three times higher than n→π* band [197]. However, from Figure 32 it is evident that π→π* and n→π* bands of alanine peptides in TFE are of comparable intensities. It is difficult to reconcile the quantitative UV-CD intensities with the two or more component mixture. The algorithms to estimate the secondary structure composition of proteins are widely used but they exclude the databases for oligopeptides. Therefore, the UV-CD spectra of peptides cannot be analysed reliably at present except under condition where one secondary structure (or conformation) predominates. Additionally to α-helix, the C_7 conformation can be appropriate candidate to accomplish the spectral characteristics of alanine peptides in TFE while this conformation has the negative n→π* CD and positive π→π* CD, both with similar intensities [106]. Therefore, the UV-CD spectrum of the C_7 conformation is reminiscent of the spectra in Figure 32. Another conformation with the characteristic positive π→π* band near 190 nm is the type II ß-turn conformation. However, UV-CD cannot distinguish between type II ß-turn and C_7 conformation. From Figure 32 it is evident that the spectra of short alanine peptides approximate the characteristic spectrum of C_7 conformation with the longer peptide chain.

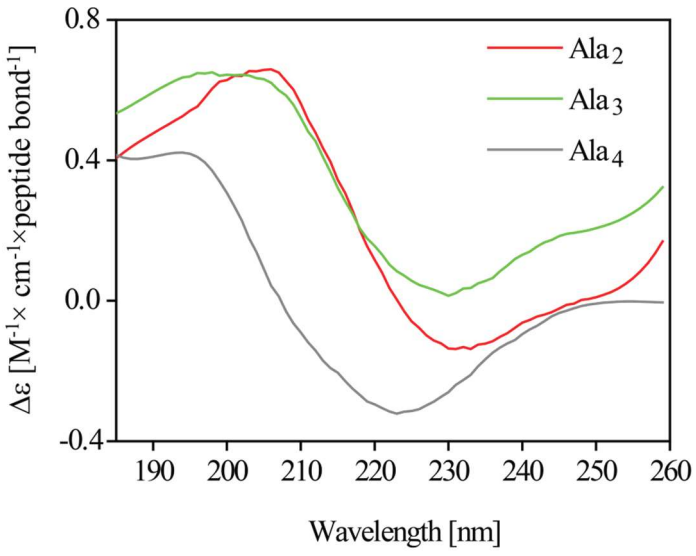


Figure 32: The UV-CD spectra of Ala_n (n = 2, 3, 4) in TFE at concentration 0.05 mM at 30 °C.

4.1.4.3 UV-CD of Short Alanine Peptides in DMSO

When Ala_2 is dissolved in DMSO the UV-CD spectrum shows the characteristics patterns of β structures with the broad negative band at 210–230 nm and stronger positive band at 190–200 nm. The features of β conformation become more pronounced in UV-CD spectrum of Ala_3 and Ala_4 (Figure 33).

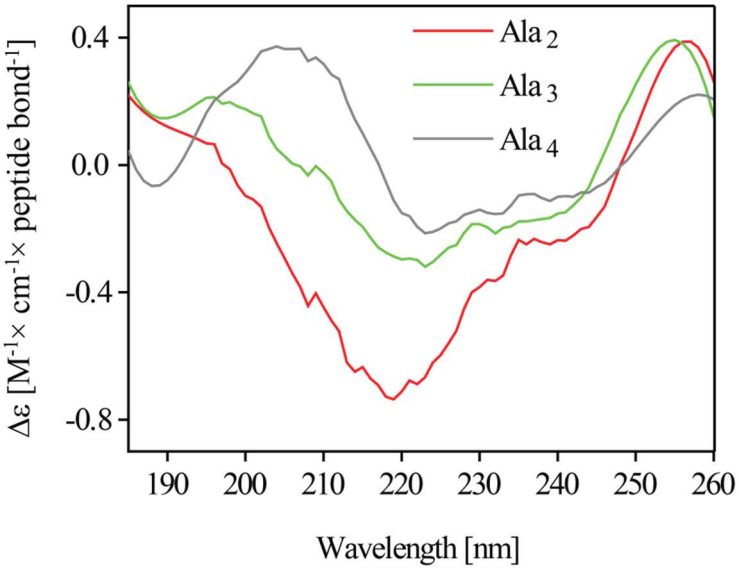


Figure 33: The UV-CD spectra of Ala_n (n = 2, 3, 4) in DMSO at concentration 0.1 mM at 30 °C.

4.1.5 NMR Spectroscopy of Short Alanine Peptides

4.1.5.1 NMR Spectroscopy of Short Alanine Peptides in Water

1D NMR spectra were measured for short alanine peptides in water at pH 3. From the spectra the coupling constants for short alanine peptides were obtained that are represented in Table 9. The average values of coupling constants do not differ significantly between alanine peptides, however, they slightly decrease with the longer peptide chain. The coupling constants were also measured at pH 5 and at concentration used in vibrational spectroscopy experiments but short alanine peptides didn't show any pH or concentrational dependence.

Table 9: The $^3J(H_\alpha,H_N)$ coupling constants of Ala_n (n = 2, 3, 4) at 20 °C in water at 20 mM concentration for Ala_2 and Ala_3 and 2 mM for Ala_4.

Ala_n n =	$^3J(H_N(1))$ [Hz]	$^3J(H_N(2))$ [Hz]	$^3J(H_N(3))$ [Hz]	$^3J(H_N(4))$ [Hz]
2	5.98	6.50	/	/
3	5.87	6.16	6.39	/
4	6.09	5.78	5.37	5.77

The temperature measurements of short alanine peptides in water showed the increasing values of $^3J(H_\alpha,H_N)$ coupling constant with the increasing temperature, indicating the increase of β conformation. Figure 34 represents both $^3J(H_\alpha,H_N)$ coupling constants for 20 mM Ala_2 in water at different temperatures.

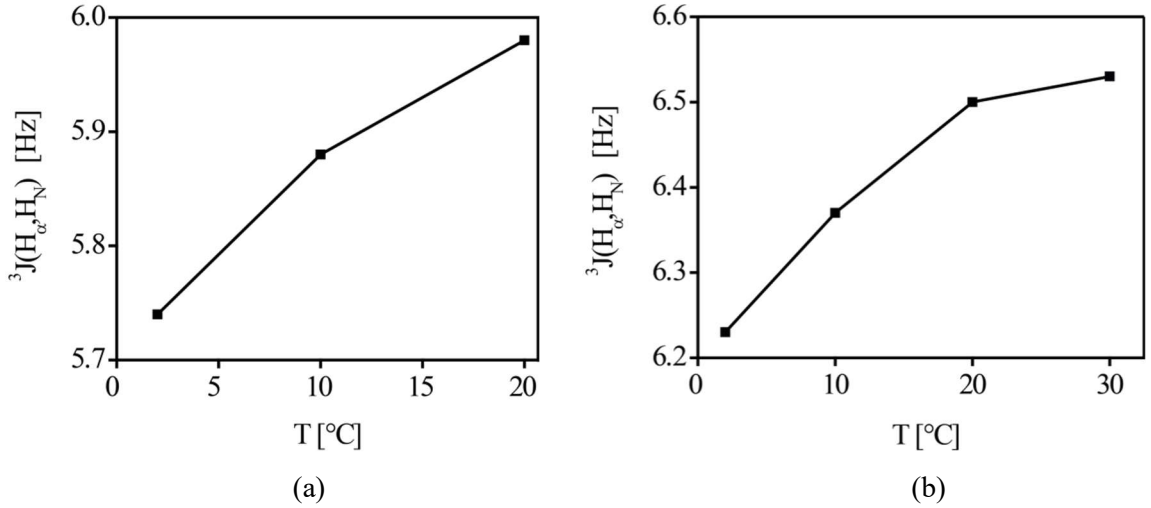


Figure 34: The $^3J(H_\alpha,H_N)$ coupling constants for amino proton a) $H_N(1)$ and b) $H_N(2)$ for 20 mM Ala_2 at pH = 3 in water (10 % D_2O) measured in temperature range from 2 to 40 °C. Above 20 °C the doublet of $H_N(1)$ proton had too low resolution to obtain the coupling constant.

Using Equation 6 the differences in β proportion of each amino proton of short alanine peptides can be estimated with the temperature increase (Table 10). The maximum melting of β conformation was observed for Ala_4, followed by Ala_2 and Ala_3.

Table 10: The differences in β proportion of each amino proton of Ala_n (n = 2, 3, 4) after temperature increase. For Ala_2 the $H_N(1)$ value was obtained from 2 to 20 °C and from 2 to 30 °C for the $H_N(2)$. For Ala_3 and Ala_4 all H_N protons values ware obtained from 2 to 30 °C.

Ala_n n =	$\Delta \chi_\beta$			
	$H_N(1)$	$H_N(2)$	$H_N(3)$	$H_N(4)$
2	7 %	9 %	/	/
3	4 %	8 %	9 %	/
4	10 %	8 %	19 %	7 %

The temperature coefficient method provides us information about H-bonding of functional groups in peptides where the slopes of the chemical shift vs. temperature are monitor. With increasing temperature the H-bond length increases and consequently the chemical shift of the amide proton changes. If an amide proton exchanges slowly and has a temperature coefficient more positive than -4.5 ppb/°C, it is H-bonded, while if it exchanges rapidly and has a temperature coefficient more negative than -4.5 ppb/°C, it is not H-bonded [198]. The temperature-dependent proton spectra were recorded over a wide temperature range in water. The amide proton temperature coefficients of Ala_n (n = 2, 3, 4) are summarized in Table 11. All the temperature coefficients are lower than -4.5 ppb/°C suggesting that NMR spectroscopy can't observe any intramolecular H-bonding of the amino protons of the short alanine peptides in the water.

Table 11: The temperature coefficient for Ala_n (n = 2, 3, 4) from 2 °C to 60 °C in water at concentration 20 mM for Ala_2 and Ala_3 and 2 mM for Ala_4.

Ala_n n =	$\Delta\delta/\Delta T(H_N(1))$ [ppb/°C]	$\Delta\delta/\Delta T(H_N(2))$ [ppb/°C]	$\Delta\delta/\Delta T(H_N(3))$ [ppb/°C]	$\Delta\delta/\Delta T(H_N(4))$ [ppb/°C]	$\Delta\delta/\Delta T$ ($H_N(5)$) [ppb/°C]
2	-7.6	-8.3	-6.2	/	/
3	-7.2	-7.6	-7.8	-6.2	/
4	-7.2	-7.5	-7.5	-7.6	-6.0

A total of 12 NOESY peaks were identified in ^1H-^1H NOESY spectrum of Ala_2 in water (Table 12). From Equation 7 we calculated the distances from measured NOE volumes in spectrum of 200 mM Ala_2 in water using fixed reference, $r_{ref}(H_N(1) \leftrightarrow CH_3(0)) = 2.92$ Å.

Table 12: The NOE distances used for the structure calculation of Ala_2 in 90 % H_2O and 10 % D_2O. The labelled rows represent those distances that are conformational sensitive.

Assignment	Normalized NOE volume	Fixed distance [Å]	Calculated distance [Å]
$H_N(1) \leftrightarrow CH_3(0)$	$4.87 \cdot 10^8$	2.93	
$H_N(1) \leftrightarrow CH_3(1)$	$6.67 \cdot 10^8$		2.78
$H_N(2) \leftrightarrow CH_3(0)$	$2.34 \cdot 10^7$		4.86
$H_N(2) \leftrightarrow CH_3(2)$	$8.70 \cdot 10^8$		2.66
$H_N(3) \leftrightarrow CH_3(2)$	$2.01 \cdot 10^8$		3.39
$H_N(3) \leftrightarrow CH_3(3)$	$4.50 \cdot 10^8$		2.97
$H_N(1) \leftrightarrow C_\alpha H(1)$	$6.16 \cdot 10^8$		2.82
$H_N(2) \leftrightarrow C_\alpha H(1)$	$2.18 \cdot 10^9$		2.28
$H_N(3) \leftrightarrow C_\alpha H(2)$	$1.13 \cdot 10^9$		2.55
$H_N(2) \leftrightarrow H_N(3)$	$3.23 \cdot 10^8$		3.14
$C_\alpha H(1) \leftrightarrow CH_3(1)$	$1.44 \cdot 10^9$		2.45
$C_\alpha H(2) \leftrightarrow CH_3(2)$	$8.39 \cdot 10^8$		2.75

Based on the average characteristic distances of each conformation, represented in Table 4, the measured NOE distances of Ala_2 that are conformational sensitive (grey rows in Table 12) are short. High values of $H_N(2) \leftrightarrow C_\alpha H(1)$ and $H_N(2) \leftrightarrow C_\alpha H(1)$ NOEs are characteristic of C_7, β and P_{II} conformation while detectable NOE between $H_N(2)$ and $H_N(3)$ protons is characteristic of α_R and C_7 conformation. However, it is difficult to interpret these results in the view of conformational distribution of Ala_2 in water due to the conformational averaging.

The analysis of NOESY spectra of Ala_2, Ala_3 and Ala_4 shows that for all measured alanine peptides NOESY cross-peaks are observed in the H_N-H_α, H_N-H_{Me} and H_α-H_{Me} regions, where the first one is the most relevant for conformational analysis. In fact, in all NOESY spectra the H_N-H_α intra-residue cross-peaks are less intense than the inter-residue ones, that is characteristic for the P_{II} and/or C_7 and/or β conformation [76]. For Ala_2 and Ala_3 the weak NOE's in the amide-amide region are observed, which suggest the presence of α_R and/or C_7 conformer.

4.1.5.2 NMR Spectroscopy of Short Alanine Peptides in DMSO-d_6

1D proton spectra were measured for Ala, Ala_2, Ala_3 and Ala_4 dissolved in DMSO-d_6 that is a solvent with a proton acceptor properties. As evident from Table 13, the solvent DMSO causes the shift of $^3J(H_\alpha,H_N)$ coupling constants toward β conformation for all studied alanine peptides compared to water. There were no concentrational dependences observed from coupling constants and chemical shifts for all studied peptides in DMSO.

Table 13: The $^3J(H_\alpha,H_N)$ coupling constants of Ala_n (n = 1, 2, 3, 4) at 30 °C in DMSO-d_6 at concentration 100 mM for Ala and Ala_2, 33 mM for Ala_3 and 1 mM for Ala_4. The arrows indicate the increasing (↑) or decreasing (↓) of the $^3J(H_\alpha,H_N)$ with the increasing temperature.

Ala_n n =	$^3J(H_N(1))$ [Hz]	$^3J(H_N(2))$ [Hz]	$^3J(H_N(3))$ [Hz]	$^3J(H_N(4))$ [Hz]
1	7.67 ↑	/	/	/
2	7.62 ↓	7.18 ↑	/	/
3	7.26 ↓	7.21 ↓	7.53 ↓	/
4	7.21 ↓	7.23 ↑	7.20 ↑	7.60 ↓

With the increasing temperature the $^3J(H_\alpha,H_N)$ coupling constants of amino proton $H_N(1)$ of the Ala increases from 7.65 Hz at 20 °C to 7.68 Hz at 40 °C. For the other alanine peptides the $^3J(H_\alpha,H_N)$ coupling constants increase or decrease inconsistently for amino protons in one peptide indicating the non-uniform environment for amino protons or the presence of dimer or small aggregated structures of those peptides.

In DMSO, $\Delta\delta/\Delta T$ values above -3 ppb/°C have been taken as indicative of solvent-shielded NH protons that usually participate in the intramolecular H-bonds, while values below -4 ppb/°C have been assigned to the exposed groups [199]. According to this criterion, any of the NH protons of Ala, Ala_2, Ala_3 and Ala_4 are not involved in the intramolecular H-bonds (Table 14) although the central amino proton in Ala_4 has $\Delta\delta/\Delta T < 4$ ppb/°C and can be slightly protected from the solvent.

Table 14: The temperature coefficient for Ala_n (n = 1, 2, 3, 4) from 20 °C to 60 °C in DMSO-d_6 at concentration 100 mM for Ala and Ala_2, 33 mM for Ala_3 and 1 mM for Ala_4.

Ala_n n =	$\Delta\delta/\Delta T(H_N(1))$ [ppb/°C]	$\Delta\delta/\Delta T(H_N(2))$ [ppb/°C]	$\Delta\delta/\Delta T(H_N(3))$ [ppb/°C]	$\Delta\delta/\Delta T(H_N(4))$ [ppb/°C]	$\Delta\delta/\Delta TH_N(4))$ [ppb/°C]
1	-6.3	-5.6	/	/	/
2	-6.2	-5.1	-4.8	/	/
3	-5.3	-5.3	-4.8	-5.1	/
4	-5.3	-5.0	-3.9	-5.3	-5.2

4.1.5.3 NMR Spectroscopy of Short Alanine Peptides in TFE-d_2

Ala, Ala_2, Ala_3 and Ala_4 where measured in TFE-d_2 that is proton donor solvent. As we can observe from Table 15 the TFE causes an increase of the $^3J(H_\alpha,H_N)$ coupling constant toward β conformation for the Ala compared it to the value of Ala in water, while Ala_2, Ala_3 and Ala_4 have lower $^3J(H_\alpha,H_N)$ coupling constants in TFE as in the water. There were no concentrational dependence of Ala, Ala_2 and Ala_3 in TFE-d_2. The $^3J(H_\alpha,H_N)$ coupling constants of all measured peptides increase with the increasing temperature suggesting the peptides populate more β conformation at higher temperature.

Table 15: The $^3J(H_\alpha,H_N)$ coupling constants of Ala_n (n = 1, 2, 3, 4) at -20 °C ([a] at 0 °C for $H_N(1)$ of Ala_2) in TFE-d_2 at 100 mM concentrations for Ala, Ala_2 and Ala_3 and 1 mM concentration for Ala_4.

Ala_n n =	$^3J(H_N(1))$ [Hz]	$^3J(H_N(2))$ [Hz]	$^3J(H_N(3))$ [Hz]	$^3J(H_N(4))$ [Hz]
1	6.61	/	/	/
2	6.77[a]	5.40	/	/
3	6.75	5.00	5.46	/
4	6.96	4.77	5.80	4.22

Regarding the low coupling constants, the alanine peptides measured in TFE can possess the conformation like α_R, C_7 or P_{II}. As suggested by Liu et al., the TFE destabilizes P_{II} conformation and concomitantly stabilizes the structures with the intramolecular H-bonds, like C_7 conformation in short chains and α-helix in chains long enough to nucleate this helix [106]. Therefore, if dominated by one conformation, the Ala_2, Ala_3 and Ala_4 in TFE most likely possess the C_7 conformation. However, this should be proved by other techniques.

4.1.5.4 NMR Spectroscopy of Ala in Chloroform

Alanine dipeptide (^{15}N labelled) was measured in chloroform at different concentrations. High concentration dependence of chemical shifts were observed, which indicates the aggregation of peptides at higher concentration (Figure 35, a). The $^3J(H_\alpha,H_N)$ coupling constant for $H_N(1)$ proton increases with increasing concentration. The difference between the $^3J(H_\alpha,H_N)$ at lowest and highest concentrations is 0.14 Hz. With increasing concentration the population of β concentration increases (Figure 35, b).

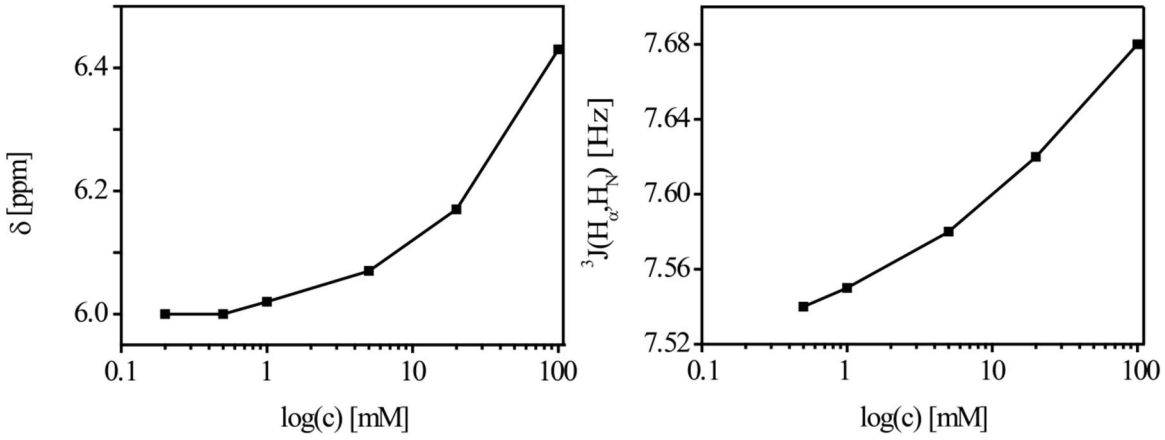


Figure 35: a) The chemical shift of $H_N(1)$ amino proton dependence over 0.2 to 100 mM concentration range of Ala in chloroform. b) The $^3J(H_\alpha,H_N)$ coupling constant dependence over 0.5 mM to 100 mM concentration range of Ala in chloroform, all at 20 °C.

The coupling constants of Ala (^{15}N labelled) in chloroform was measured at different temperatures and concentrations. Figure 36 shows the $^3J(H_\alpha,H_N)$ coupling constants from -20 °C to 50 °C for the H_N proton of the Ala at 100 mM concentration. From Figure 36 it is evident that from -20 to 0 °C the $^3J(H_\alpha,H_N)$ increases while further heating induce the decrease of the $^3J(H_\alpha,H_N)$ coupling constant, suggesting the environment of amino proton at low temperatures differ from the one at temperature above °C and can indicate the different intermolecular H-bond pattern of Ala in chloroform. For 0.5 and 5 mM concentration of Ala the increase of $^3J(H_\alpha,H_N)$ coupling constant was observed for all measurable temperatures.

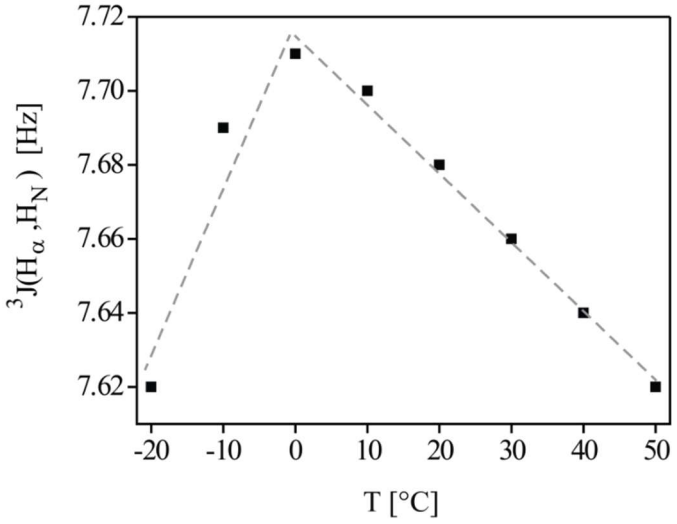


Figure 36: The $^3J(H_\alpha,H_N)$ coupling constants measured in temperature range from -20 to 50 °C for 100 mM Ala in chloroform.

Contrary to the case in aqueous solutions, a small temperature dependence of the chemical shifts in chloroform does not mean the NH group is shielded nor does a large temperature dependence mean the NH group is exposed [200]. The temperature coefficient of 100 mM Ala in chloroform is not constant over the entire temperature range. We can define two temperature regions, the first from -40 to 0 °C and the second one from 0 to 50 °C. At lower temperatures, the values of temperature coefficient for 100 mM Ala are -22.1 and -19.9 ppb/°C for $H_N(1)$ and $H_N(2)$, respectively, while at higher temperatures the $\Delta\delta/\Delta T$ is -8.1 and -8.8 ppb/°C for $H_N(1)$ and $H_N(2)$, respectively (Table 16). The values for temperature coefficients observed for Ala at 0.5 and 5 mM concentration were quite different. Similarly as at 100 mM, at 5 mM concentration the temperature coefficients are again divided into two temperature ranges. At lower temperatures the $\Delta\delta/\Delta T$ values were -7.9 and -8.5 for $H_N(1)$ and $H_N(2)$, respectively, whereas at higher temperatures the $\Delta\delta/\Delta T$ was -3.8 and -4.0 for those amino protons. At 0.5 mM concentration, the Ala in chloroform shows the resolved amino bands only above 0 °C, where the temperature coefficients were -3.2 and -2.9 for $H_N(1)$ and $H_N(2)$, respectively. All these temperature coefficient measurements indicate that at least three different processes take place with the increasing temperature and concentration. Precisely, the changes in inter- or intramolecular H-bonding of amino protons can take place where NH protons become bonded or non-bonded. Since the lack of intramolecular H-bond in Ala in chloroform was proved by experiment of titration of Ala in chloroform with DMSO-d_6, the different temperature coefficients can be explained by assembly of peptides into small aggregates. While the $^3J(H_\alpha,H_N)$ coupling constant of Ala in chloroform is high and it even

increases with peptide concentration the dimer or any other assembly formed at higher concentration predominantly possesses β conformation.

Table 16: The temperature coefficients for Ala in chloroform at 0.5, 5 and 100 mM concentrations. The temperature coefficients are divided in two temperature ranges, first from -40 to -10 °C and second from 0 to 50 °C. At low temperatures and 0.5 mM concentration the amino protons are overlapped, which precludes obtaining their chemical shifts.

Temperature	Δδ/ΔT [ppb/°C]					
	0.5 mM		5 mM		100 mM	
	$H_N(1)$	δ $H_N(2)$	$H_N(1)$	δ $H_N(2)$	$H_N(1)$	δ $H_N(2)$
from -40 to -10 °C	/	/	-7.9	-8.5	-22.1	-19.9
from 0 to 50 °C	-3.2	-2.9	-3.8	-4.0	-8.1	-8.8

The presence of intramolecular H-bond can be proved by adding small amounts of DMSO to the 0.5 mM solution of Ala in $CDCl_3$. If the H-bond is present no or very small chemical shift of amino protons should occur after titration with DMSO while the significant chemical shift induced by DMSO indicates the non-H-bonded NH protons [201]. As evident from Figure 37 the DMSO-d_6 induces a large chemical shift of amino protons of Ala in chloroform indicating that NH groups are not involved in the intramolecular H-bonding.

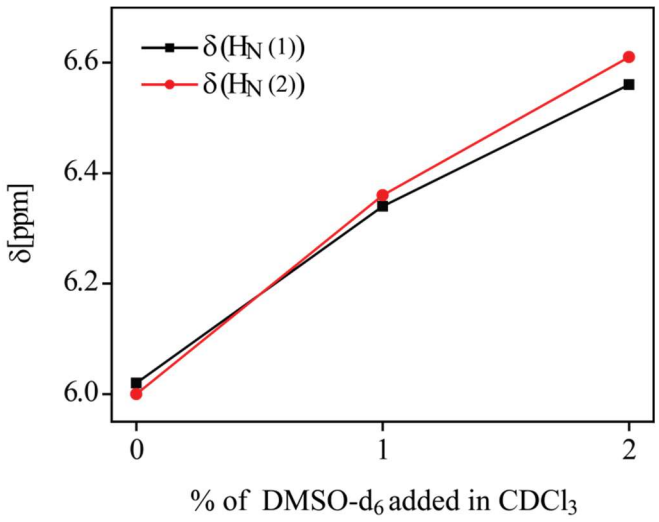


Figure 37: Chemical shifts of amino protons of Ala in $CDCl_3$ at 0.5 mM concentration and 20 °C after adding small amounts of DMSO-d_6.

4.2 Discussion about Short Alanine Peptides

Our investigation of the conformational ensemble sampled by the short alanine peptides suggested a conformational energy landscape of several minima that can be affected by different solvents by stabilization or destabilization of inter- and intramolecular H-bonds.

4.2.1 Short Alanine Peptides in Water

When Ala is placed into the water, the conformational preferences depend on the local electrostatic energy of the peptide backbone and the electrostatic solvation free energy at that site, which depends on the screening interaction of water dipoles with the polar peptide group [51]. In contrary to nonpolar solvent, the local electrostatic energy difference between P_{II} and β conformations in water is close to zero. The reason lies in the extraordinary ability of water to screen electrostatic interactions by the backbone solvation where electrostatic solvation free energy (ESF) is dependent upon side chain [51, 54]. A small residue like alanine provides a better access to water leading to a larger negative ESF [51]. Therefore, water interaction to the peptide backbone must have a crucial role in the stabilization of Ala conformation. To show the water effect on conformational population of Ala the DFT calculations were performed with explicit adding 4 water molecules. The most stabilized conformation was the P_{II} where water molecules oriented in the way to form water bridges connecting the C=O and NH groups as indicated in Figure 23. By adding four explicit water molecules to the B3LYP/6-31G(d,p) method we have accounted for the effect of the hydrogen bonding and we have explicitly shown that the apparent enthalpic interaction between water and the alanine peptide is significantly stronger in the P_{II} conformation than in the extended β, α_R or C_7 conformation. Thus, the P_{II} structure fully utilizes the hydrogen bonding capacity of the C=O and NH groups, maximizes peptide-water cooperativity, and leaves the first solvation layer of hydration able to participate in further hydrogen bonding with the next solvation shell [202]. Mezei et al. suggested that P_{II}-helices produce less disruptive effect on surrounding water organization as compared to β-strands [203]. With respect to the stabilization of the P_{II} conformations by water, Poon et al. [103], hypothesized that bridging water molecules are responsible for Ala adopting the P_{II} conformation, whereas contrary opinion was provided by Law and Daggett who suggested that water bridges do not stabilize the P_{II} structures [204]. We suggest that P_{II} conformation enables the most favourable solvation conditions of the Ala with a water bridges connecting the C=O and NH groups that produce a large permanent

dipole moment. While alanine residue has small side chain that has minor effect on the shielding of the backbone, the water dipoles can exhibit a great screening of the local backbone electrostatic interaction, and thus stabilizing the P_{II} conformation. Our results are in agreement to ESM theory that predicted high P_{II} content for alanine in water based on large contribution of ESF [51].

In addition to the *ab initio* calculations, experimental methods like IR, Raman, VCD, NMR and UV-CD spectroscopy were employed for the characterization of conformational populations of Ala, Ala_2, Ala_3 and Ala_4 in water. NMR measurements revealed that Ala_2 populates more β conformation than Ala_3 which populates more β than Ala_4. To obtain residual information about the peptide conformation, we determined the $^3J(H_\alpha,H_N)$ coupling constants that directly correlate to the ϕ dihedral angel. Based on Equation 6 it is evident that $H_N(1)$ amino proton in Ala_2 possesses less β conformation that $H_N(2)$ amino proton while together Ala_2 do not exceed the 33 % of β conformation of overall population. When comparing the $^3J(H_\alpha,H_N)$ coupling constants that are connected to the β basin and the fractions of β conformation obtained from amide III region in IR and Raman spectra we get a very good correlation ($R^2 = 0.98$) (Figure 38).

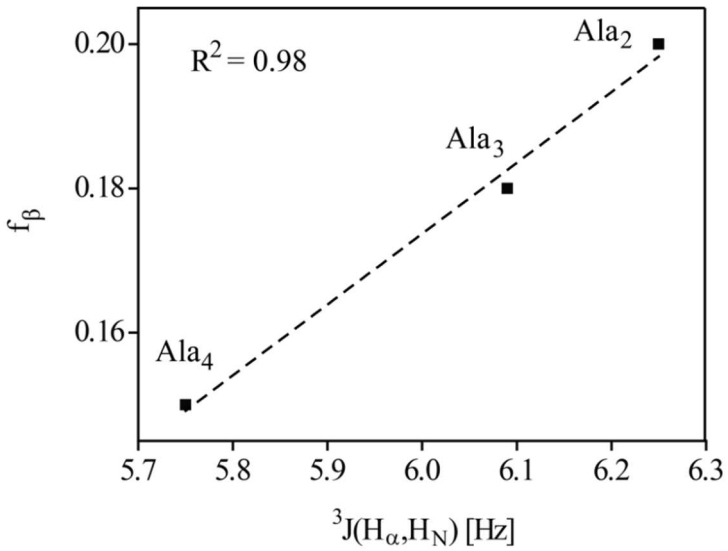


Figure 38: The fraction of β conformation in Ala_2, Ala_3 and Ala_4 in water obtained from IR and Raman spectral fitting algorithm of amide III region plotted against the average NMR $^3J(H_\alpha,H_N)$ coupling constants.

The assignment of the structural sensitive bands in amide III region in IR and Raman spectra of short alanine peptides in water is based on previous results on alanine dipeptide [45] while the assignment of the band for C_7 conformation is not so conclusive. The assignation for C_7

conformation was suggested based on the results obtained from measurements of short alanine peptides in different solvents (chloroform, TFE) and DFT calculations. The fitted model bands in amide III region of short alanine peptides show the increase in C_7 conformation with longer peptide chain. Moreover, loosing of the band intensity of the negative CD band at 196 nm (Figure 39) and, simultaneously, increasing of negative CD band at 215 nm indicates the losing of P_{II} conformation together with the increasing of C_7 conformation with longer peptide chain. However, based on UV-CD spectra the presence of β-turns that also have the negative band around 215 nm [106] cannot be excluded. Additionally, the VCD spectral bands in amide III region again indicate the losing of P_{II} conformation with longer peptide chain. However, the data from the literature suggest the increase of P_{II} population with the increasing length of short alanine peptides obtained from the vibrational spectra from amide I region [60, 205] or from UV-CD spectra [46]. The argument to these contradict data are the small sensitivity of the amide I band of IR spectra of short peptides whereas the UV-CD spectrum shows an average spectral shape of all conformations sampled by one molecule leading to the overlapping or subtractions of the UV-CD bands. Additionally, the UV-CD spectral band shapes for conformations like C_7 are not well established for the short peptides. In the literature, short alanine peptides usually have the charged terminus resulting in repulsion between both ends of peptides and thus preferred population of the extended P_{II} conformation [40, 61, 106].

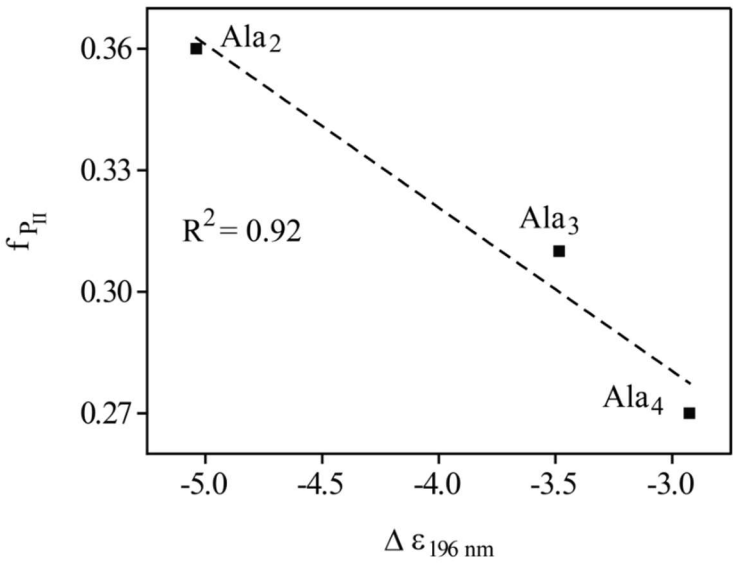


Figure 39: The fraction of P_{II} conformation in Ala_2, Ala_3 and Ala_4 in water obtained from IR and Raman spectral fitting algorithm of amide III region plotted against the minimum of UV-CD band at 196 nm.

We observed the thermal transition from P_{II} to β conformation for all measured alanine peptides. The increase in $^3J(H_\alpha,H_N)$ coupling constants with the temperature indicates a partial thermal transition from one backbone conformation with lower coupling constant (P_{II}, C_7, α_R) to the β conformation. Additionally, the UV-CD spectra between 15 and 60 °C show the decreasing in the band intensity at 196 nm indicating the losing of P_{II} conformation while the difference spectrum resembles of β structure (Figure 31, *Inset*). The increase in $^3J(H_\alpha,H_N)$ coupling constant for $H_N(1)$ amino proton of Ala_2 from 6.18 Hz at 2 °C to 6.51 Hz at 40 °C corresponds to the additional population of 10 % of β conformation. The increased temperatures have the major impact on Ala_4 that transforms above 18 % of P_{II} conformation at 2 °C into the β conformation at 60 °C. The melting of P_{II} conformation to β structure of short alanine peptides was also proposed by Chen et al [52]. Tiffany and Krimm [36] suggested that at a high enough temperature, a statistical random coil-like conformation can be attained. However, our studies of short blocked alanine peptides do not support this prediction.

The observation of C_7 conformation in the Ala_2, Ala_3 and Ala_4 in water was never reported before. We caution that some results presented in this chapter are contradictory. For example, the low-frequency amide I band of Ala_3 and especially of Ala_2 indicate a strong intramolecular H-bond, while the NMR values of the temperature coefficients rather suggest the solvent exposed amino protons in these peptides. Therefore, we cannot say that C_7 conformation in water is stabilized with intramolecular H-bond rather it stands as more opened structure as it is observed from DFT calculations with explicit water molecules.

The increase in proportion of C_7 conformation with the longer peptide chain of short alanine peptides can be explained with the cooperative effect. The cooperative formation of C_7 conformation in short alanine peptides cannot be elucidated with the Flory's random coil model [28] or with the empirical coil library model of unfolded proteins [57]. These two models use Flory's isolated-pair hypothesis that forbids coupling between neighbour residues. On the other hand, the ESM theory emphasizes the importance of the screening of the backbone electrostatic interactions by water that is most influenced by neighbouring residue and has a big effect on the conformation [110]. This model can explain the cooperative formation of local structures in unfolded ubiquitin where the fluctuating β-strands were observed [54]. The electrostatic interactions between C=O and NH groups are the most favourable in C_7 conformation whereas the water screening shouldn't exceed this tendency. In contrary to alanine dipeptide, where solvation plays the dominant role in conformational preference, the stabilization of Ala_2, Ala_3 and Ala_4 comes from the electrostatic stabilization

between C=O and NH groups in C_7 conformation (in literature also known as inverse γ-turn). Schweitzer-Stenner and Measey [63] reported that the turn population is length-dependent, with only a small fraction in trimers and tetramers but a large fraction in sevenmers. Their focus was on β-turns, but the same conclusion would hold for inverse γ-turns. Kallenbach and co-workers [106] found that the short pentapeptide $AcGGAGGNH_2$ can be transformed from P_{II} toward an increasing population of H-bonded γ- or β-turns by adding the selected co-solvent that produce graduated reductions in solvent quality. Furthermore, Gong and Rose [206] showed the detectable population of γ-turn in blocked polyalanyl heptamers in poor solvent. However, more recent work–both experimental and theoretical–reports a significant population of P_{II} conformation in unfolded proteins whereas the C_7 conformation has often been ignored or was put into the P_{II}-basin when subdividing the ϕ,ψ-map. Ramachandran and Sasisekharan [207] distinguished the P_{II} conformation from C_7 since they occupy distinct, nonintersecting regions of the ϕ,ψ-map, and they are chemically distinct as well; while P_{II} conformation is stabilized by the solvent, the stability of C_7 conformation is accomplished with formation of intramolecular H-bond or electrostatic interaction.

Altogether, both intermolecular H-bonds and electrostatic interactions contribute to the stabilization of conformations in short alanine peptides in water where the conformational distribution arises from the competition between these effects. It is suggested that the Ala prefers a conformation that would produce a large permanent dipole moment, so that it can be solvated well, like is in the case of P_{II} conformation, while by increasing the peptide length of alanine peptides the intrachain electrostatic interaction predominate. To extrapolate these results to the unfolded state with a dynamic ensemble of conformations, our results suggest the need for a reassessment of electrostatic interaction between the backbone groups in unfolded proteins under solvent conditions of experimental interest such as molten globules [208].

4.2.2 Ala in Chloroform

Polar molecule like Ala undergoes large changes in dipole moment on transfer from polar (water) to nonpolar (chloroform) solvent. A nonpolar solvent induces weaker screening of the backbone electrostatic, thus the local electrostatic energy plays the dominant role in the conformational preferences of peptide in chloroform. ESM suggests that the local backbone electrostatic energy is more favourable for a residue in β conformation that has an antiparallel alignment of peptide dipoles than in the α_R conformation where the peptide dipoles are

aligned parallel [54]. Using PCM method we calculated the total ground state energy of the alanine dipeptide in chloroform where the local electrostatic energy determines the conformational distribution. The results showed that the most favourable structure is C_7 conformation with 82 % of the total population. The C_7 conformation enables the orientation of methyl group in the equatorial position resulting in favoured internally H-bonded structure as observed in Figure 22. In the remaining 18 % of the structures, the Ala occupies the β conformation which is for 0.89 kcal/mol less energetically favourable than C_7 conformation. Similar results were obtained with calculation the Ala in gas phase where even more population of C_7 conformation was observed. These results are in agreement with ESM theory that supports the dominate role of Columb's electrostatic interactions in the conformational equilibrium of Ala in the gas phase [100]. By application of IR and NMR spectroscopy we observed strong concentration dependence of Ala in chloroform. High value of $^3J(H_\alpha,H_N)$ coupling constants and low-frequency amide III bands at 1235 and 1222 cm^{-1} indicate the formation of small soluble aggregates in β conformation. We propose the formation of low-molecular-weight oligomer structure of Ala at low concentration and formation of higher oligomeric structures at high concentration of peptide. In contrary to the monomer structure of Ala, where C_7 conformation is preferred, DFT calculations revealed that the β dimer is for 0.64 kcal/mol energetically more favoured than the C_7 dimer, suggesting the β oligomerization as an onset of the early nucleation structure in the process of self-aggregation. The temperature coefficients, obtained from 1D NMR spectra, show the three different processes when increasing the peptide concentration and the temperature. We suggest that at very low concentration (0.5 mM) the Ala in chloroform stands as a monomer while with increasing the concentration the β oligomerization takes place. This process is well described by the values of the temperature coefficients. Amino protons in monomer structure have the lowest value of temperature coefficient (~3 ppb/°C). The low-molecular-weight oligomer structure of Ala (5 mM and low temperatures) has the value of temperature coefficient around 8 ppb/°C while at higher oligomeric structure of Ala (100 mM and low temperatures) this value increases to ~20 ppb/°C (Table 11).

4.2.3 Short Alanine Peptides in DMSO

DMSO is strong proton acceptor and tends to break H-bonds between the C=O and NH groups, thus stabilizing a conformer with a relatively open structure and allowing the coordination of the solvent molecules around the amide groups. Using DFT calculations with

explicit DMSO molecules we showed that H-bonding between DMSO and NH peptide groups have strong effect on the calculation of VCD spectrum. The calculated spectrum has big similarity with experimental one that is assigned for the β conformation. Pohl et al. also showed that Ala in DMSO possesses the β conformation obtained from analysing the amide I region of IR spectrum [196]. Additionally, all investigated short alanine peptides have the amide III band at 1277 cm^{-1}, characteristic pattern of UV-CD spectrum and high $^3J(H_\alpha,H_N)$ coupling constants that are all evident for a high population of β conformers.

Contrary to Raman, IR spectroscopy is very sensitive for determination of H-bonds of NH groups in amide II and amide III region. The band at 1240 cm^{-1} is not assigned to any conformations rather it represents the aggregated form of peptide. The band increases with the longer peptide chain suggesting the aggregation process is more pronounced at longer peptides. Additionally, the temperature coefficient measurements of alanine peptides in DMSO obtained from NMR spectra showed the presence of slightly solvent shielded NH protons for Ala_4 indicating the formation of aggregated structure of peptide. The intensity of the band at 1240 cm^{-1} correlates with the average temperature coefficient of the individual alanine peptide (Figure 40). This suggests that the aggregated forms of peptides are soluble aggregates with only a small number of participated peptides. Additionally, this plot suggests that the temperature coefficient is sensitive to aggregation process as was also observed in the case of Ala in chloroform.

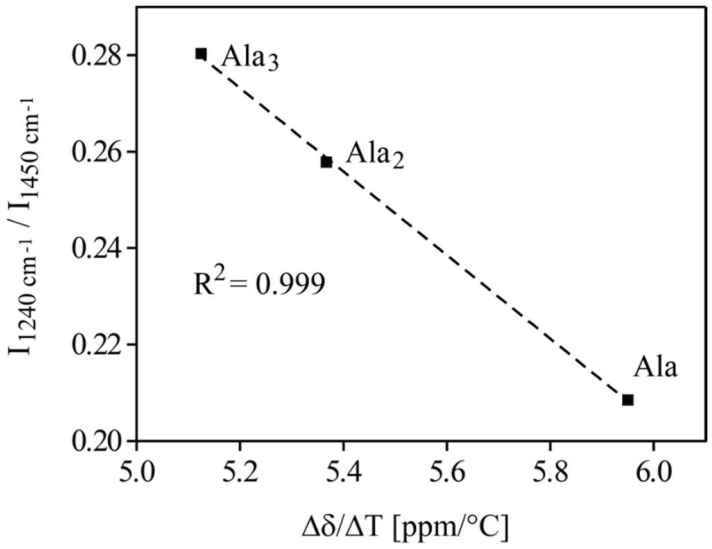


Figure 40: The fraction of aggregated peptide obtained from the intensity of the band at 1240 cm^{-1} relative to the reference band at 1450 cm^{-1} (CH_3 deformation) divided to the number of CH_3 groups in individual peptide vs. average temperature coefficient obtained from the chemical shifts of amino protons in 1D NMR spectra of the Ala, Ala_2 and Ala_3 in DMSO-d_6.

4.2.4 Short Alanine Peptides in TFE

TFE is a good proton donor that decrease accessibility of water by clustering the amide bond leading to stronger intramolecular H-bonds of helices and turns and promote formation of bifurcated H-bonds (one carbonyl group accepting two hydrogens) or H-bonds between amide groups and solvent molecules [209]. Luo and Baldwin [210] have proposed that desolvation of the peptide backbone in the helical conformation strengthens the intrahelical H-bonding, thereby stabilizing the helix. It was suggested that TFE destabilizes the P_{II} conformation resulting in formation of internally H-bonded alternative structures, C_7-turn in short chains and α-helix in chains long enough to nucleate this helix [106]. In agreement to the work published by Liu et al. [106], we propose the presence of C_7 conformation in short alanine peptides in TFE, as revealed by low $^3J(H_\alpha,H_N)$ coupling constants, UV-CD spectral shape and amide III band at 1260 cm^{-1}. Generally, alanine is an α-helix promoting amino acid [211], however, we showed that till four residues of alanine, the peptide chain is not long enough to nucleate the α-helix. Kallenbach and co-workers shoved that in the presence of TFE the peptide with seven alanine residues forms α-helix [106]. To obtain the exact number of alanine residues that can nucleate the α-helix, further studies needs to be performed with five and six alanine residues.

4.2.5 Methodology Used for Structural Characterization of Short Alanine Peptides

All the results of short alanine peptides obtained in our study demonstrate that the combination of vibrational spectroscopy approach with NMR and UV-CD data are 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 III band profiles concerning even modest variations of the conformational population. It is conceivable that due to the average values of measured NMR and UV-CD signals over all conformations populated of one amino acid residue and the large differences in the time scales of different spectroscopic methods, variations in conformational populations were obtained between them. In contrary to the NMR spectroscopy, which is capable of giving structural information of peptide at the atomic level of resolution, UV-CD is a low resolution structural technique in which overall structural features are described.

The experimentally obtained results were supplemented by the theoretical approach using

DFT method as complementary tool to determine the conformational equilibrium and mechanisms of structural stabilization of Ala. An agreement between the experimental and calculated VCD spectra improved significantly upon the addition of explicit solvent molecules to the computational models in case of polar solvents, providing hydrogen bonding between solvent and NH and C=O peptide groups. For the first time the VCD spectra in amide III region were characterized for all conformations found in short alanine peptides (Table 17).

Table 17: Assignment of the amide III region in which four backbone conformations (P_{II}, α_R, β and C_7) are found for Ala with the characteristic VCD band shape.

Conformation	Amide III [cm^{-1}]	VCD band shape of amide III
P_{II}	1317–1306	Positive (+)
α_R	1304–1294	Positive (+)
β	1294–1270	Negative (-)
C_7	1265–1258	Negative (-)

5 Structural Characterization of Poly-L-Lysine

5.1 Results of PLL

5.1.1 FTIR Spectral Assignment of Different Conformers of PLL

5.1.1.1 PLL in Water (pH 11.6)

A high pH value (11.6) and low temperature stabilize the α-helical conformation of PLL in water. The band fitting algorithm resolves the amide I band on two band components at 1664 cm^{-1} and 1644 cm^{-1} respectively (Figure 41). The strongest model band at lower wavenumbers is ascribed to the α-helix conformation. The band frequency of this amide I band is located close to the α-helix band found in the IR spectrum published by Dzwolak et al. [212]. However, in the α-helical state, there is an additional band in the amide I region at 1664 cm^{-1}, the assignment of which is not conclusive. One of the possibilities, which should be proved by the presence of corresponding bands in the amide II and amide III regions, is that this amide I band belongs to turns [209]. The decomposition of the amide II band revealed two prevailing components located at 1550 cm^{-1} and 1572 cm^{-1}. These model bands are assigned to the α-helix conformation and the amide II counterpart of the 1664 cm^{-1} amide I band. The remaining bands presented in Figure 41 are due to the NH_2 symmetric and antisymmetric deformation of the lysine side chain located at 1602 cm^{-1} and 1532 cm^{-1} respectively. The remaining band in the amide I region located at 1628 cm^{-1} is attributed to the bending vibration of water molecules, which remains after the subtraction of the bulk water.

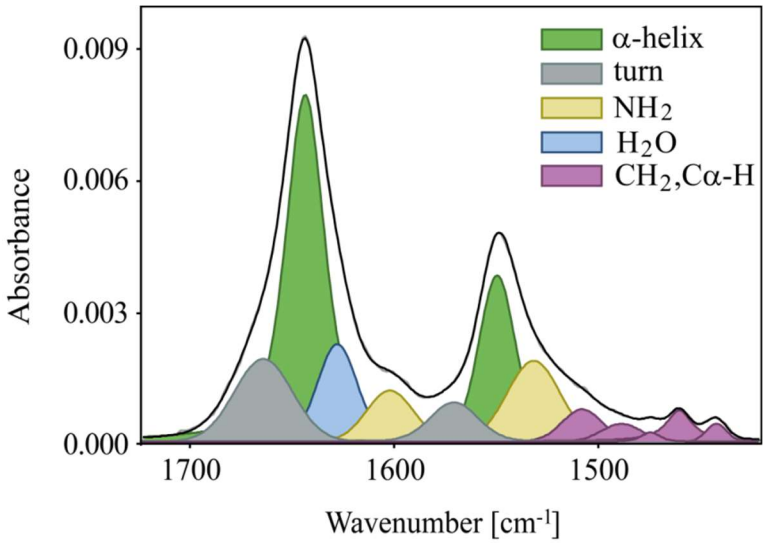


Figure 41: The IR spectrum of PLL at pH 11.6 and T = 10 °C with fitted model bands that represent: the α-helix (1644 cm^{-1}, 1550 cm^{-1}), the turn conformation (1664 cm^{-1}, 1572 cm^{-1}), the CH_2 and C_α-H vibrations (1508 cm^{-1}, 1489 cm^{-1}, 1474 cm^{-1}, 1461 cm^{-1}, 1443 cm^{-1}), the NH_2 groups of the lysine side chain (1602 cm^{-1}, 1532 cm^{-1}) and the water bending (1628 cm^{-1}).

The band structure of the extended amide III region is presented in Figure 42. At the high frequency region, two sharp model bands are assigned to the CH_2 (1394 cm^{-1}) and C_α-H (1348 cm^{-1}) bending vibrations. The present assignment of the band at 1348 cm^{-1} is not unique. While Barron et al. [88] assigned this band as a component of the amide III region representing the vibration of hydrated α-helices, vibrational studies of short dipeptides showed that this band is already out of the amide III region [45]. It is predominately attributed to $C_\alpha H$ bending mixed with the amide III mode but has no potentials for the structural characterizations of peptides [213]. Therefore, the amide III region of the PLL is composed of two band components. The strongest band at 1295 cm^{-1} has been assigned to the α-helical conformation of the polypeptide. The band near 1300 cm^{-1} is generally found in the vibrational spectra of globular proteins composed mainly of α-helices [214-216]. Moreover, the component of the amide III vibration in the IR spectra of dipeptides, which was assigned to an α conformation of the peptide backbone [45], was found at a similar position. The intensity of the second band in the amide III region located at 1257 cm^{-1} (Figure 42) is much lower compared to the central band at 1295 cm^{-1}. Its appearance is correlated to the bands at 1664 cm^{-1} and 1572 cm^{-1} in the amide I and amide II regions respectively, which were assigned to turn structures.

The area of the component bands can be efficiently applied for a rough estimation of the

populations of a particular conformation. This analysis revealed that the population of α-helix in PLL at 10 °C and at pH 11.6 is between 78 % and 81 % as obtained from the amide I or amide III region, respectively. These results are in accordance with results published by Jiji et al. [49].

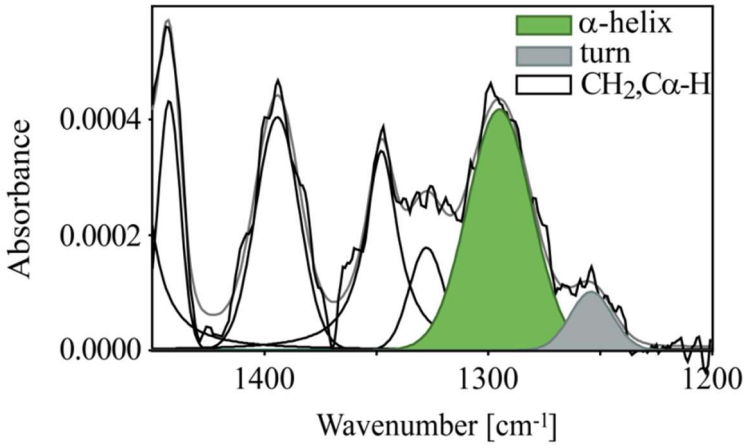


Figure 42: The IR spectrum in an amide III region of PLL at pH 11.6 and T = 10 °C with fitted model bands that represent: the α-helix (1295 cm^{-1}), the turn conformation (1257 cm^{-1}) and the CH_2 and C_α-H vibrations.

At higher temperatures and pH 11.6, helical PLL gradually transforms into the β-sheet conformation. The IR spectrum of β-sheet PLL (Figure 43) shows two characteristic amide I bands, weaker at 1692 cm^{-1} and stronger at 1618 cm^{-1} that are typical for the (aggregated) antiparallel β-sheet.

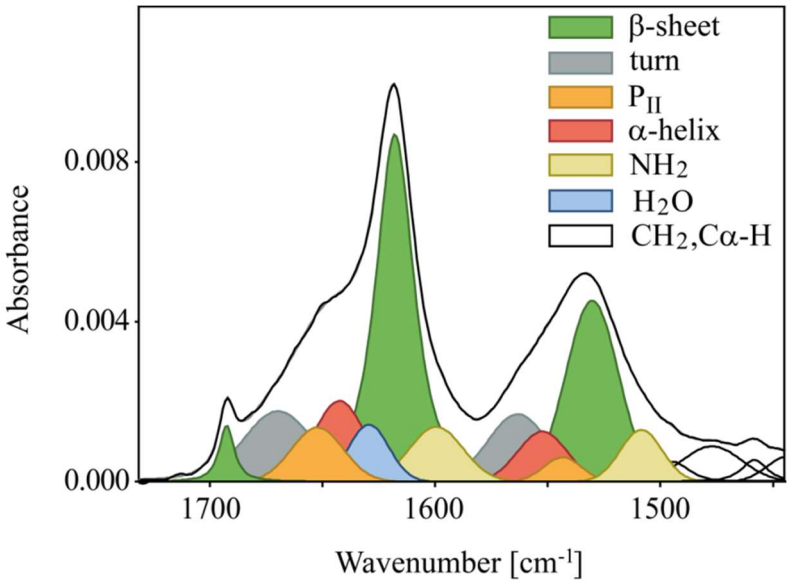


Figure 43: The IR spectrum of PLL at pH 11.6 and T = 80 °C with fitted model bands that represent: the β-sheet (1692 cm^{-1}, 1618 cm^{-1}, 1530 cm^{-1}), the turn conformation (1669 cm^{-1}, 1563 cm^{-1}), the P_{II} conformation (1652 cm^{-1}, 1543 cm^{-1}), the α-helix (1642 cm^{-1}, 1553 cm^{-1}), CH_2 and C_α-H vibrations (1495 cm^{-1}, 1477 cm^{-1}, 1458 cm^{-1}, 1444 cm^{-1}), the NH_2 groups of the lysine side chain (1600 cm^{-1}, 1508 cm^{-1}), the water bending (1629 cm^{-1}) and the CH_2 and C_α-H vibrations.

In the decomposed amide III region presented in Figure 44, three components of different β conformations can be observed; the first component is located at 1270 cm^{-1}, which is assigned solely to amino acids in the β conformation that are not involved in the secondary structure [45], the second is located at 1243 cm^{-1} and is assigned to the β-strands and third is located at 1222 cm^{-1} and is assigned to the strong hydrogen bonded β-sheets of aggregated PLL. The frequency of the second component is close to the value of the amide III band found in the IR spectra of β structured proteins [87], while the amide III components near 1220 cm^{-1} have already been found in the spectra of aggregated proteins [87, 217-219]. The bands representing the β-sheet conformation populates 61 % and 58 % of all the conformations obtained from the amid I and III region respectively. The structure of the amide I, amide II and amide III bands revealed that besides the prevailing β-sheets, the final state of the fibrillation of PLL also contains some amino acids in the turn, α-helical and P_{II} conformations.

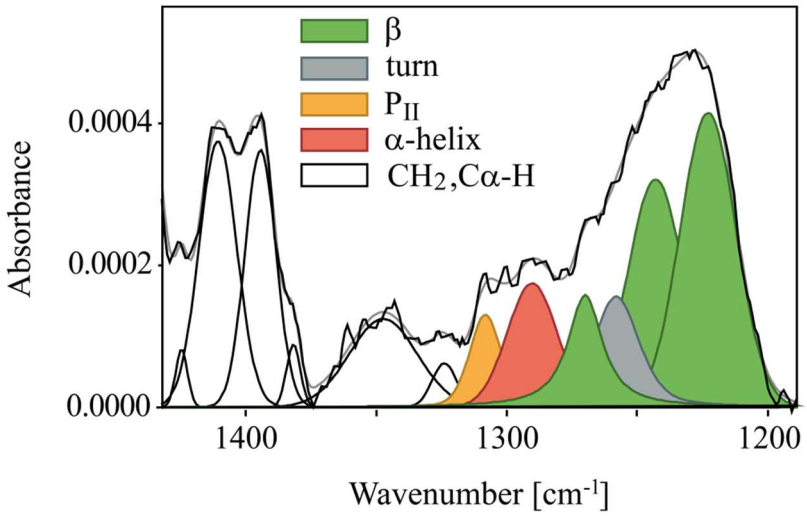


Figure 44: The IR spectrum in an amide III region of PLL at pH 11.6 and T = 80 °C with fitted model bands that represent: β conformations (β conformation at 1270 cm^{-1}, β-strand at 1243 cm^{-1}, hydrogen bonded β-sheet at 1222 cm^{-1}), the turn conformation (1258 cm^{-1}), the P_{II} conformation (1308 cm^{-1}), the α-helix (1290 cm^{-1}) and the CH_2 and C_α-H vibrations.

5.1.1.2 PLL in Water (pH 4)

The PLL side chains at pH 4 are fully ionized, so electrostatic repulsion prevents the formation of the α-helical conformation. The characteristic IR spectrum of PLL in a water solution at 10 °C and a low pH shows broadened and asymmetrical amide I and amide II bands (Figure 45). Similar to at a higher pH, the fitting algorithm is used to determine the intrinsic components of the amide bands as shown in Figure 45. The band analysis shows that the major population of PLL at pH 4 is the P_{II} conformation (45 %) followed by the β-structures. The amide I and II band components attributed to P_{II} are located at 1649 cm^{-1} and 1547 cm^{-1} respectively. The corresponding bands for β structures can be found at 1618 cm^{-1} and 1522 cm^{-1}, while small bands at 1644 cm^{-1} and 1671 cm^{-1} were assigned to the α-helix and turn structures. The amide II counterpart of the later structure is located at 1570 cm^{-1}. The bands at 1594 cm^{-1} and 1506 cm^{-1} are assigned to the NH_3^+ stretching of the lysine side chains. The assignment of those bands was checked with deuterium substitution of the solvent.

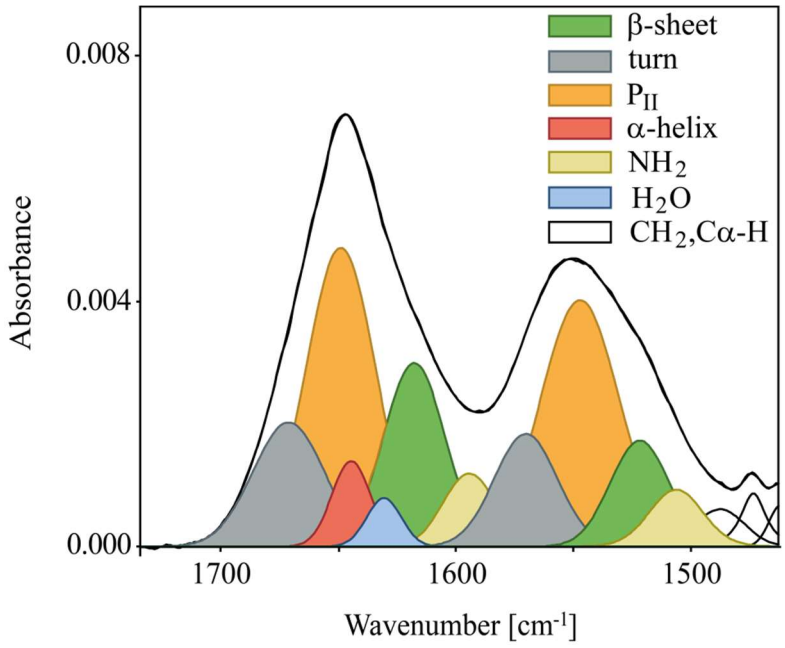


Figure 45: The IR spectrum of PLL at pH 4 and T = 25 °C with fitted model bands that represent: the β-sheet (1618 cm^{-1}, 1522 cm^{-1}), the turn conformation (1671 cm^{-1}, 1570 cm^{-1}), the P_{II} conformation (1649 cm^{-1}, 1547 cm^{-1}), the α-helix (1644 cm^{-1}), the NH_3^+ groups of lysine side chain (1594 cm^{-1}, 1506 cm^{-1}), the water bending band (1630 cm^{-1}) and the CH_2 and C_α-H vibrations.

Similar structural elements of PLL conformation in water at a low pH was found by decomposing the amide III region. The extended amide III region of the FTIR spectrum of PLL in water at pH 4 shows at least four well-defined bands with the frequencies 1311 cm^{-1} (P_{II}), 1291 cm^{-1} (α-helix), 1272 cm^{-1} (β conformation), 1257 cm^{-1} (turn), 1243 cm^{-1} (extended β-strand) and 1219 cm^{-1} (aggregated β-sheet). However, it needs to be emphasised that the decomposition of the amide III region presented in Figure 46 yielded slightly different distributions of the presented conformers. By comparison of the integral intensity, the ratio of the P_{II} conformation falls to 25 %. The reason for this inconsistency is not known. The comparison of the results obtained from the analysis of amide I and amide III region reveals that the decomposition of the latter region retrieves more complete information about the populations of different conformers of PLL. The band components of the amide III region are more resolved already in the original spectrum. Consequently, the application of the band fitting algorithm is therefore more reliable. This is the reason why we are able to determine some conformations connected to the β-structure, in which the characteristic patterns are strongly overlapped in the amide I region with water and NH_3^+ bending.

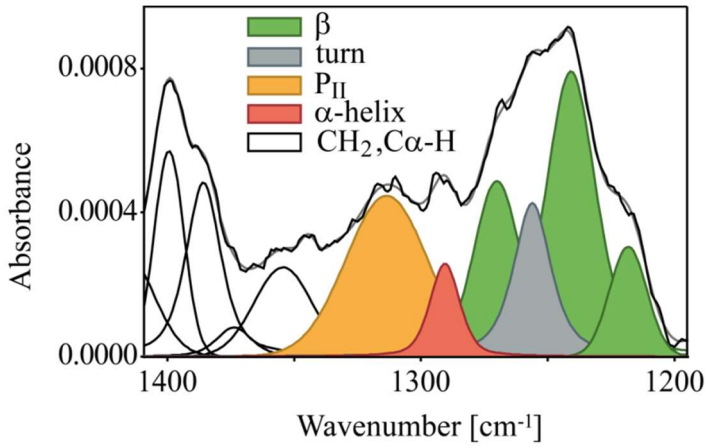


Figure 46: The amide III region of the IR spectrum of PLL, pH 4, at 20 °C. fitted with model bands that represent: β conformations (β conformations at 1272 cm^{-1}, β-strands at 1243 cm^{-1}, aggregated β-sheet at 1219 cm^{-1}), the turn conformation (1257 cm^{-1}), the P_{II} conformation (1311 cm^{-1}), the α-helix (1291 cm^{-1}) and the CH_2 and C_α-H vibrations.

The temperature dependence of the PLL conformation at pH 4 is not as extensive as at pH 11.6. It culminates in a small upshift of the amide I band that corresponds to a small downshift of the amide II band. The frequency of the amide II band is the most sensitive for the formation of hydrogen bonds and consequently to hydration [220]. Therefore the change in hydration of PLL causes a downshift of this band by 16 cm^{-1}. This temperature dependence derives from the hydrogen bonds of water molecules to amide carbonyl and N–H groups [220-222]. With a higher temperature, the water–amide hydrogen bond strengths decrease, which consequently form stronger carbonyl bond resonance and weaker C(O)–N bonding resonance of the peptide bond. Comparison of the band intensities of the bands corresponding to the P_{II} conformation reveals that heating PLL at a low pH from 20 °C to 80 °C decreases the population of P_{II} by roughly 10 %.

5.1.1.3 pKa Calculation in Different Solvents

The experimental values of pKa for PLL in TFE and ethylene glycol are not known. The protonation state of the side chain amino group of PLL in different solvents was determined indirectly by comparison with the calculated pKa values of model molecule propylamine using the DFT approach. Propylamine was used as a model molecule for the side chain of lysine possessing a similar pKa value in water as PLL. Solvation free energy was calculated using the polarized continuum model (PCM) with the B3LYP/6-31++G(d,p) method and calculated using the equation [223]:

$$\Delta\Delta G = \Delta G_{GP}(PropylNH_3^+) - \Delta G_{GP}(PropylNH_2) + \Delta G_{solvent}(PropylNH_3^+) - G_{solvent}(PropylNH_2) + \Delta G_{GP}(solvent\text{-}H^+) - \Delta G_{GP}(solvent) + \Delta G_{solvent}(solvent\text{-}H^+) - \Delta G_{solvent}(solvent) \quad (10)$$

The obtained ΔΔG values were then used to calculate the pKa values, which were compared to experimentally determined pKa values for propylamine in different solvents (Table 18).

Table 18: pKa values for the propylamine model molecule in different solvents obtained *ab initio* from Eq. 1 and experimental values.

Solvent	*pKa calculated*	*pKa experimental*
Water	10.56	10.6[a]
TFE	4.54	/
DMSO	9.76	10.4[b]
Ethylene glycol	1.82	/

[a] Obtained from Perrin et al. [224].

[b] Obtained from Makowska et al. [65].

A very good agreement with the experimental values was obtained for water and DMSO, suggesting a rather good match between the calculated pKa value for ethylene glycol and TFE. The low values of the calculated pKa of propylamine in ethylene glycol and TFE suggest the deprotonation of amino groups from side chains in both solvents.

5.1.1.4 PLL in Different Solvents

PLL was titrated with different ratios of water:TFE at pH 4. With an increasing TFE concentration from 0 % to 80 % (v:v), the spectra of PLL remain unaffected. An additional increase of the TFE concentration induces a shift of the amide I 3band maximum from 1649 cm^{-1} to 1652 cm^{-1}, indicative of P_{II}/α-helix transformation, i.e. the manly P_{II} conformation of PLL in water converts to the α-helical conformation in 100 % TFE. The calculated amount of α-helix conformation in 100 % TFE is 84 %. It should be mentioned that the amide I band characterizing an α-helix of PLL is blue-shifted in TFE due to the low dielectric constant of TFE compared to water. Heating PLL in 100 % TFE from 10 °C to 70 °C results in a decrease of the α-helix conformation by 13 %, which runs parallel with the formation of the P_{II} conformation at 1659 cm^{-1} (Figure 47).

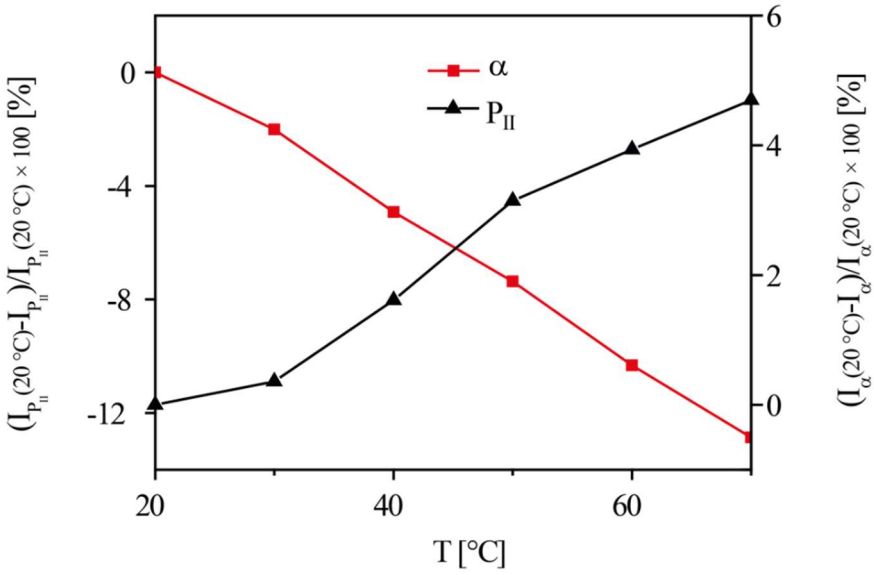


Figure 47: The melting of the α-helix and the formation of the P_{II}-helix with increasing temperature relative to the α-helix and P_{II} content of PLL in TFE at 20 °C.

We measured PLL in DMSO as a solvent with only hydrogen bond acceptor abilities. At low temperatures, PLL in DMSO exhibits an even higher α-helical population than in TFE, i.e. a solvent with hydrogen bond donor abilities. The spectrum of PLL in DMSO is characterized by a sharp amide I band at 1650 cm^{-1}, an amide II band at 1548 cm^{-1} (Figure 48) and an amide III band at 1295 cm^{-1}. The amide III region is only applicable when the DMSO solvent is substituted with DMSO-d_6. The amide I band is upshifted due to the low dielectric constant of DMSO.

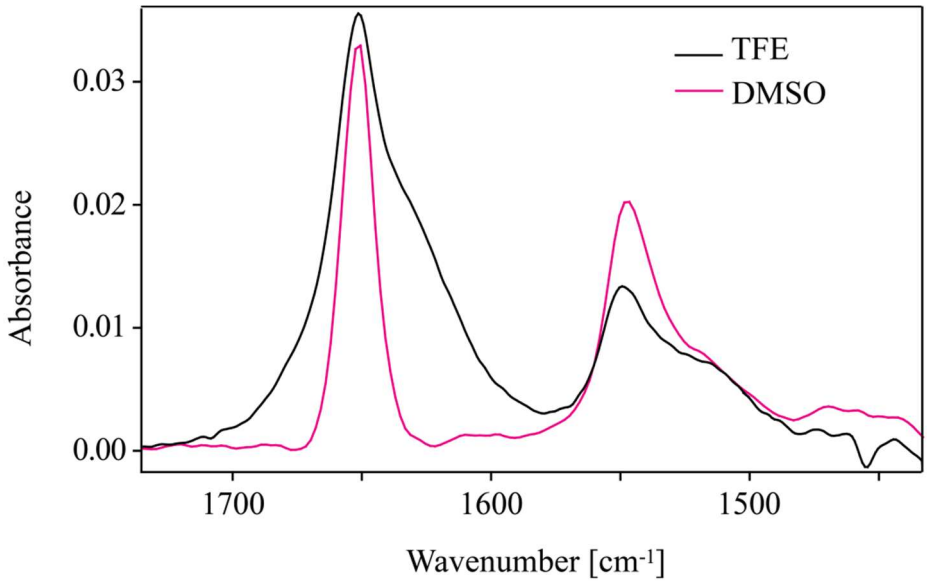


Figure 48: Spectra of PLL in 100 % TFE and in DMSO at 20 °C.

Ethylene glycol was the next solvent where the structure of PLL was probed. As a solvent it has many applications especially as a cryogenic liquid. Calculation of the pKa values showed that the ε-amino group of the lysine side chain in ethylene glycol has a significant potential to be deprotonated. The reduced electrostatic repulsions of the side chains of PLL in ethylene glycol enable the stabilization of P_{II} conformation compared to water, where repulsions between ionized ε-amino groups promotes extended β-strand conformation [93]. PLL in ethylene glycol populates 88 % of the P_{II} conformation, as can be seen in Figure 49. The bands characteristic of the P_{II} conformation are located at 1656 cm^{-1} in the amide I region, 1546 cm^{-1} in the amide II region and 1311 cm^{-1} in the amide III region. Due to the low dielectric constant of ethylene glycol, the amide I band characterizing the P_{II} conformation of PLL is blue-shifted compared to the amide I position in water. From the amide III region in Figure 49, a band is observed at 1238 cm^{-1} that corresponds to the β-strand structure. After heating PLL in ethylene glycol from 10 to 65 °C, 10 % of the P_{II} conformation melts as observed from the difference spectra.

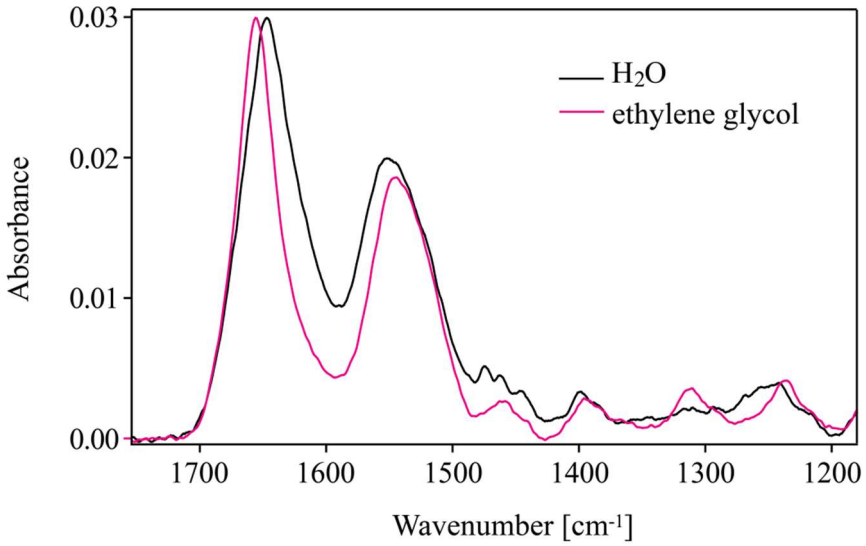


Figure 49: Spectra of PLL in ethylene glycol and in water at pH 4.

5.1.2 Structural Characterization of PLL during Fibrillation

5.1.2.1 Fibrillation of PLL in Water

The IR spectra of PLL at pH 11.6 were measured over a range of the temperatures between 10 °C and 80 °C, as shown in Figure 50. The temperature induced structural transformation from α-helix to β-sheet structure starts at temperatures above 40 °C. Following the spectral changes

the transition temperature was determined near 51 °C, slightly higher than previously reported [87, 225].

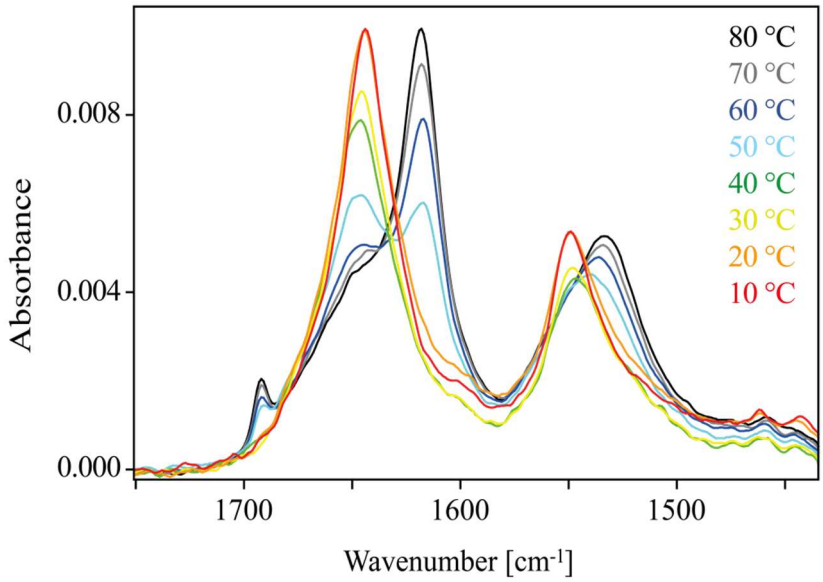


Figure 50: α-helix-to-β-sheet transition of PLL upon gradually heating from 10 °C to 80 °C, monitored by FTIR, 4 wt% solution in H_2O, pH = 11.6. The spectrum of bulk water is subtracted.

The presence of amyloid fibrils of PLL, detected by IR spectroscopy, was confirmed by TEM as shown in Figure 51. The fibrils are assembled to lengths over 1 µm and exhibit widths on the order of 7–12 nm. They have twisted-bundle morphology that was already reported before [226].

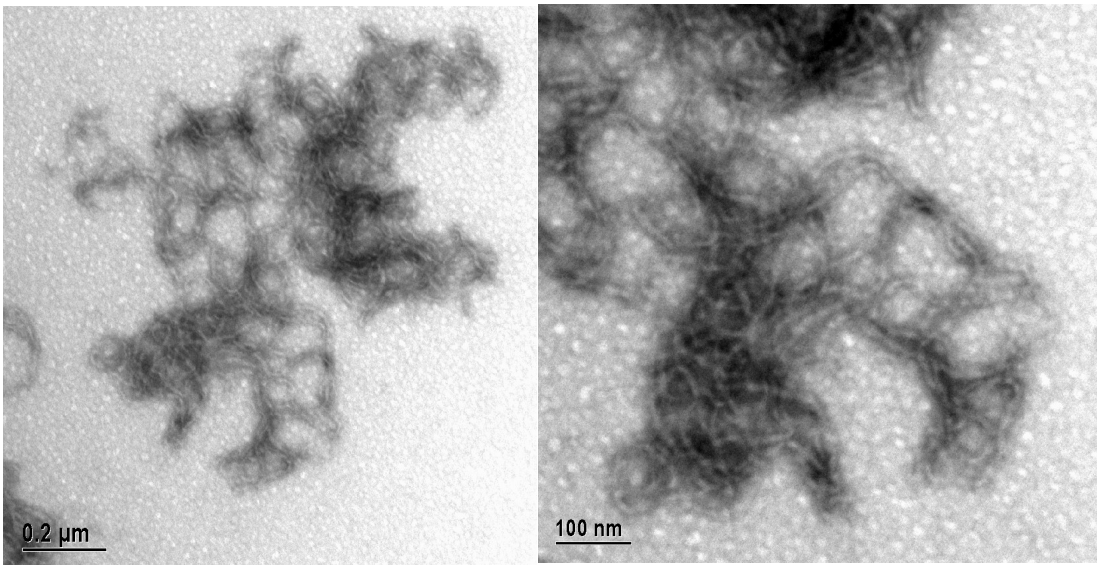


Figure 51: TEM images show the fibrils of PLL at pH 11.6 preheated at 70 °C for 15 min.

Knowing the basic assignments of the amide's components we are able to follow the structural changes during the PLL amyloid fibril formation. To simplify the spectra we calculated the difference spectra for every 5 °C during fibrillation. In the temperature range between 10 °C and 40 °C difference spectra showed manly the decrease of the intensity of the spectral features characteristic for the α-helix structure observed in amide I (1643 cm^{-1}) and amide II (1550 cm^{-1}) regions. However, at temperatures above 45 °C the first bands characteristic for antiparallel β-sheet structure of fibrils appeared. In the amide I region bands at 1692 cm^{-1} and 1616 cm^{-1} gain in intensity. Similar increase of the intensity can be observed for the band at 1529 cm^{-1} in amide II region, assigned to β-sheets. Besides the appearance of characteristic bands of β-sheet, difference spectrum 45–40 °C in Figure 52 also shows the negative band at 1652 cm^{-1} characteristic for P_{II} conformation and negative band at 1643 cm^{-1} characteristic for α-helix. The negative bands indicate loosing of these two conformations during fibrillation and suggesting the role of these two conformations in formation of β-sheets. From the Figure 42 it is evident that α-helical structure of PLL in water (pH 11.6 and low temperature) do not include any P_{II} conformation before fibrillation process takes place implying that P_{II} forms after melting of α-helix. Positive band at 1652 cm^{-1} would be expected for the P_{II} conformation in the difference spectrum (Figure 52). In contrary, negative band is observed, indicating a decrease in the P_{II} conformation already at the beginning of fibrillation. This implies that P_{II} conformation, after its formation from the melted α-helix, immediately transforms further into β-sheet.

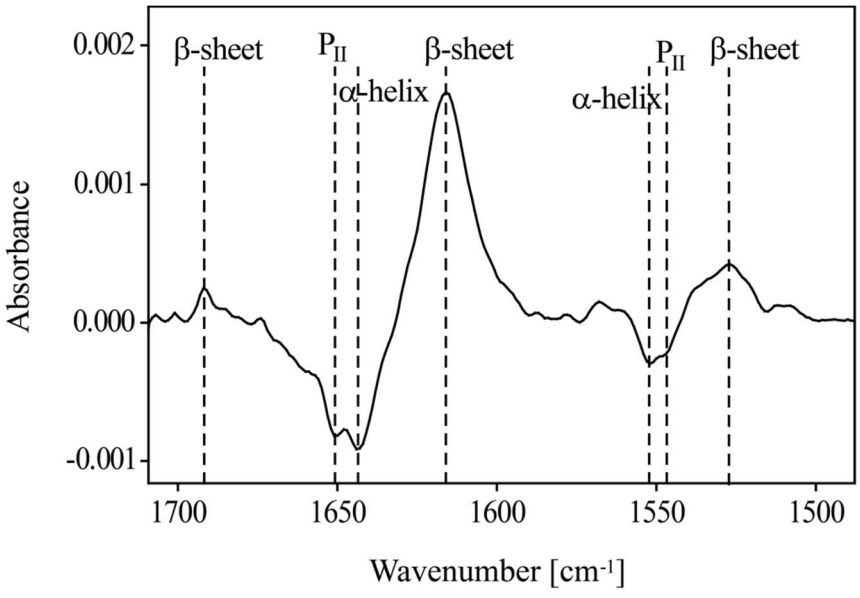


Figure 52: The difference spectrum of spectra recorded at 45 °C and 40 °C of PLL in water at pH 11.6.

However, above 65 °C only rearrangements within antiparallel β-sheets occur. The band at 1616 cm^{-1} shifts to 1620 cm^{-1} suggesting weakening of H-bonds within β-sheets. The changes of the intensity of the bands at 1616 cm^{-1} and 1620 cm^{-1} as a function of temperature (Figure 53) clearly demonstrate that the most of β-sheet fibrils are formed in the temperature range between 40 °C and 50 °C.

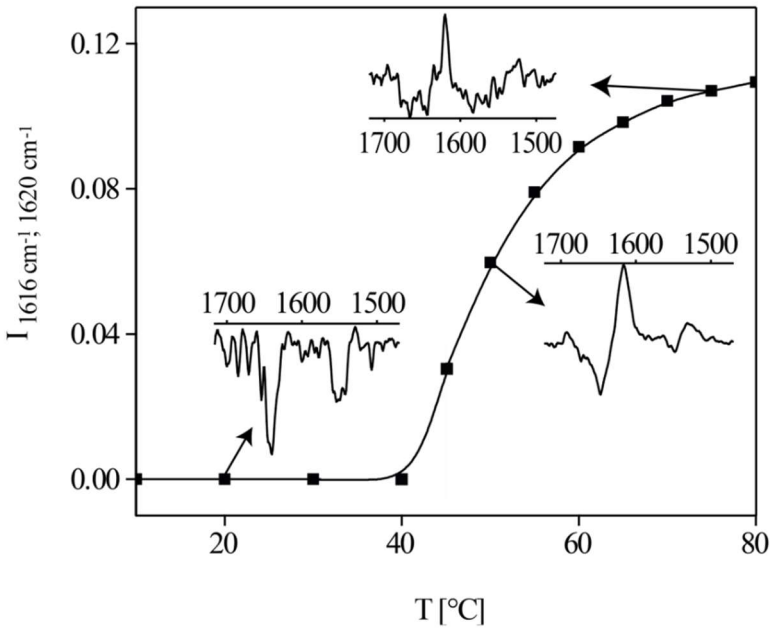


Figure 53: Plot represents the proportions of difference bands at 1616 cm^{-1} and 1620 cm^{-1} during temperature measurements from 10 °C to 80 °C for PLL in water at pH 11.6.

5.1.2.2 Fibrillation of PLL with $NaClO_4$

The titration of PLL at pH 4 with $NaClO_4$ reveals the conformational changes according to the level of concentration of sodium perchlorate. Below 400 mM $NaClO_4$ spectra match the spectrum of PLL in P_{II}-helical conformation, like in water at low pH, whereas above the concentration of 400 mM $NaClO_4$ the salt induces conformational changes to α-helical PLL with similar spectra as spectrum of PLL in water at pH 11.6 and low temperature. Asher and co-workers [227] reported that $NaClO_4$ converts charged PLL to a highly α-helical conformation by forming ion pairs between ClO_4^- ions and the NH_3^+ groups of the side chains of PLL to neutralize the repulsions that destabilize α-helical conformation. They showed that stabilization of α-helix of PLL is stronger with ClO_4^- ions than with Cl^- ions of acid. However, we showed that at 10 °C and pH 4 addition of 0.83 M $NaClO_4$ induces the formation of 61 % of the α-helix in PLL with additional 10 % of the P_{II}-helix. With increasing

the pH value to 11.6 the PPL with 0.83 M $NaClO_4$ showed increased α-helix (1643 cm^{-1}, 1550 cm^{-1}) content for 10 % while P_{II}-helix and turn conformation decrease for 4 % and 6 %, respectively (Table 19). However, compared to PLL in water at pH 11.6 it is notable that addition of $NaClO_4$ at high pH increases the P_{II} conformation for 6 %, suggesting that $NaClO_4$ stabilizes P_{II} conformation.

Table 19: Population of the conformations obtained by fitting ATR-absorbance spectra in the amide I region of PLL in pure water and with added NaClO4 at pH 4 and pH 11.6, all at 10 °C.

	α-helix	*P_{II}-helix*	*turn*
PLL in water pH = 11.6	0.77	/	0.23
PLL + 0.83 M $NaClO_4$ pH = 4	0.61	0.1	0.29
PLL + 0.83 M $NaClO_4$ pH = 11.6	0.71	0.06	0.23

Heating of PLL at pH 4 with 0.83 M $NaClO_4$ has no influence on aggregation while heating of PLL at pH 11.6 with 0.83 M $NaClO_4$ behave similar as in solutions in pure water. It results in transformation from α-helical structure to β-sheet fibrils with characteristic transition temperature around 41 °C, what is roughly 10 °C lower as without $NaClO_4$. The conformational changes during fibrillation are quite similar to those observed in pure water. From 10 °C to 30 °C changes in helical structure occur, while further heating up to 80 °C promotes bands characteristic for antiparallel β-sheet with concomitant reducing the intensity of the bands assigned to vibrations of α-helices. The difference spectrum 30–20 °C (Figure 54) also shows the increase in band 1654 cm^{-1} characteristic for P_{II} conformation, suggesting formation of P_{II} from melted α-helix and stabilization of this conformation with $NaClO_4$ before it is transformed into β-sheet. Above 30 °C the bands representing antiparallel β-sheets of amyloid fibrils appear on the cost of α-helix, P_{II} and turn conformation. The majority of fibrils are formed in the temperature range from 45 °C to 55 °C characterized by appearance of the bands at 1618 cm^{-1}, 1693 cm^{-1} and 1535 cm^{-1}. These bands are blue shifted compared to β-sheet bands in water, suggesting fibrillation of PLL with $NaClO_4$ induces weaker H-bond network within the β-sheets of fibrils. From 55 °C to 80 °C the β-sheet band at 1619 cm^{-1}

upshifts to 1623 cm^{-1} while band at 1693 cm^{-1} upshifts for 2 cm^{-1}.

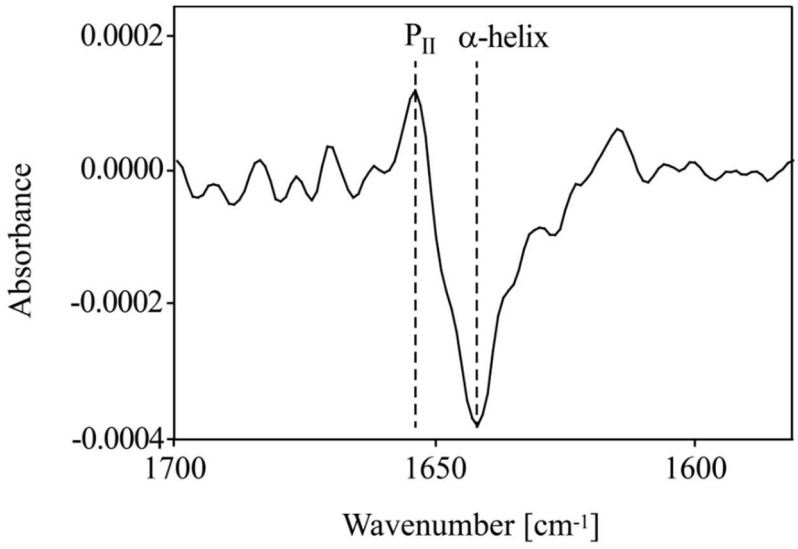


Figure 54: The difference spectrum 30-20 °C of PLL with 0.83 M $NaClO_4$ at pH 11.6.

5.1.2.3 Fibrillation of PLL with Phospholipids

Vesicles composed of dipalmitoyl phosphatidic acid (DPPA) and dipalmitoyl phosphatidyl choline (DPPC) in equivalent ratio are added to PLL at pH 11.6 leading to stabilization of helical PLL and results in increasing of the transition temperature from α-helix to β-sheet conformation of PLL for 21 °C ($T_{\alpha\rightarrow\beta}$ = 72 °C). At low temperatures PLL with DPPA+DPPC vesicles shows predominantly α-helical conformation with some P_{II}-helix and the conformation characterized by the spectral band at 1280 cm^{-1}, assigned as β conformation of those amino acid residues that are not included into secondary structure elements [45]. The band pattern of the amide I region in the difference spectra (Figure 55) shows that heating up to 40°C manly results in the growth of α-helix (band 1642 cm^{-1}) and P_{II} (band 1650 cm^{-1}) conformation. At 50 °C the α-helix starts to melt while the melting of P_{II}-helix starts at 70 °C. Additionally, the first bands characteristic for antiparallel β-sheet appear at 60 °C. A high transition temperature implies that vesicles stabilize both P_{II} and α-helical conformers. The reason for such evident elevation of the transition temperature in the presence of DPPA+DPPC vesicles may lay in interaction between the helical PLL molecules and vesicles. Generally, the lipid ester C=O group is rather exposed to solvent and thus influenced by hydrogen bonding from water hydrogens or other proton donor groups [228]. The formation of H-bonds significantly red-shifted the C=O stretching frequency. Such a red-shift (6 cm^{-1}) was indeed observed by adding the PLL to the vesicles, suggesting that PLL molecules in

helical conformation interact with DPPA+DPPC vesicles. Therefore, we propose that this type of interaction stabilizes helical conformation while heating the system above 50 °C (that is close to temperature of melting of DPPA+DPPC vesicles in water; T_m = 48 °C) results in the melting of α-helix (dotted line in Figure 55). However, interaction and consequently stabilization of helical structures of PLL with DPPA+DPPC vesicles has also some consequences on the properties of the vesicles: the CH_2 stretching frequencies of the phospholipid's aliphatic chains are shifted to the higher wavenumbers for 15 cm^{-1} in both gel and the liquid–crystalline phases of the vesicles indicating that binding of helical PLL probably reduces the order of aliphatic chains in the hydrophobic part of the vesicles. Interaction of phospholipids with PLL also decreases the T_m of vesicles for 20 °C (from 68 to 48 °C). Therefore, such significant spectral changes indicate the integration of helical PLL into vesicles that consequently results in the stabilization of the α-helices and P_{II}-helices of PLL. After the vesicles are melted (above 48 °C), peptide no longer interact with phospholipids and starts to transform into β-sheet fibrils.

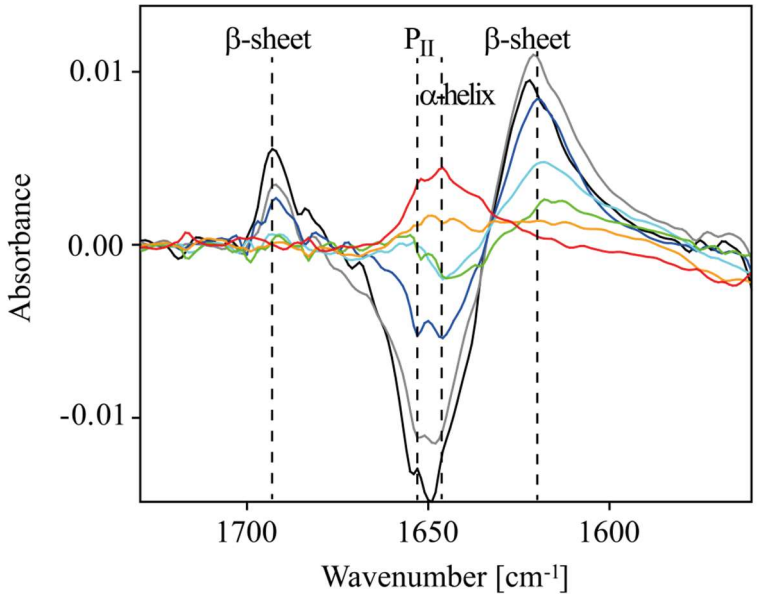


Figure 55: The difference spectra of PLL with DPPA+DPPC vesicles, pH 11.6. Each line represents the difference spectrum of 10 °C, started with 20-30 °C (red curve), 40-30 °C (orange curve), 50-40 °C (green curve), 60-50 °C (blue curve), 70-60 °C (dark blue curve), 80-70 °C (grey curve) and 90-80 ° C (black curve).

The formation of particular conformations during the heating process can be easily monitored by analysing the amide III region as presented in Figure 56. Four different conformational

changes can be followed during fibrillation in amide III region: the melting of α-helix (band at 1296 cm^{-1}), decreasing of population of amino acid residues in β conformation that are not included in secondary structure (band 1280 cm^{-1}) and formation of β-strand and β-sheet (bands at 1239 cm^{-1} and 1222 cm^{-1}). The P_{II} conformation was excluded from analysis due to the overlapping with CH_2 and CH_3 deformation bands of lipids. However, all conformations significantly change their population above 70 °C resulting in formation of majority of β-sheets on the cost of α-helix and some amino acids in β conformation. The β-strands formation starts before the formation of β-sheet, suggesting the role of β-strands structure as prerequisite for β-sheet formation and may serve as a building block for fibril growth [54].

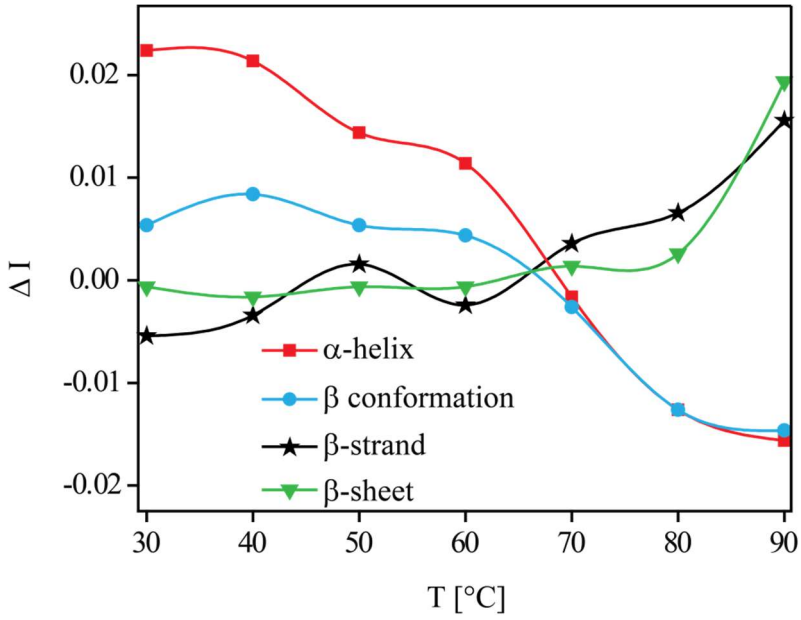


Figure 56: Plot represents the difference integrals of spectral bands in amide III region at different temperatures of PLL with DPPA+DPPC vesicles, pH 11.6, for different bands in amide III region.

5.2 Discussion about PLL

5.2.1 Spectral Assignment of PLL Conformations

We used IR spectroscopy to characterize the spectral bands of individual conformational states of PLL, stabilized by different solvents. PLL changes its conformation depending on the solvent, pH value and temperature. In water at a low pH value, it is populated predominately by the P_{II} conformation that is consistent with the literature [37, 93]. The increase of the pH value stabilizes the α-helical structure, which transforms into the

aggregated β-sheet structure after heating the peptide. The conformational analysis of PLL in various solvents showed that the decomposition of the amide III region provides the most detailed distribution of the peptide conformations. Less extensive overlapping of the characteristic bands that belong to a particular conformation improve the accuracy of the structural determination. This is especially true for two typical helical secondary structure elements, i.e. P_{II}- and α-helix. Both conformations have well separated bands in the amide III region, the frequencies of which are less sensitive to the type of solvent compared to the corresponding amide I bands. The assigned spectral bands are represented in Table 20. The assignment of the structural sensitive bands in amide I, II and III regions is based on previous results on blocked dipeptides [45] and vibrational spectra of proteins in solution with known structure [212, 217, 218, 229, 230]. This ability to discriminate between conformations may prove to be useful for early detection of amyloidal fibril formation in solutions of proteins.

Table 20: The assignment of the model bands from the amide I, II and III region, retrieved with the band fitting algorithm from the corresponding spectral regions of PLL in water.

Conformation	*Frequency [cm^{-1}]*		
	Amide I	Amide II	Amide III
turn	1664-1674	1563-1572	1257-1260
P_{II}	1648-1654	1543-1546	1308-1311
α	1643-1645	1550-1552	1290-1295
β conformation	/	/	1270-1280[a]
β-strand	1625-1630[b]	/	1240-1243
β-sheet	1630-1640[c]	/	1230
aggregated β-sheets	1692, 1618	1530	1219-1222

[a] Obtained from Grdadolnik et al. [45].

[b] Obtained from Huang et al. [229].

[c] Obtained from Baumruk et al. [230].

To confirm the presented assignment and to elucidate the origin of the bands of individual conformation, we extended the conformational studies to cover various solvents that enable to promote particular types of conformations. We used common α-helix promoting solvent TFE to stabilize the α-helix conformation of PLL [231, 232]. The analysis of the amide I, II and III bands revealed that TFE induced the α-helix conformation of PLL with 84 % of population.

Even more uniform α-helical conformation was observed when PLL was dissolved in DMSO. Contrary to the data that DMSO disrupts or weakens the intramolecular hydrogen bonds [233], our results indicate that DMSO strongly stabilizes the α-helix conformation, reaching 97 % of the population in PLL. It was shown that polar amino acid side chains have dipole-dipole interactions with the oxygen atom of DMSO and form hydrogen bonds, whereas apolar side chain alkyl groups become solvated by DMSO through the formation of a hydrophobic pocket [234]. The dual solvation properties of DMSO cause it to be a good membrane-mimicking solvent [234] that can interact with the polar ε-amino group and apolar alkyl groups of the side chain of PLL. Additionally, a uniform conformation for PLL (88 %) was found in ethylene glycol with the characteristic amide bands for P_{II}-helix.

5.2.2 The Mechanism of TFE Stabilization

The mechanism by which TFE stabilizes α-helices is unresolved, although two possible mechanisms have been discussed: helix stabilization by direct binding of TFE and an indirect mechanism in which the peptide H-bonds in the α-helix are stabilized by weakening the H-bonding of H_2O to peptide C=O and NH groups in the coil form (P_{II}-helix) [235]. However, in the literature the data show induction of α-helix at low concentration of TFE, where helix propensity for a number of model peptides increases from 0 to 25 % of TFE but levels off at higher TFE concentrations [235, 236]. Based on DFT calculations of pKa, the side chains of PLL in TFE are deprotonated, resulting in the reduced effect of decreased bulk dielectric constant of the medium [237]. Moreover, PLL exists in an α-helical conformation only in high concentrations of TFE (>80 % v:v) suggesting that the attraction forces among the backbone NH and C=O groups in α-helix overcome the strengthened repulsion forces among the charged side chain amino groups in P_{II}-helix only after sufficient high concentration of TFE is present that can change the protonation state of the side chains. Thus, above 80 % of TFE the transformation of the PLL from a P_{II}-helical to an α-helical structure occurs. We did not observe any destabilization of P_{II}-helix till 80 % of TFE, suggesting that TFE do not weaken the water-peptide H-bonds as it was proposed as a mechanism of TFE stabilization of α-helix [235]. Similar as already observed in short alanine peptides (Page 62), TFE rather stabilizes the intramolecular H-bonds, i. e. H-bond of C_7 conformation in short peptides and NH_i—CO_{i+4} H-bond in the α-helices in long peptide PLL.

5.2.3 Melting of α-helix of PLL

The α-helix melting process has been extensively studied for over 50 years. The classical view, which describes α-helix melting, was proposed by Zimm and Bragg [238], Lifson and Roig [239] and Polland and Scheraga [240]. This theory envisions the α-helix melting as a transition between an ordered α-helix structure and a disordered random coil structure with uncorrelated, but allowed ψ and ϕ Ramachandran angles. Such a melting involves the breakage of intra-α-helical H-bonds. First challenge comes from the recent observations that simple α-helix-forming peptides often melt to a P_{II} conformation [47, 50, 52, 202, 203, 221]. In fact, there is no evidence in favour of melting to, or even the existence of random coil peptides in solution. Even single alanine peptide bonds in aqueous solutions, which are intrinsically expected to be disordered, were recently reported to show P_{II}-like spectra [45]. Thus, the fact that α-helix melting (formation) is rather a transition between two ordered conformations than that between ordered and completely disordered ones, must be incorporated into the new α-helix melting theories. We showed the spectroscopic evidence that the α-helix of PLL melts into the P_{II}-helix where intramolecular H-bonds are replaced by the intermolecular H-bonds with the solvent.

Heating the PLL in 80 % TFE from 25 to 70 °C results in linear decreasing of α-helix conformation for 8 % and formation of some C=O groups in P_{II} conformation that interact with the solvent with characteristic band at 1659 cm^{-1} as observed from Figure 17 (The frequencies of band extremes in a difference spectrum may not always correspond to the real band maximum frequencies in the original spectrum and thus may be shifted due to the proximity of the bands [178] like in the case of P_{II}- and α-helix in amide I region). These results are consistent with data published by Xiong and Asher [241] where poly-alanine peptide in 50 % TFE, which possess mainly helical conformation, melts 9.1 % of its α-helices into P_{II} after heating the peptide from 25 to 40 °C.

5.2.4 Stabilization of P_{II} Conformation

Recent study revealed that lysine in short dipeptide possesses a high propensity to P_{II} conformation with the 55 % of population [45]. This high propensity of lysine to adopt the P_{II} conformation was explained by the strong backbone electrostatic interactions and by the screening of those interactions with water molecules, which is further influenced by the side chain [45, 51]. The P_{II} conformation is supposed to be stabilized by water hydrogen bonding

to the backbone amide and carbonyl group [242]. However, Drozdov et al. found no evidence for a role of water bridges in stabilizing P_{II} [48]. Contrary, Nerenberg and Head-Gordon found that the P_{II} conformation is optimizing the packing of water molecules in the hydration shell of the peptide [243]. The water molecules reduce the attractive electrostatic interactions between the peptide atoms, leaving the steric interactions as the main force for determining conformational preferences [46]. P_{II} minimizes these interactions and thus becomes the most prominent conformation. However, in this study the P_{II} conformation was determined in a non-aqueous solvent for the first time. Moreover, Law and Daggett [204] found no correlation between the dielectric constant of the solvent and the P_{II} structure. Thus we suggest that the stabilization of the P_{II} structure in ethylene glycol occurs due to the deprotonated side chain of the lysine leading to a more sterically optimized structure of the P_{II}-helix. The question of "Why there is no formation of the P_{II}-helix in TFE and DMSO, where PLL also have a deprotonated ε-amine" definitely needs an answer. Ethylene glycol has, like water molecule, both proton donor and proton acceptor groups that are able to form intermolecular hydrogen bonds with the NH and C=O group of the peptide backbone and thus hinder the formation of the intramolecular hydrogen bonds characteristic of α-helices. We suppose that P_{II}-helix stabilization has similar roots as predicted for water, i.e. steric interactions of the uncharged side chains of PLL and the establishment of hydrogen bonds between the C=O and NH groups from the peptide backbone and the solvent's proton donor and acceptor groups.

5.2.5 Intermediate Structure in PLL Fibrillation

We have introduced the difference spectra calculations to follow amyloid fibril formation of PLL in water, with added salt $NaClO_4$ and DPPA+DPPC vesicles. The difference spectra and temperature measurements show that during PLL amyloid fibril formation the α-helices of PLL melt to the extended P_{II}-helix replacing the α-helix intramolecular hydrogen bonds with intermolecular hydrogen bonds between water molecules and C=O and NH group of peptide backbone (Figure 57). The P_{II}-helix further transforms into β-strand which acts as building block for amyloid fibrils. Similar model was proposed by Jiji et.al [49]. To the best of our knowledge, this work shows the first spectral evidence of the structural changes during PLL fibrillation, i. e. melting of α-helices into the P_{II}-helix and further formation of β-strands that represent the building blocks for the formation of β-sheet fibril. The evidences of the intermediate structures in the PLL fibrillation process were already proposed by other authors [49, 244].

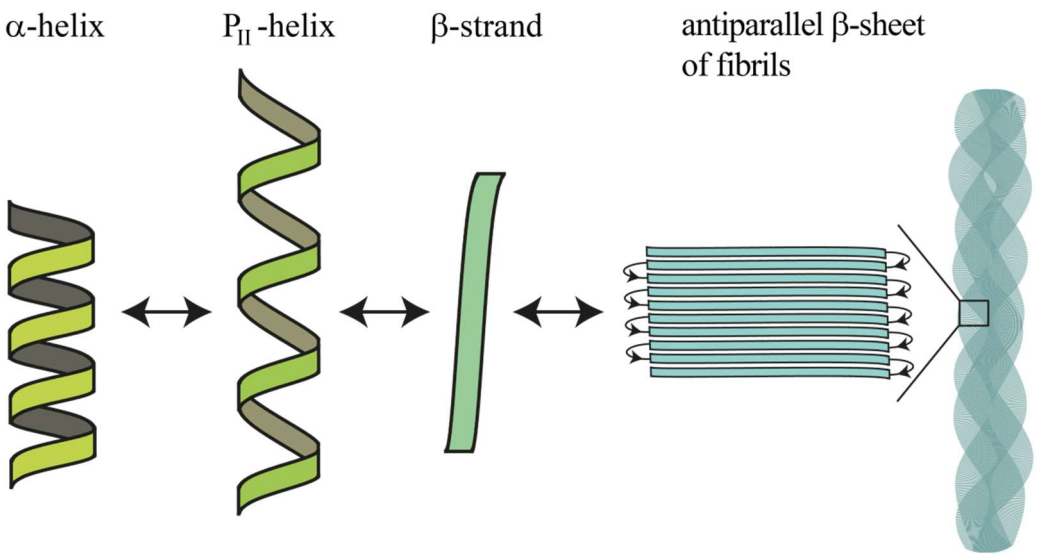


Figure 57: The model describes the amyloid fibril formation of PLL. The conformational structures α-helix, P_{II}-helix and β-strand are in equilibrium where β-strand plays as building block of the β-sheets of fibrils.

We suggest that the rate determining step during the PLL fibrillation is the formation of P_{II}-helix regarding the significant lower transition temperature after P_{II}-stabilizing salt $NaClO_4$ is added. In addition, the charged vesicles composed of equimolar ratio of DPPC and DPPA interact and thus stabilize α-helix and P_{II}-helix of PLL, preventing further transformation of helical PLL to β-sheet amyloid fibrils. P_{II}-helix of PLL represents intermediate structure that determines the kinetic of PLL fibrillation, while its interaction with the DPPA+DPPC vesicles affects the membrane integrity.

6 Fibrillation of Insulin

6.1 Results of Insulin

6.1.1 Structure of Native Insulin

Fitted model bands in infrared spectrum of the insulin in acid solutions at pH 1.6 were dominated by the band at 1642 cm^{-1} (Figure 58), indicating a large content of α-helix. The other conformations found in the insulin are β-sheet (1616 cm^{-1}), turns (1665 cm^{-1}) and P_{II} conformation (1654 cm^{-1}). The assignment of the structural sensitive bands in amide I, II and III regions is based on results obtained from the model peptide PLL. In amide III region the spectral bands of all conformations were found that that were already assigned in amide I region. The band representing the major populated α-helix has wavenumber 1291 cm^{-1}, the band that is assigned to P_{II} conformation is found at 1312 cm^{-1}, the corresponding band for turn structures is at 1266 cm^{-1} and for β-strand at 1247 cm^{-1}. The low-frequency band at 1231 cm^{-1} may represent the dimer form of insulin which is dominant at low pH [165].

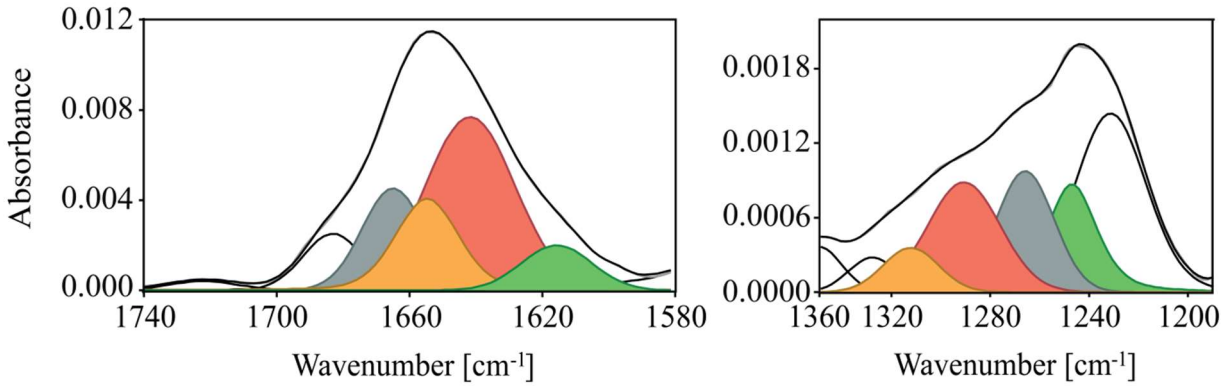


Figure 58: a) The amide I and b) amide III region of IR spectrum of insulin at pH 1.6 at 25 °C. The fitted model bands are coloured as α-helix (red), P_{II} conformation (orange), β-sheet (green) and turns (grey).

The analysis of the model bands integrals revealed that the population of the α-helix in insulin at 25 °C and pH 1.6 is about 50 % as revealed from the amide I region. The model bands for turns and P_{II} conformation have the same intensity and each represents 20 % of conformational population. The band that corresponds to β-sheet has the lowest population (10 %). Based on characteristic dihedral angles of individual conformation (Table 1) the analysis of the dihedral angles of the crystal structure of insulin, represented by PDB code 2CEU [150] (Figure 7), with the Swiss PDB Viewer [245] the following population of conformations were obtained: 62 % of α-helix, 17 % of P_{II} conformation, 10 % of turn structure and 12 % of β-sheet. These proportions are in good agreement with our results obtained from IR spectroscopy.

6.1.2 Fibrillation of Insulin

Thermally induced fibrillation of insulin in HCl at 60 °C was accompanied by a progressive increase in the β-sheet bands at 1630 and 1618 cm^{-1} in the IR spectra that can be monitored as integrals of those bands during the heating, as represented in Figure 59. The kinetics of insulin fibril formation can be described as sigmoidal curves defined by an initial lag phase where no change in the band intensity characteristic for β-sheet was observed, a subsequent growth phase in which the bands for β-sheet increased, and final equilibrium phase, where intensities of those bands reached a plateau, indicating the end of fibril formation (Figure 59). The lag time was determined to last 25 min followed by exponential growth of insulin amyloid fibrils with the apparent rate constant (k_{app}) 2.7 h^{-1} that is in good agreement with the results found in the literature [156], suggesting the IR spectroscopy is a good method to observe the fibrillation kinetic.

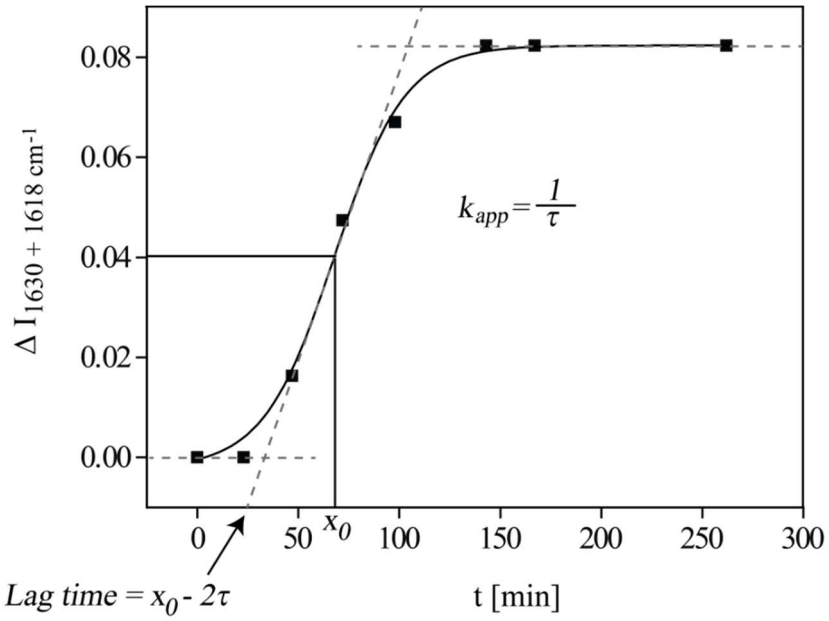


Figure 59: The sigmoidal increase in intensity of the band at 1630 and 1618 cm^{-1} representing the insulin fibril formation during the heating at pH 1.6 and 60 °C. The lag time is approximated by $x_0 - 2\tau$ and the apparent rate constant, k_{app}, for the fibril growth is given by $1/\tau$.

In the first 20 min of the heating of the insulin at 60 °C no spectral changes were observed. Further heating resulted in gradually increasing of the bands at 1630, 1618 and 1680 cm^{-1} that represent the formation of amyloid fibrils. Simultaneously, the intensities of the bands at 1642 and 1654 cm^{-1} were gradually diminished indicating the melting of α-helix and P_{II} conformation. The difference spectra showed that after the fibrillation starts first the band characteristic of P_{II} conformation decreases in intensity followed by the decrease of the band of α-helix. No further changes in amide I region were seen after 3 h of heating, where the final state of insulin fibrillation was composed of 65 % of β-sheet, 20 % of turns, 7 % of P_{II} conformation and 8 % of α-helix (Figure 60).

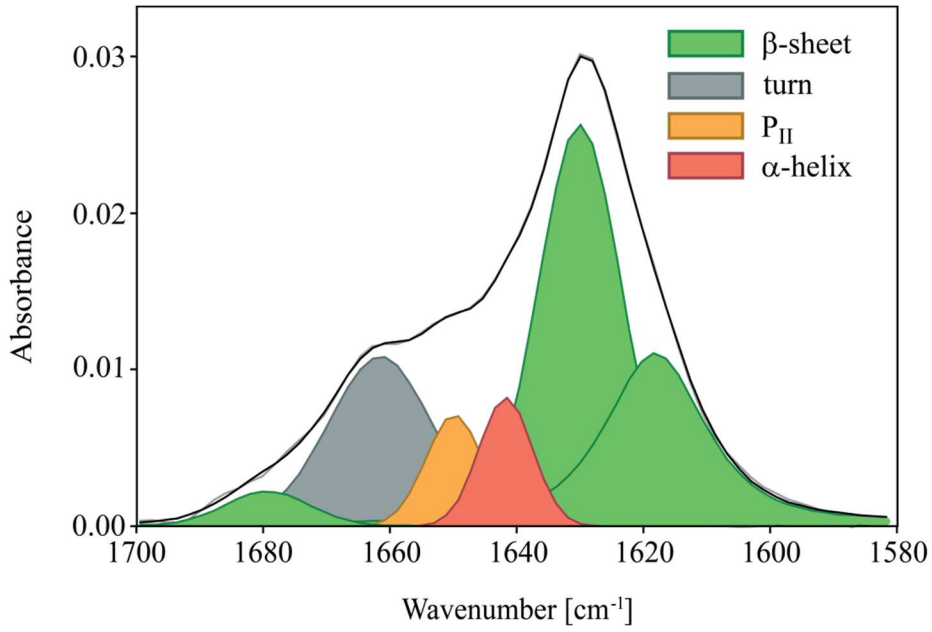


Figure 60: The amide I region of IR spectrum of insulin at pH 1.6 heated for 3h at 60 °C with fitted model bands of individual conformation.

VCD spectra of insulin at pH 1.6 were measured at 60 °C to observe the fibrillation process. From Figure 62 it is evident that VCD spectral band characteristic of β-sheet significantly gain in intensity. The enhancement of the VCD band of amyloid fibrils has not been explained in details. It is proposed that VCD intensity enhancement results from intrasheet as well as from inter-sheet vibrational coupling between stacked β-sheets [177]. From Figure 61 it is observed that the enhancement of the negative band of β-sheet is at least 3 ordered of magnitude higher compared to the band of non-fibrillar insulin, suggesting the VCD can be used as a probe to determine the kinetics of the insulin fibril formation. The VCD spectrum of the native insulin contained a positive VCD band at 1648 cm^{-1} that represents α-helix and smaller positive coupled band with maximum at 1673 cm^{-1} and minimum at 1668 cm^{-1} which is assigned to turn structure.

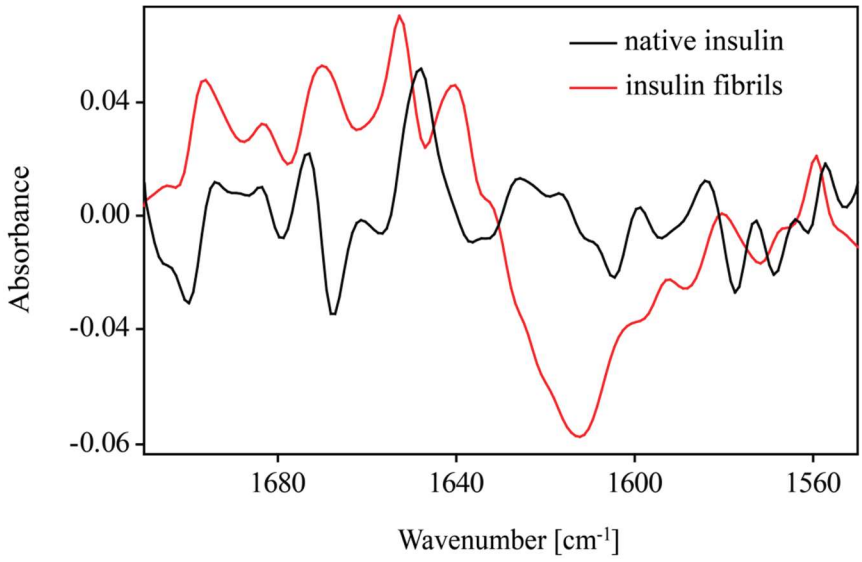


Figure 61: VCD spectra of insulin in D_2O at pH 1.6 measured at the beginning (black curve) and after 5 h of heating at 60 °C.

As revealed from Figure 62 the fibrillation starts after 42 min of heating the insulin at 60 °C. The negative VCD band characteristic of β-sheet amyloid fibril gains in intensity another 2 hours followed by steady state phase as a last phase of insulin fibril growth. Obtained curve has characteristic sigmoidal shape with the k_{app} 2.5 h^{-1}.

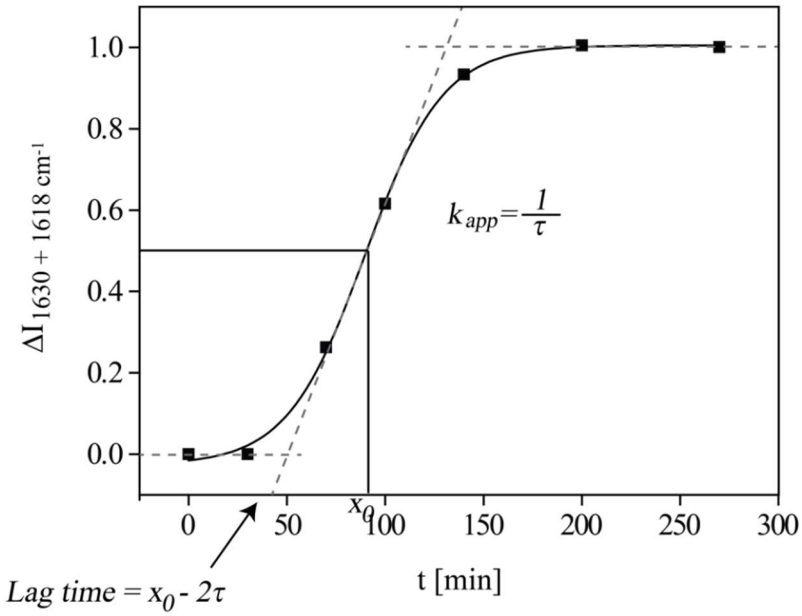


Figure 62: The sigmoidal increase in intensity of the negative band between 1630 and 1615 cm^{-1} in amide I region of VCD spectra representing the insulin fibril formation during the heating at pH 1.6 and 60 °C. The lag time is approximated by $x_0 - 2\tau$ and the apparent rate constant, k_{app}, for the fibril growth is given by $1/\tau$.

6.1.3 Effect of Insulin Concentration on Fibril Formation

The kinetics of insulin fibril formation was measured using fluorescent dye molecule ThT that specifically bind to amyloid fibrils. When insulin solutions were incubated under conditions leading to fibrils, the increase in fluorescence of ThT at 482 nm was observed. For all preformed measurements of insulin at different concentrations a sigmoidal curves were obtained defined by an initial lag phase, where no change in ThT fluorescence intensity was observed, a subsequent growth phase in which ThT fluorescence increased, and a final equilibrium phase, where ThT fluorescence reached a plateau indicating the end of fibril formation. Like in the case of vibrational kinetic measurements of insulin fibrillation, the lag time and k_{app} were determined from the sigmoidal curves. The influence of insulin concentration on the kinetics of fibril formation at 60 °C is reviewed in Table 21 with the lag times and apparent rate constant for fibril growth. There was an inverse linear correlation between the logarithm of insulin concentration and the duration of the lag time (Figure 63, a). The rate of elongation of fibrils is consistent with a first-order reaction, based on correlation coefficient (R^2 = 0.97). There was also linear dependency of the apparent rate constant for fibril growth on insulin concentration (Figure 63, b).

Table 21: Effect of concentration on lag time and growth rate constant for insulin fibrillation. Fibril formation was monitored by ThT fluorescence at 60 °C at pH 1.6.

Concentration [mg/mL]	Lag time [h]	k_{app} [h^{-1}]
1	3.1	1.5
5	2.1	2.5
20	0.8	3.1

During incubation at 60 °C, increasing viscosity of the sample was observed with time, indicating that insulin fibrils were formed. The end state of the fibrillated insulin solutions can be described as firm, turbid gel. Additionally, very good linear correlation was obtained when comparing the initial insulin concentration and the values of ThT fluorescence of fibrillated sample.

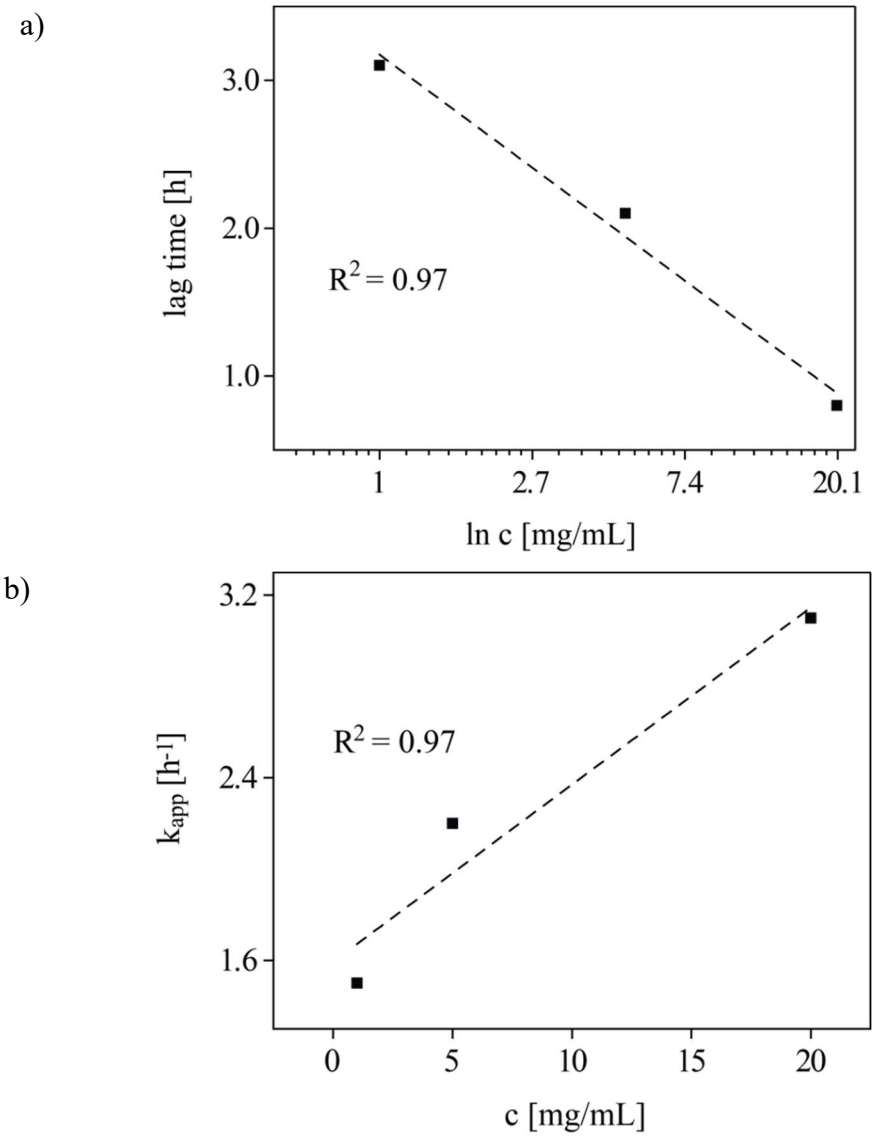


Figure 63: Effects of insulin concentration on the kinetics of fibrillation. a) Inverse linear dependency of the logarithm to insulin concentration as a function of lag time and b) linear dependency of the apparent rate constant, k_{app}, for fibril growth on the insulin concentration at pH 1.6 and 60 °C.

6.2 Discussion about Insulin

The present FTIR results for native insulin agree well with the earlier X-ray, NMR and FTIR spectroscopic studies. The major component of the amide I spectrum is the α-helix band, whereas the β-sheet, P_{II} conformation and turn bands contribute less to the secondary structure.

Insulin fibrils prepared by heating the sample at 60 °C showed an increased amount of β-sheet, whereas the α-helix and P_{II} conformation bands decreased significantly compared with the native state, indicating major conformational changes during fibrillation. The increase of

the bands in the β-sheet region with concomitant loss of the α-helical structure, induced by fibril formation, agrees well with reported studies of insulin [156, 219, 246]. In the amide I region of insulin fibrils two low-frequency bands were found at 1630 and 1618 cm^{-1} whereas in fibrils of PLL only the band at 1618 cm^{-1} represents the antiparallel β-sheets. The different β-sheet bands have been suggested to reflect subtle differences in hydrogen bonding between β-strands [247], precisely, it has been shown that the bands at 1620–1630 cm^{-1} arise due to intermolecular β-sheet contacts [248, 249]. It was suggested that single protofibril seen in electron micrographs of insulin fibrils correlate with the formation of the lower frequency β-sheet band at 1618 cm^{-1} in the FTIR spectrum, whereas thicker laterally aggregated fibrils showed β-sheet bands at higher frequencies around 1635 and 1629 cm^{-1} [250]. Thus, the H-bonds in the perpendicular β-sheets apparently have weaker interactions compared to the strong H-bonding in protofibrils. We suggest that the new band at 1630 cm^{-1} in the IR spectra of insulin fibrils in Figure 60 is due to laterally intermolecular contacts of β-sheet in aggregated fibrils. The difference in the ratio and the position of the bands at 1630 and 1618 cm^{-1} provides insight into the nature of the intermolecular contacts in the insulin fibrils and consequently into the morphology of the fibrils, suggesting that insulin aggregation revealed more lateral growth of the fibrils. However, more studies are needed to further elucidate the role of intermolecular H-bonding on the fibril structure and morphology.

The characteristic sigmoidal curves of measured intensities of β-sheet IR bands and the values of ThT fluorescence, seen when insulin fibrils are formed, was observed in most other amyloid systems. The kinetics of insulin fibril formation was characterized by two independent factors that determined the overall rate of fibril formation: lag time (related to how fast the nuclei are formed) and fibril growth (the rate of fibril elongation). The rate of elongation of fibrils is consistent with a first-order reaction, based on the good fit of the data to the sigmoidal curve. The kinetics of elongation have been shown to follow first-order kinetics for other proteins, including α-synuclein and amyloid β [251, 252]. The simplest explanation for the linear dependency of the apparent first-order rate constant for fibril growth on insulin concentration is that increasing the concentration of insulin leads to increasing numbers of fibrils, due to the increasing concentration of nuclei [156]. The exponential growth of fibrils stops after 100 min of heating following the steady state phase where the growth and end-to-end annealing of the fibrils are in the equilibrium [136]. This curve is consistent with a nucleation-dependent elongation model, in which the three phases (lag, exponential increase and final levelling off) correspond to nucleation, extension and

equilibrium phases [126, 253]. In order to understand the fibrillation mechanism, we determine the kinetic parameters of insulin fibril formation. The lag phase showed high concentrational dependence as is characteristic of nucleation-dependent fibrillation model [133, 252, 254], but more studies should be done to exclude other mechanisms of amyloid fibril formation.

We showed the different spectroscopic techniques give similar kinetic description of insulin fibril formation. Dzwolak et al. [255] also used both IR spectroscopy and ThT fluorescence to determine the formation of insulin fibrils and showed that both techniques gave similar sigmoidal curve. IR spectroscopy is very convenient method to follow fibril formation because it provides the kinetic information as well as the information about the structural changes that occurs during fibrillation.

7 Discussion

Nearly five decades ago, Ramachandran and co-workers [99] elucidated the steric map for alanine dipeptide. Today, the restrictions that sterics impose on the conformation of dipeptide are widely accepted. However, the hard sphere models cannot be used in protein structure prediction or in analysis of unfolded state. The reason lies in the intrinsic backbone preferences of each amino acid [51, 56] as well as in the fact that each ϕ,ψ-pair is not independent of its neighbours as it was proposed by Flory [28]. Earlier work showed that alanine dipeptide in water preferentially populates P_{II} conformation [45]. To determine the neighbour alanine residue effect we studied the conformational distribution of Ala_2, Ala_3 and Ala_4. We showed that with the increasing chain length of the short alanine peptides in water the P_{II} conformation decreases in population while the conformation that was assigned to C_7 conformation increases in population. The stabilization of Ala comes from the favourable solvation conditions while in Ala_2, Ala_3 and Ala_4 the favourable electrostatic interactions between backbone groups are suggested to have the dominant contribution to the stability of short alanine peptides. Based on the study of the coil library, Baldwin and Avbelj calculated the electrostatic free energy for the P_{II} and β structures. They interpret the neighbour effect as due to shielding from water by the $i \pm 1$ side chain so that branched neighbours favour β structure over P_{II} [110]. Chen et al. found that the Ala* residue in the peptide $AcGGAA^*AGGNH_2$ (* represents ^{15}N label) is in P_{II} conformation while its conformation is not affected by flanking alanines [52]. In contrary to these results, we found a significant neighbour effect of alanine residue on the conformational population of neighbour alanine residue. To assign spectral features in terms of conformational population of short alanine peptides we used several spectroscopic techniques, such as IR, Raman, VCD, NMR and UV-CD, as well as computational method DFT. The stabilization of individual conformations was achieved with the change of the solvent. However, when studying dynamic systems, such as short alanine peptides, much precaution should be taken when interpreting the results. Similar approach as in the short alanine peptides was used in the study of PLL. The spectral

assignments of PLL conformations were obtained from the decomposition of amide I, II and especially amide III region where the specific conformation was characterized by dissolving the PLL in the solvent that promote particular type of conformation. We showed that P_{II} conformation is not connected only to water as a solvent and the hydration of peptide bonds as previously assumed [47, 48]. Ethylene glycol even more efficiently stabilizes the P_{II}-helix. These results illustrates the unique physical properties of P_{II} conformation that was observed as an intermediate structure in amyloid fibril formation [49, 50] and is recognized as a general structure of natively unfolded proteins [41, 44].

Understanding the rules governing protein folding should lead to a better understanding of protein misfolding. The conversion to the cross-β structure observed in mature amyloid fibrils takes place starting from an intermediate conformation which, in the case of globular proteins, forms after partial unfolding and, in natively unfolded proteins, after partial folding [126, 256]. In general, amyloid formation *in vitro* can be achieved by destabilizing the native state of the protein under conditions in which non-covalent interactions still remain favourable [126]. In the case of globular proteins, the formation of partially unfolded intermediates populated from the native state or accessible during refolding is the first critical step of the pathway of amyloid fibril formation [126, 128]. Our results show that fibrillation of α-helical peptides PLL and insulin proceeds through the formation of P_{II}-helix that represents the rate determining step in the formation of amyloid fibrils. P_{II} conformation is very exposed to the solvent and therefore being well suited for participating in protein-protein interactions [257]. We showed that P_{II}-helix is stabilized by acid lipid vesicles and this interaction stabilize its structure and consequently slow down the fibrillation process. In the study of the PLL interacting with phospholipid bilayer Spurlin and Gewirth [258] observed the formation of spherical domains that are inserted into the bilayers. They proposed the mechanism of PLL insertion that is similar to those of membrane pores [258] that allows the molecule transport. However, we showed that PLL in the helical conformation interacts and probably gets inserted into the phospholipid vesicles but further studies needs to be performed to show the formation of pores in those vesicles. In order to understand the general mechanism of amyloid fibril formation the non-specific sequence of PLL can demonstrate a generic property of polypeptide main chain interaction and suggests that amyloid formation arises primarily from the main chain interaction that are, in some environments, overruled by specific side chain contacts [259]. The P_{II} conformation was observed as intermediate structure in the process of fibrillation of α+β structured lysozyme [50]. Moreover, P_{II}-helix was proposed as a most

populated structure in natively unfolded Aβ peptide [144, 260], tau and α-synuclein [55]. However, we cannot postulate that P_{II} conformation is a general intermediate structure in the fibrillation pathway, as we did not study any α+β peptide and all-β peptide nor any globular protein that forms amyloid fibrils.

8 Conclusions

This thesis aimed at improving our understanding of the conformational aspects of unfolded peptides and proteins and how they relate to intrinsic propensities of amino acids in a polypeptide chain. The values reported herein reflect a valuable intrinsic structural propensity for alanine residue within a limited neighbouring sequence context. Furthermore, this work emphasizes the necessity to have an indispensable intrinsic propensity scale for dialanine, trialanine and tetraalanine as a basis for comparison of further context dependent studies and the calibration of molecular mechanics force fields for computational biochemistry.

We have shown how lower resolution techniques in the view of structural details such as IR, VCD, Raman and UV-CD spectroscopy can be used to quantitatively probe the conformational changes as well as the H-bonding connectivity. We employed the combined use of different spectroscopic methods as a tool to explore and differentiate between the conformations and thus assign the characteristic spectral bands for each conformation. We also categorized the role of certain solution conditions in their connection to structural effects and thus, alterations of spectroscopic data, identified and characterized conformational population in short alanine peptides. We have re-examined in detail the assignment of the amide III region of the P_{II}, β, C_7 and α_R conformations using IR and Raman spectra of a short alanine peptides and determine the characteristic VCD band shape for each conformation. Furthermore, concentrational measurements of Ala in chloroform showed the oligomerization of Ala in β conformation. DFT calculations indicate that β dimer is the most favourable and suggest the β oligomerization as an onset of the early nucleation structure in the process of self-aggregation.

Our study indicates that the unfolded states of PLL exist as a mixture of P_{II} conformation and the extended β structures. The P_{II} conformation was previously assumed to be the conformation that is most populated in unfolded peptides and proteins and that is mainly stabilized by water molecules. We have indeed shown that P_{II} conformation was the most favourable conformation of alanine dipeptide in water where the stabilization of peptide backbone proceeded through the water bridges connecting the C=O and NH groups. However, PLL exhibits even stronger propensity to adopt P_{II} conformation when solvent was changed to

the ethylene glycol. Therefore, we showed that P_{II} conformation is not connected only with water as a solvent and the hydration of peptide bonds, rather it is stabilized with the intermolecular H-bonds between the C=O and NH groups from the peptide backbone and the solvent's proton donor and acceptor groups.

We used the model peptide PLL and human insulin to study the conformational changes during the fibrillation process. Before fibrillation takes place both, peptides mainly possess an α-helical structure. Upon heating, α-helix and P_{II} conformation started to melt and, simultaneously, anti-parallel β-sheet formed. We showed that P_{II} conformation represents the rate-determining step in amyloid fibril formation. Our results reaffirm the notion that homopolypeptide PLL not affiliated with any known debilitating human diseases 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.

Amyloid fibrils display a downshift of the amide I band of the IR spectrum as well as an enhanced signal in this region in the VCD spectrum. The increase in intensity of the difference IR and VCD amide I band can be used to probe the kinetics of fibril formation. We introduced new structural parameters retrieved from the IR spectra, which are able to distinguish between the α-helix and P_{II} conformations, as well as between the various types of β structures, such as β conformation, β-strands and β-sheets, by analysing the amide I and amide II region of IR spectra, as well as amide III. Thus, difference IR and VCD spectroscopy are a convenient and novel probe of the fibril formation kinetics and structure of peptide fibrils.

9 Acknowledgements

I would like to express my sincere gratitude to my advisor and mentor, Asst. Prof. Dr. Jože Grdadolnik for his continued support and guidance through my doctoral program, and for all discussions and advices that have moulded my thinking.

I would like to acknowledge various people who have contributed to my research and the work presented herein. I would like to thank Asst. Prof. Dr. Franci Merzel for his assistance and help with theoretical calculations. I would like to thank Asst. Prof. Dr. Simona Golič Grdadolnik for her help with NMR experiments and for all helpful suggestions. I would like to thank Dr. Franc Avbelj for the proof reading part of the thesis and for the useful comments.

I would like to thank Dr. Magda Tušek Žnidarič for preforming TEM measurements.

I would like to thank Metka Novak for her assistance and sharing her laboratory experiences with me.

I would like to thank all the colleagues from Laboratory for Biomolecular Structure for helpful discussions and friendly environment.

I would like to thank my committee members for devoting their time to read the dissertation and for their valuable suggestions.

I would also like to thank my family and friends for their love and support. Finally, I want to express my deepest love and thanks to Martin for his support, understanding and encouragement.

10 References

[1] Mirsky, A. E.; Pauling, L. On the Structure of Native, Denatured, and Coagulated Proteins. *Proceedings of the National Academy of Sciences* **22**, 439–447 (1936).

[2] Anson, M. L.; Mirsky, A. E. The Reversibility of Protein Coagulation. *The Journal of Physical Chemistry* **35**, 185–193 (1930).

[3] Anfinsen, C. B. Principles That Govern Folding of Protein Chains. *Science* **181**, 223–230 (1973).

[4] Levinthal, C. How to fold graciously. In: *Mossbauer Spectroscopy in Biological Systems, Proceedings of a meeting held at Allerton House*. 22–24 (University of Illinois Press, Monticello, 1969).

[5] Levinthal, C. Are There Pathways For Protein Folding? *Extrait du Journal de Chimie Physique* **65**, 44–45 (1968).

[6] Dill, K. A. Dominant Forces in Protein Folding. *Biochemistry* **29**, 7133–7155 (1990).

[7] Kauzmann, W. Some factors in the interpretation of protein denaturation. *Advances in Protein Chemistry* **14**, 1–63 (1959).

[8] Avbelj, F.; Moult, J. Role of electrostatic screening in determining protein main chain conformational preferences. *Biochemistry* **34**, 755–764 (1995).

[9] Brant, D. A.; Flory, P. J. The Role of Dipole Interactionsin Determining Polypeptide Conformation. *Journal of the American Chemical Society* **87**, 663–664 (1965).

[10] Perutz, M. F. Electrostatic Effects in Proteins. *Science* **201**, 1187–1191 (1978).

[11] Klapper, M. H. On the Nature of the Protein Interior. *Biochimica et Biophysica Acta* **229**, 557–566 (1971).

[12] Richards, F. M. Areas, Volumes, Packing and Protein Structure. *Annual Review of Biophysics and Bioengineering* **6**, 151–176 (1977).

[13] Schellman, J. A. The stability of hydrogen-bonded peptide structures in aqueous solution. *Comptes Rendus des Travaux du Laboratoire Carlsberg* **29**, 230–259 (1955).

[14] Stickle, D. F.; Presta, L. G.; Dill, K. A.; Rose, G. D. Hydrogen bonding in globular proteins. *Journal of Molecular Biology* **226**, 1143–1159 (1992).

[15] Scholtz, J. M. et al. Calorimetric determination of the enthalpy change for the alpha-helix to coil transition of an alanine peptide in water. *Proceedings of the National Academy of Sciences* **88**, 2854–2858 (1991).

[16] Richardson, J. M.; Lopez, M. M.; Makhatadze, G. I. Enthalpy of helix–coil transition: Missing link in rationalizing the thermodynamics of helix-forming propensities of the amino acid residues. *Proceedings of the National Academy of Sciences of the United States of America* **102**, 1413–1418 (2005).

[17] Myers, J. K.; Pace, C. N. Hydrogen bonding stabilizes globular proteins. *Biophysical Journal* **71**, 2033–2039 (1996).

[18] Baskakov, I. V. et al. Pathway complexity of prion protein assembly into amyloid. *Journal of Biological Chemistry* **277**, 21140–21148 (2002).

[19] Hammarstrom, P.; Wiseman, R. L.; Powers, E. T.; Kelly, J. W. Prevention of transthyretin amyloid disease by changing protein misfolding energetics. *Science* **299**, 713–716 (2003).

[20] Kiefhaber, T.; Rudolph, R.; Kohler, H. H.; Buchner, J. Protein Aggregation Invitro and Invivo - a Quantitative Model of the Kinetic Competition between Folding and Aggregation. *Bio-Technology* **9**, 825–829 (1991).

[21] Gazit, E. The "Correctly Folded" State of Proteins: Is It a Metastable State? *Angewandte Chemie International Edition* **41**, 257–259 (2002).

[22] Dunker, A. K. et al. Intrinsically disordered protein. *Journal of Molecular Graphics and Modelling* **19**, 26–59 (2001).

[23] Shortle, D. The denatured state (the other half of the folding equation) and its role in protein stability. *FASEB Journal* **10**, 27–34 (1996).

[24] Tanford, C. Protein denaturation. *Advances in Protein Chemistry* **23**, 121–282 (1968).

[25] Kohn, J. E. et al. Random-coil behavior and the dimensions of chemically unfolded proteins. *Proceedings of the National Academy of Sciences of the United States of America* **101**, 12491–12496 (2004).

[26] Fitzkee, N. C.; Rose, G. D. Reassessing random-coil statistics in unfolded proteins. *Proceedings of the National Academy of Sciences of the United States of America* **101**, 12497–12502 (2004).

[27] Tran, H. T.; Wang, X.; Pappu, R. V. Reconciling observations of sequence-specific conformational propensities with the generic polymeric behavior of denatured proteins. *Biochemistry* **44**, 11369–11380 (2005).

[28] Flory, P. J. *Statistical Mechanics of Chain Molecules* (Wiley, New York, 1969).

[29] Tanford, C.; Kawahara, K.; Lapanje, S. Proteins in 6-M guanidine hydrochloride. Demonstration of random coil behavior. *Journal of Biological Chemistry* **241**, 1921–1923 (1966).

[30] Plaxco, K. W. et al. Chain collapse can occur concomitantly with the rate-limiting step in protein folding. *Nature Structural Biology* **6**, 554–556 (1999).

[31] Wilkins, D. K. et al. Hydrodynamic Radii of Native and Denatured Proteins Measured by Pulse Field Gradient NMR Techniques. *Biochemistry* **38**, 16424–16431 (1999).

[32] Ohkubo, Y. Z.; Brooks, C. L. Exploring Flory's isolated-pair hypothesis: Statistical mechanics of helix–coil transitions in polyalanine and the C-peptide from RNase A. *Proceedings of the National Academy of Sciences* **100**, 13916–13921 (2003).

[33] Baldwin, R. L. Seeding protein folding. *Trends in Biochemical Sciences* **11**, 6–9 (1986).

[34] McGee, W. A. et al. Thermodynamic cycles as probes of structure in unfolded proteins. *Biochemistry* **35**, 1995–2007 (1996).

[35] Creighton, T. E. On the relevance of non-random polypeptide conformations for protein folding. *Biophysical Chemistry* **31**, 155–162 (1988).

[36] Tiffany, M. L.; Krimm, S. New chain conformations of poly(glutamic acid) and polylysine. *Biopolymers* **6**, 1379–1382 (1968).

[37] Dukor, R. K.; Keiderling, T. A. Reassessment of the random coil conformation: vibrational CD study of proline oligopeptides and related polypeptides. *Biopolymers* **31**, 1747–1761 (1991).

[38] Woutersen, S.; Hamm, P. Structure Determination of Trialanine in Water Using Polarization Sensitive Two-Dimensional Vibrational Spectroscopy. *The Journal of Physical Chemistry B* **104**, 11316–11320 (2000).

[39] Schweitzer-Stenner, R.; Eker, F.; Huang, Q.; Griebenow, K. Dihedral angles of trialanine in D2O determined by combining FTIR and polarized visible Raman spectroscopy. *Journal of the American Chemical Society* **123**, 9628–9633 (2001).

[40] Shi, Z. et al. Polyproline II structure in a sequence of seven alanine residues. *Proceedings of the National Academy of Sciences of the United States of America* **99**, 9190–9195 (2002).

[41] Shi, Z. S.; Woody, R. W.; Kallenbach, N. R. Is polyproline II a major backbone conformation in unfolded proteins? *Unfolded Proteins* **62**, 163–240 (2002).

[42] Keiderling, T. A. Protein and peptide secondary structure and conformational determination with vibrational circular dichroism. *Current Opinion in Chemical Biology* **6**, 682–688 (2002).

[43] Rucker, A. L.; Creamer, T. P. Polyproline II helical structure in protein unfolded states: lysine peptides revisited. *Protein Science* **11**, 980–985 (2002).

[44] Barron, L. D.; Blanch, E. W.; Hecht, L. Unfolded proteins studied by raman optical activity. *Unfolded Proteins* **62**, 51–90 (2002).

[45] Grdadolnik, J.; Mohacek-Grosev, V.; Baldwin, R. L.; Avbelj, F. Populations of the three major backbone conformations in 19 amino acid dipeptides. *Proceedings of the National Academy of Sciences of the United States of America* **108**, 1794–1798 (2011).

[46] Toal, S.; Amidi, O.; Schweitzer-Stenner, R. Conformational Changes of Trialanine Induced by Direct Interactions between Alanine Residues and Alcohols in Binary Mixtures of Water with Glycerol and Ethanol. *Journal of the American Chemical Society* **133**, 12728–12739 (2011).

[47] Pappu, R. V.; Rose, G. D. A simple model for polyproline II structure in unfolded states of alanine-based peptides. *Protein Science* **11**, 2437–2455 (2002).

[48] Drozdov, A. N.; Grossfield, A.; Pappu, R. V. Role of Solvent in Determining Conformational Preferences of Alanine Dipeptide in Water. *Journal of the American Chemical Society* **126**, 2574–2581 (2004).

[49] JiJi, R. D.; Balakrishnan, G.; Hu, Y.; Spiro, T. G. Intermediacy of poly(L-proline) II and beta-strand conformations in poly(L-lysine) beta-sheet formation probed by temperature-jump/UV resonance Raman spectroscopy. *Biochemistry* **45**, 34–41 (2006).

[50] Blanch, E. W. et al. Is polyproline II helix the killer conformation? A Raman optical activity study of the amyloidogenic prefibrillar intermediate of human lysozyme. *Journal of Molecular Biology* **301**, 553–563 (2000).

[51] Avbelj, F.; Grdadolnik, S. G.; Grdadolnik, J.; Baldwin, R. L. Intrinsic backbone preferences are fully present in blocked amino acids. *Proceedings of the National Academy of Sciences of the United States of America* **103**, 1272–1277 (2006).

[52] Chen, K.; Liu, Z.; Kallenbach, N. R. The polyproline II conformation in short alanine peptides is noncooperative. *Proceedings of the National Academy of Sciences of the United States of America* **101**, 15352–15357 (2004).

[53] Dyson, H. J.; Wright, P. E. Peptide conformation and protein folding. *Current Opinion in Structural Biology* **3**, 60–65 (1993).

[54] Avbelj, F.; Grdadolnik, S. G. Electrostatic screening and backbone preferences of amino acid residues in urea-denatured ubiquitin. *Protein Science* **16**, 273–284 (2007).

[55] Syme, C. D. et al. A Raman optical activity study of rheomorphism in caseins, synucleins and tau. New insight into the structure and behaviour of natively unfolded proteins. *European Journal of Biochemistry* **269**, 148–156 (2002).

[56] Avbelj, F.; Baldwin, R. L. Role of backbone solvation and electrostatics in generating preferred peptide backbone conformations: distributions of phi. *Proceedings of the National Academy of Sciences of the United States of America* **100**, 5742–5747 (2003).

[57] Serrano, L. Comparison between the phi distribution of the amino acids in the protein database and NMR data indicates that amino acids have various phi propensities in the random coil conformation. *Journal of Molecular Biology* **254**, 322–333 (1995).

[58] Swindells, M. B.; MacArthur, M. W.; Thornton, J. M. Intrinsic phi, psi propensities of amino acids, derived from the coil regions of known structures. *Nature Structural Biology* **2**, 596–603 (1995).

[59] Whittington, S. J.; Chellgren, B. W.; Hermann, V. M.; Creamer, T. P. Urea promotes polyproline II helix formation: implications for protein denatured states. *Biochemistry* **44**, 6269–6275 (2005).

[60] McColl, I. H.; Blanch, E. W.; Hecht, L.; Barron, L. D. A Study of α-Helix Hydration in Polypeptides, Proteins, and Viruses Using Vibrational Raman Optical Activity. *Journal of the American Chemical Society* **126**, 8181–8188 (2004).

[61] Eker, F.; Griebenow, K.; Schweitzer-Stenner, R. Stable conformations of tripeptides in aqueous solution studied by UV circular dichroism spectroscopy. *Journal of the American Chemical Society* **125**, 8178–8185 (2003).

[62] Schweitzer-Stenner, R. et al. The conformation of tetraalanine in water determined by polarized Raman, FT-IR, and VCD spectroscopy. *Journal of the American Chemical Society* **126**, 2768–2776 (2004).

[63] Schweitzer-Stenner, R.; Measey, T. J. The alanine-rich XAO peptide adopts a heterogeneous population, including turn-like and polyproline II conformations. *Proceedings of the National Academy of Sciences* **104**, 6649–6654 (2007).

[64] Zagrovic, B. et al. Unusual compactness of a polyproline type II structure. *Proceedings of the National Academy of Sciences of the United States of America* **102**, 11698–11703 (2005).

[65] Makowska, J. et al. Polyproline II conformation is one of many local conformational states and is not an overall conformation of unfolded peptides and proteins. *Proceedings of the National Academy of Sciences of the United States of America* **103**, 1744–1749 (2006).

[66] Rucker, A. L. et al. Host-guest scale of left-handed polyproline II helix formation. *Proteins: Structure, Function, and Bioinformatics* **53**, 68–75 (2003).

[67] Hagarman, A. et al. Intrinsic Propensities of Amino Acid Residues in GxG Peptides Inferred from Amide I′ Band Profiles and NMR Scalar Coupling Constants. *Journal of the American Chemical Society* **132**, 540–551 (2009).

[68] Ding, L. et al. The Pentapeptide GGAGG Has PII Conformation. *Journal of the American Chemical Society* **125**, 8092–8093 (2003).

[69] Woutersen, S. et al. Peptide conformational heterogeneity revealed from nonlinear vibrational spectroscopy and molecular-dynamics simulations. *The Journal of Chemical Physics* **117**, 6833–6840 (2002).

[70] Graf, J.; Nguyen, P. H.; Stock, G.; Schwalbe, H. Structure and dynamics of the homologous series of alanine peptides: a joint molecular dynamics/NMR study. *Journal of the American Chemical Society* **129**, 1179–1189 (2007).

[71] Schweitzer-Stenner, R. et al. Conformations of alanine-based peptides in water probed by FTIR, Raman, vibrational circular dichroism, electronic circular dichroism, and NMR spectroscopy. *Biochemistry* **46**, 1587–1596 (2007).

[72] Eker, F.; Cao, X.; Nafie, L.; Schweitzer-Stenner, R. Tripeptides adopt stable structures in water. A combined polarized visible Raman, FTIR, and VCD spectroscopy study. *Journal of the American Chemical Society* **124**, 14330–14341 (2002).

[73] Jha, A. K. et al. Helix, sheet, and polyproline II frequencies and strong nearest neighbor effects in a restricted coil library. *Biochemistry* **44**, 9691–9702 (2005).

[74] Fiebig, K. M. et al. Toward a Description of the Conformations of Denatured States of Proteins. Comparison of a Random Coil Model with NMR Measurements. *The Journal of Physical Chemistry* **100**, 2661–2666 (1996).

[75] Gnanakaran, S.; Garcia, A. E. Validation of an All-Atom Protein Force Field: From Dipeptides to Larger Peptides. *The Journal of Physical Chemistry B* **107**, 12555–12557 (2003).

[76] Duan, Y. et al. A point-charge force field for molecular mechanics simulations of proteins based on condensed-phase quantum mechanical calculations. *Journal of computational chemistry* **24**, 1999–2012 (2003).

[77] Smith, L. J. et al. Analysis of main chain torsion angles in proteins: prediction of NMR coupling constants for native and random coil conformations. *Journal of Molecular Biology* **255**, 494–506 (1996).

[78] Peti, W.; Hennig, M.; Smith, L. J.; Schwalbe, H. NMR Spectroscopic Investigation of ψ Torsion Angle Distribution in Unfolded Ubiquitin from Analysis of 3J(Cα,Cα) Coupling Constants and Cross-Correlated Relaxation Rates. *Journal of the American Chemical Society* **122**, 12017–12018 (2000).

[79] Eker, F. et al. Preferred peptide backbone conformations in the unfolded state revealed by the structure analysis of alanine-based (AXA) tripeptides in aqueous solution. *Proceedings of the National Academy of Sciences of the United States of America* **101**, 10054–10059 (2004).

[80] Han, W.-G.; Jalkanen, K. J.; Elstner, M.; Suhai, S. Theoretical Study of Aqueous N-Acetyl-l-alanine N'-Methylamide: Structures and Raman, VCD, and ROA Spectra. *The Journal of Physical Chemistry B* **102**, 2587–2602 (1998).

[81] Fandrich, M.; Fletcher, M. A.; Dobson, C. M. Amyloid fibrils from muscle myoglobin. *Nature* **410**, 165–166 (2001).

[82] Grigsby, J. J.; Blanch, H. W.; Prausnitz, J. M. Effect of secondary structure on the potential of mean force for poly-L-lysine in the alpha-helix and beta-sheet conformations. *Biophysical Chemistry* **99**, 107–116 (2002).

[83] Painter, P. C.; Koenig, J. L. Solution Conformation of Poly(L-Lysine) - Raman and Infrared Spectroscopic Study. *Biopolymers* **15**, 229–240 (1976).

[84] Darke, A.; Finer, E. G. Nmr-Studies of Mixtures of Poly-L-Lysine Hydrobromide with Water. *Biopolymers* **14**, 441–455 (1975).

[85] Joubert, F. J.; Lotan, N.; Scheraga, H. A. A nuclear magnetic resonance study of the helix coil transition of poly L lysine in methanol water solvents. *Physiological Chemistry and Physics* **1**, 348 (1969).

[86] Perly, B.; Chevalier, Y.; Chachaty, C. Nmr and Electron-Spin-Resonance Study of the Conformations and Dynamical Properties of Poly(L-Lysine) in Aqueous-Solutions. *Macromolecules* **14**, 969–975 (1981).

[87] McColl, I. H. et al. A new perspective on beta-sheet structures using vibrational Raman optical activity: from poly(L-lysine) to the prion protein. *Journal of the American Chemical Society* **125**, 10019–10026 (2003).

[88] Barron, L. D.; Hecht, L.; Blanch, E. W.; Bell, A. F. Solution structure and dynamics of biomolecules from Raman optical activity. *Progress in Biophysics & Molecular Biology* **73**, 1–49 (2000).

[89] Wilson, G.; Hecht, L.; Barron, L. D. Vibrational Raman optical activity of alpha-helical and unordered poly(L-lysine). *Journal of the Chemical Society-Faraday Transactions* **92**, 1503–1509 (1996).

[90] Keiderling, T. A.; Silva, R. A.; Yoder, G.; Dukor, R. K. Vibrational circular dichroism spectroscopy of selected oligopeptide conformations. *Bioorganic & Medicinal Chemistry* **7**, 133–141 (1999).

[91] Song, S.; Asher, S. A. UV Resonance Raman Studies of Peptide Conformation in Poly(L-lysine), Poly(L-glutamic acid), and Model Complexes: The Basis for Protein Secondary Structure Determinations. *Journal of the American Chemical Society* **111**, 4295–4305 (1988).

[92] Wang, Y.; Purrello, R.; Georgiou, S.; Spiro, T. G. UVRR spectroscopy of the peptide bond. 2. Carbonyl H-bond effects on the ground- and excited-state structures of N-methylacetamide. *Journal of the American Chemical Society* **113**, 6368–6377 (1991).

[93] Mikhonin, A. V.; Myshakina, N. S.; Bykov, S. V.; Asher, S. A. UV resonance Raman determination of polyproline II, extended 2.5(1)-helix, and beta-sheet Psi angle energy landscape in poly-L-lysine and poly-L-glutamic acid. *Journal of the American Chemical Society* **127**, 7712–7720 (2005).

[94] Anderluh, G.; Zerovnik, E. Pore formation by human stefin B in its native and oligomeric states and the consequent amyloid induced toxicity. *Frontiers in Molecular Neuroscience* **5**, (2012).

[95] Hartmann, W.; Galla, H. J. Binding of polylysine to charged bilayer membranes: molecular organization of a lipid.peptide complex. *Biochimica et Biophysica Acta* **509**, 474–490 (1978).

[96] Takahashi, H.; Yasue, T.; Ohki, K.; Hatta, I. Structure and phase behaviour of dimyristoylphosphatidic acid/poly(L-lysine) systems. *Molecular Membrane Biology* **13**, 233–240 (1996).

[97] Ben-Tal, N. et al. Binding of small basic peptides to membranes containing acidic lipids: theoretical models and experimental results. *Biophysical Journal* **71**, 561–575 (1996).

[98] Carrier, D.; Mantsch, H. H.; Wong, P. T. T. Protective Effect of Lipidic Surfaces against Pressure-Induced Conformational-Changes of Poly(L-Lysine). *Biochemistry* **29**, 254–258 (1990).

[99] Ramachandran, G. N.; Ramakrishnan, C.; Sasisekharan, V. Stereochemistry of polypeptide chain configurations. *Journal of Molecular Biology* **7**, 95–99 (1963).

[100] Avbelj, F. Solvation and Electrostatics as Determinants of Local Structural Order in Unfolded Peptides and Proteins. In: *Protein and Peptide Folding, Misfolding, and Non-Folding*. 131–158 (John Wiley & Sons, New Jersy, 2012).

[101] Perskie, L. L.; Street, T. O.; Rose, G. D. Structures, basins, and energies: A deconstruction of the Protein Coil Library. *Protein Science* **17**, 1151–1161 (2008).

[102] Shi, Z.; Chen, K.; Liu, Z.; Kallenbach, N. R. Conformation of the backbone in unfolded proteins. *Chemical Reviews* **106**, 1877–1897 (2006).

[103] Poon, C.-D.; Samulski, E. T.; Weise, C. F.; Weisshaar, J. C. Do Bridging Water Molecules Dictate the Structure of a Model Dipeptide in Aqueous Solution? *Journal of the American Chemical Society* **122**, 5642–5643 (2000).

[104] Kuemin, M.; Schweizer, S.; Ochsenfeld, C.; Wennemers, H. Effects of Terminal Functional Groups on the Stability of the Polyproline II Structure: A Combined Experimental and Theoretical Study. *Journal of the American Chemical Society* **131**, 15474–15482 (2009).

[105] Milner-White, E. J. Situations of gamma-turns in proteins: Their relation to alpha-helices, beta-sheets and ligand binding sites. *Journal of Molecular Biology* **216**, 385–397 (1990).

[106] Liu, Z. et al. Solvent dependence of PII conformation in model alanine peptides. *Journal of the American Chemical Society* **126**, 15141–15150 (2004).

[107] Gnanakaran, S.; Hochstrasser, R. M.; García, A. E. Nature of structural inhomogeneities on folding a helix and their influence on spectral measurements. *Proceedings of the National Academy of Sciences of the United States of America* **101**, 9229–9234 (2004).

[108] Walsh, S. T. et al. The hydration of amides in helices; a comprehensive picture from molecular dynamics, IR, and NMR. *Protein Science* **12**, 520–531 (2003).

[109] Avbelj, F.; Baldwin, R. L. Origin of the change in solvation enthalpy of the peptide group when neighboring peptide groups are added. *Proceedings of the National Academy of Sciences* **106**, 3137–3141 (2009).

[110] Avbelj, F.; Baldwin, R. L. Origin of the neighboring residue effect on peptide backbone conformation. *Proceedings of the National Academy of Sciences of the United States of America* **101**, 10967–10972 (2004).

[111] Avbelj, F. Amino Acids Conformational Preferences and Solvation of the Backbone Atoms in Peptides and Proteins. *Journal of Molecular Biology* **300**, 1335–1359 (2000).

[112] Ross, C. A.; Poirier, M. A. Protein aggregation and neurodegenerative disease. *Natural Medicines* **10**, 10–17 (2004).

[113] Morris, A. M.; Watzky, M. A.; Finke, R. G. Protein aggregation kinetics, mechanism, and curve-fitting: a review of the literature. *Biochimica et Biophysica Acta* **1794**, 375–397 (2009).

[114] Irvine, G. B.; El-Agnaf, O. M.; Shankar, G. M.; Walsh, D. M. Protein aggregation in the brain: the molecular basis for Alzheimer's and Parkinson's diseases. *Molecular Medicine* **14**, 451–464 (2008).

[115] Uversky, V. N. Mysterious oligomerization of the amyloidogenic proteins. *FEBS Journal* **277**, 2940–2953 (2010).

[116] Chiti, F.; Dobson, C. M. Protein misfolding, functional amyloid, and human disease. *Annual Review of Biochemistry* **75**, 333–366 (2006).

[117] Horwich, A. L.; Weissman, J. S. Deadly conformations–protein misfolding in prion disease. *Cell* **89**, 499–510 (1997).

[118] Lansbury, P. T. Inhibition of amyloid formation: a strategy to delay the onset of Alzheimer's disease. *Current Opinion in Chemical Biology* **1**, 260–267 (1997).

[119] Conway, K. A.; Harper, J. D.; Lansbury, P. T. Accelerated in vitro fibril formation by a mutant alpha-synuclein linked to early-onset Parkinson disease. *Natural Medicines* **4**, 1318–1320 (1998).

[120] Sunde, M. et al. Common core structure of amyloid fibrils by synchrotron X-ray diffraction. *Journal of Molecular Biology* **273**, 729–739 (1997).
[121] Stefani, M. Protein Folding and Misfolding, Relevance to Disease, and Biological Function. In: *Protein Misfolding in Neurodegenerative Diseases*. 1–66 (CRC Press, Boca Raton, 2007).
[122] Dobson, C. M. Protein misfolding, evolution and disease. *Trends in Biochemical Sciences* **24**, 329–332 (1999).
[123] Pertinhez, T. A. et al. Amyloid fibril formation by a helical cytochrome. *FEBS Letters* **495**, 184–186 (2001).
[124] DuBay, K. F. et al. Prediction of the absolute aggregation rates of amyloidogenic polypeptide chains. *Journal of Molecular Biology* **341**, 1317–1326 (2004).
[125] Dobson, C. M. Protein folding and misfolding. *Nature* **426**, 884–890 (2003).
[126] Zerovnik, E. et al. Mechanisms of amyloid fibril formation - focus on domain-swapping. *FEBS Journal* **278**, 2263–2282 (2011).
[127] Guijarro, J. I. et al. Amyloid fibril formation by an SH3 domain. *Proceedings of the National Academy of Sciences of the United States of America* **95**, 4224–4228 (1998).
[128] Kelly, J. W. The alternative conformations of amyloidogenic proteins and their multi-step assembly pathways. *Current Opinion in Structural Biology* **8**, 101–106 (1998).
[129] Rochet, J. C.; Lansbury, P. T. Amyloid fibrillogenesis: themes and variations. *Current Opinion in Structural Biology* **10**, 60–68 (2000).
[130] Lansbury, P. T. Evolution of amyloid: What normal protein folding may tell us about fibrillogenesis and disease. *Proceedings of the National Academy of Sciences of the United States of America* **96**, 3342–3344 (1999).
[131] Žerovnik, E. Amyloid-Fibril Formation. *European Journal of Biochemistry* **269**, 3362–3371 (2002).
[132] Kelly, J. W. Mechanisms of amyloidogenesis. *Nature Structural Biology* **7**, 824–826 (2000).
[133] Jarrett, J. T.; Lansbury, P. T., Jr. Seeding "one-dimensional crystallization" of amyloid: a pathogenic mechanism in Alzheimer's disease and scrapie? *Cell* **73**, 1055–1058 (1993).
[134] Harper, J. D.; Wong, S. S.; Lieber, C. M.; Lansbury, P. T. Observation of metastable A² amyloid protofibrils by atomic force microscopy. *Chemistry and Biology* **4**, 119–125 (1997).
[135] Walsh, D. M. et al. Amyloid beta-protein fibrillogenesis. Detection of a protofibrillar intermediate. *Journal of Biological Chemistry* **272**, 22364–22372 (1997).
[136] Aggeli, A. et al. Hierarchical self-assembly of chiral rod-like molecules as a model for peptide beta -sheet tapes, ribbons, fibrils, and fibers. *Proceedings of the National Academy of Sciences of the United States of America* **98**, 11857–11862 (2001).
[137] Fändrich, M. Oligomeric Intermediates in Amyloid Formation: Structure Determination and Mechanisms of Toxicity. *Journal of Molecular Biology* **421**, 427–440 (2012).
[138] Kayed, R. et al. Common structure of soluble amyloid oligomers implies common mechanism of pathogenesis. *Science* **300**, 486–489 (2003).
[139] Quist, A. et al. Amyloid ion channels: A common structural link for protein-misfolding disease. *Proceedings of the National Academy of Sciences of the United States of America* **102**, 10427–10432 (2005).
[140] Lasagna-Reeves, C. A.; Glabe, C. G.; Kayed, R. Amyloid-β Annular Protofibrils Evade Fibrillar Fate in Alzheimer Disease Brain. *Journal of Biological Chemistry* **286**, 22122–22130 (2011).
[141] Hebda, J. A.; Miranker, A. D. The interplay of catalysis and toxicity by amyloid intermediates on lipid bilayers: insights from type II diabetes. *Annual review of biophysics* **38**, 125–152 (2009).
[142] Lashuel, H. A.; Lansbury, P. T. Are amyloid diseases caused by protein aggregates that mimic bacterial pore-forming toxins? *Quarterly Reviews of Biophysics* **39**, 167–201 (2006).
[143] Middleton, E. R.; Rhoades, E. Effects of Curvature and Composition on Alpha-Synuclein Binding to Lipid Vesicles. *Biophysical Journal* **99**, 2279–2288 (2010).

[144] Eker, F.; Griebenow, K.; Schweitzer-Stenner, R. Abeta(1-28) fragment of the amyloid peptide predominantly adopts a polyproline II conformation in an acidic solution. *Biochemistry* **43**, 6893–6898 (2004).

[145] Darnell, G.; Orgel, J. P. R. O.; Pahl, R.; Meredith, S. C. Flanking Polyproline Sequences Inhibit β-Sheet Structure in Polyglutamine Segments by Inducing PPII-like Helix Structure. *Journal of Molecular Biology* **374**, 688–704 (2007).

[146] Darnell, G. D.; Derryberry, J.; Kurutz, J. W.; Meredith, S. C. Mechanism of Cis-Inhibition of PolyQ Fibrillation by PolyP: PPII Oligomers and the Hydrophobic Effect. *Biophysical Journal* **97**, 2295–2305 (2009).

[147] Bryant, C. et al. Acid stabilization of insulin. *Biochemistry* **32**, 8075–8082 (1993).

[148] Halban, P. A. Structural domains and molecular lifestyles of insulin and its precursors in the pancreatic Beta cell. *Diabetologia* **34**, 767–778 (1991).

[149] Dodson, G.; Steiner, D. The role of assembly in insulin's biosynthesis. *Current Opinion in Structural Biology* **8**, 189–194 (1998).

[150] Whittingham, J. L. et al. I222 crystal form of despentapeptide (B26-B30) insulin provides new insights into the properties of monomeric insulin. *Acta Crystallographica Section D* **62**, 505–511 (2006).

[151] Abel, J. J. Crystalline Insulin. *Proceedings of the National Academy of Sciences* **12**, 132–136 (1926).

[152] Blatherwick, N. R. et al. STUDIES ON INSULIN. *Journal of Biological Chemistry* **72**, 57–89 (1927).

[153] Krogh, A.; Hemmingsen, A. M. The destructive action of heat on insulin solutions. *The Biochemical journal* **22**, 1231–1238 (1928).

[154] Waugh, D. F. The linkage of corpuscular protein molecules. I. a bibrous modification of insulin. *Journal of the American Chemical Society* **66**, 663–663 (1944).

[155] Waugh, D. F.; Wilhelmson, D. F.; Commerford, S. L.; Sackler, M. L. Studies of the Nucleation and Growth Reactions of Selected Types of Insulin Fibrils. *Journal of the American Chemical Society* **75**, 2592–2600 (1953).

[156] Nielsen, L. et al. Effect of Environmental Factors on the Kinetics of Insulin Fibril Formation: Elucidation of the Molecular Mechanism. *Biochemistry* **40**, 6036–6046 (2001).

[157] Sluzky, V.; Klibanov, A. M.; Langer, R. Mechanism of insulin aggregation and stabilization in agitated aqueous solutions. *Biotechnology and Bioengineering* **40**, 895–903 (1992).

[158] Brange, J. et al. Toward understanding insulin fibrillation. *Journal of Pharmaceutical Sciences* **86**, 517–525 (1997).

[159] Millican, R. L.; Brems, D. N. Equilibrium Intermediates in the Denaturation of Human Insulin and Two Monomeric Insulin Analogs. *Biochemistry* **33**, 1116–1124 (1994).

[160] Hong, D.-P.; Ahmad, A.; Fink, A. L. Fibrillation of Human Insulin A and B Chains. *Biochemistry* **45**, 9342–9353 (2006).

[161] Lee, C.-C. et al. A Three-Stage Kinetic Model of Amyloid Fibrillation. *Biophysical Journal* **92**, 3448–3458 (2007).

[162] Hua, Q. X.; Weiss, M. A. Mechanism of insulin fibrillation: the structure of insulin under amyloidogenic conditions resembles a protein-folding intermediate. *Journal of Biological Chemistry* **279**, 21449–21460 (2004).

[163] Dobson, C. M.; Karplus, M. The fundamentals of protein folding: bringing together theory and experiment. *Current Opinion in Structural Biology* **9**, 92–101 (1999).

[164] Ahmad, A.; Uversky, V. N.; Hong, D.; Fink, A. L. Early Events in the Fibrillation of Monomeric Insulin. *Journal of Biological Chemistry* **280**, 42669–42675 (2005).

[165] Nielsen, L. et al. Probing the Mechanism of Insulin Fibril Formation with Insulin Mutants. *Biochemistry* **40**, 8397–8409 (2001).

[166] Störkel, S.; Schneider, H. M.; Müntefering, H.; Kashiwagi, S. Iatrogenic, insulin-dependent, local amyloidosis. *Laboratory investigation; a journal of technical methods and pathology* **48**, 108–111 (1983).

[167] Dische, F. E. et al. Insulin as an amyloid-fibril protein at sites of repeated insulin injections in a diabetic patient. *Diabetologia* **31**, 158–161 (1988).

[168] Swift, B.; Hawkins, P. N.; Richards, C.; Gregory, R. Examination of insulin injection sites: an unexpected finding of localized amyloidosis. *Diabetic Medicine* **19**, 881–882 (2002).

[169] Shikama, Y. et al. Localized Amyloidosis at the Site of Repeated Insulin Injection in a Diabetic Patient. *Internal Medicine* **49**, 397–401 (2010).

[170] Chi, E.; Krishnan, S.; Randolph, T.; Carpenter, J. Physical Stability of Proteins in Aqueous Solution: Mechanism and Driving Forces in Nonnative Protein Aggregation. *Pharmaceutical Research* **20**, 1325–1336 (2003).

[171] Brange, J.; Havelund, S.; Hougaard, P. Chemical Stability of Insulin. 2. Formation of Higher Molecular Weight Transformation Products During Storage of Pharmaceutical Preparations. *Pharmaceutical Research* **9**, 727–734 (1992).

[172] Williams, R. H.; Melmed, S. *Williams textbook of endocrinology* (Elsevier/Saunders, Philadelphia, 2011).

[173] Wild, S. et al. Global Prevalence of Diabetes: Estimates for the year 2000 and projections for 2030. *Diabetes Care* **27**, 1047–1053 (2004).

[174] Brange, J.; Langkjaer, L. Chemical stability of insulin. 3. Influence of excipients, formulation, and pH. *Acta Pharmaceutica Nordica* **4**, 149–158 (1992).

[175] Reichenbächer, M.; Popp, J. *Challenges in Molecular Structure Determination* (Springer, Berlin, 2012).

[176] Barron, L. D.; Buckingham, A. D. Vibrational optical activity. *Chemical Physics Letters* **492**, 199–213 (2010).

[177] Measey, T. J.; Schweitzer-Stenner, R. Vibrational Circular Dichroism as a Probe of Fibrillogenesis: The Origin of the Anomalous Intensity Enhancement of Amyloid-like Fibrils. *Journal of the American Chemical Society* **133**, 1066–1076 (2010).

[178] Grdadolnik, J. Infrared difference spectroscopy: Part I. Interpretation of the difference spectrum. *Vibrational Spectroscopy* **31**, 279–288 (2003).

[179] Austin, J. C.; Jordan, T.; Spiro, T. G. Ultraviolet resonance Raman studies of proteins and related model compounds, in Biomolecular Spectroscopy. In: *Biomolecular spectroscopy*. 55–127 (Wiley & Sons Ltd., New York, 1993).

[180] Grdadolnik, J. ATR-FTIR spectroscopy: Its advantages and limitations. *Acta Chimica Slovenica* **49**, 631–642 (2002).

[181] Frisch, M. J. *Gaussian 09 User's Reference* (Gaussian, Incorporated, Wallingford, 2009).

[182] Lee, C.; Yang, W.; Parr, R. G. Development of the Colle-Salvetti correlation-energy formula into a functional of the electron density. *Physical Review B* **37**, 785–789 (1988).

[183] Becke, A. D. Density-functional exchange-energy approximation with correct asymptotic behavior. *Physical Review A* **38**, 3098–3100 (1988).

[184] Becke, A. D. Density-functional thermochemistry. III. The role of exact exchange. *The Journal of Chemical Physics* **98**, 5648–5652 (1993).

[185] Tomasi, J.; Mennucci, B.; Cammi, R. Quantum Mechanical Continuum Solvation Models. *Chemical Reviews* **105**, 2999–3094 (2005).

[186] Brooks, B. R. et al. CHARMM: the biomolecular simulation program. *Journal of computational chemistry* **30**, 1545–1614 (2009).

[187] Foresman, J. B.; Frisch, A. E. *Exploring chemistry with electronic structure methods* (Gaussian, Inc., Pittsburgh, 1996).

[188] Karplus , M. Contact Electron-Spin Coupling of Nuclear Magnetic Moments. *Journal of Chemical Physics* **30**, 11–15 (1959).

[189] Vuister, G. W.; Bax, A. Quantitative J correlation: a new approach for measuring homonuclear three-bond J(HNH.alpha.) coupling constants in 15N-enriched proteins. *Journal of the American Chemical Society* **115**, 7772–7777 (1993).

[190] Hanwell, M. et al. Avogadro: an advanced semantic chemical editor, visualization, and analysis platform. *Journal of Cheminformatics* **4**, 17 (2012).

[191] Levine, H. Thioflavine T interaction with synthetic Alzheimer's disease β-amyloid peptides: Detection of amyloid aggregation in solution. *Protein Science* **2**, 404–410 (1993).

[192] Rabzelj, S.; Turk, V.; Žerovnik, E. In vitro study of stability and amyloid-fibril formation of two mutants of human stefin B (cystatin B) occurring in patients with EPM1. *Protein Science* **14**, 2713–2722 (2005).

[193] Biancalana, M.; Koide, S. Molecular mechanism of Thioflavin-T binding to amyloid fibrils. *Biochimica et Biophysica Acta (BBA) - Proteins and Proteomics* **1804**, 1405–1412 (2010).

[194] Vass, E.; Kurz, M.; Konat, R. K.; Hollosi, M. FTIR and CD spectroscopic studies on cyclic penta- and hexa-peptides. Detailed examination of hydrogen bonding in beta- and gamma-turns determined by NMR. *Spectrochimica Acta Part A: Molecular and Biomolecular Spectroscopy* **54**, 773–786 (1998).

[195] Vass, E.; Majer, Z.; Kohalmy, K.; Hollosi, M. Vibrational and chiroptical spectroscopic characterization of gamma-turn model cyclic tetrapeptides containing two beta-Ala residues. *Chirality* **22**, 762–771 (2010).

[196] Pohl, G. et al. A matrix isolation study on Ac-Gly-NHMe and Ac-l-Ala-NHMe, the simplest chiral and achiral building blocks of peptides and proteins. *Physical Chemistry Chemical Physics* **9**, 4698–4708 (2007).

[197] Kelly, S. M.; Jess, T. J.; Price, N. C. How to study proteins by circular dichroism. *Biochimica et Biophysica Acta* **1751**, 119–139 (2005).

[198] Baxter, N. J.; Williamson, M. P. Temperature dependence of 1H chemical shifts in proteins. *Journal of Biomolecular NMR* **9**, 359–369 (1997).

[199] Belvisi, L. et al. Conformational Preferences of Peptides Containing Reverse-Turn Mimetic Bicyclic Lactams: Inverse γ-Turns versus Type-II′ β-Turns – Insights into β-Hairpin Stability. *European Journal of Organic Chemistry* **1999**, 389–400 (1999).

[200] Stevens, E. S.; Sugawara, N.; Bonora, G. M.; Toniolo, C. Conformational analysis of linear peptides. 3. Temperature dependence of NH chemical shifts in chloroform. *Journal of the American Chemical Society* **102**, 7048–7050 (1980).

[201] Chatterjee, S. et al. Multiple conformational states in crystals and in solution in alphagamma hybrid peptides. Fragility of the C12 helix in short sequences. *The Journal of organic chemistry* **73**, 6595–6606 (2008).

[202] Garcia, A. E. Characterization of non-alpha helical conformations in Ala peptides. *Polymer* **45**, 669–676 (2004).

[203] Mezei, M.; Fleming, P. J.; Srinivasan, R.; Rose, G. D. Polyproline II helix is the preferred conformation for unfolded polyalanine in water. *Proteins: Structure, Function, and Bioinformatics* **55**, 502–507 (2004).

[204] Law, P. B.; Daggett, V. The relationship between water bridges and the polyproline II conformation: a large-scale analysis of molecular dynamics simulations and crystal structures. *Protein Engineering Design and Selection* **23**, 27–33 (2010).

[205] Sharma, B.; Asher, S. A. UV Resonance Raman Investigation of the Conformations and Lowest Energy Allowed Electronic Excited States of Tri- and Tetraalanine: Charge Transfer Transitions. *The Journal of Physical Chemistry B* **114**, 6661–6668 (2010).

[206] Gong, H.; Rose, G. D. Assessing the solvent-dependent surface area of unfolded proteins using an ensemble model. *Proceedings of the National Academy of Sciences* **105**, 3321–3326 (2008).

[207] Ramachandran, G. N.; Sasisekharan, V. Conformation of polypeptides and proteins. *Advances in Protein Chemistry* **23**, 283–438 (1968).

[208] Ptitsyn, O. B. et al. Evidence for a molten globule state as a general intermediate in protein folding. *FEBS Letters* **262**, 20–24 (1990).

[209] Vass, E.; Hollosi, M.; Besson, F.; Buchet, R. Vibrational spectroscopic detection of beta- and gamma-turns in synthetic and natural peptides and proteins. *Chemical Reviews* **103**, 1917–1954 (2003).

[210] Luo, P.; Baldwin, R. L. Interaction between water and polar groups of the helix backbone: An important determinant of helix propensities. *Proceedings of the National Academy of Sciences* **96**, 4930–4935 (1999).

[211] Lyu, P. C.; Sherman, J. C.; Chen, A.; Kallenbach, N. R. Alpha-helix stabilization by natural and unnatural amino acids with alkyl side chains. *Proceedings of the National Academy of Sciences* **88**, 5317–5320 (1991).

[212] Dzwolak, W.; Smirnovas, V. A conformational α-helix to β-sheet transition accompanies racemic self-assembly of polylysine: an FT-IR spectroscopic study. *Biophysical Chemistry* **115**, 49–54 (2005).

[213] Kubelka, J.; Keiderling, T. A. Differentiation of beta-sheet-forming structures: Ab initio-based simulations of IR absorption and vibrational CD for model peptide and protein beta-sheets. *Journal of the American Chemical Society* **123**, 12048–12058 (2001).

[214] Bandekar, J. Amide modes and protein conformation. *Biochimica et Biophysica Acta* **1120**, 123–143 (1992).

[215] Baello, B. I.; Pancoska, P.; Keiderling, T. A. Vibrational circular dichroism spectra of proteins in the amide III region: Measurement and correlation of bandshape to secondary structure. *Analytical Biochemistry* **250**, 212–221 (1997).

[216] Fu, F. N. et al. Secondary Structure Estimation of Proteins Using the Amide-Iii Region of Fourier-Transform Infrared-Spectroscopy - Application to Analyze Calcium Binding-Induced Structural-Changes in Calsequestrin. *Applied Spectroscopy* **48**, 1432–1441 (1994).

[217] Peternel, S.; Grdadolnik, J.; Gaberc-Porekar, V.; Komel, R. Engineering inclusion bodies for non denaturing extraction of functional proteins. *Microbial Cell Factories* **7**, 34 (2008).

[218] Jevsevar, S. et al. Production of nonclassical inclusion bodies from which correctly folded protein can be extracted. *Biotechnology Progress* **21**, 632–639 (2005).

[219] Yamamoto, S.; Kaminský, J.; Bouř, P. Structure and Vibrational Motion of Insulin from Raman Optical Activity Spectra. *Analytical Chemistry* **84**, 2440–2451 (2012).

[220] Torii, H.; Tatsumi, T.; Tasumi, M. Effects of hydration on the structure, vibrational wavenumbers, vibrational force field and resonance Raman intensities of N-methylacetamide. *Journal of Raman Spectroscopy* **29**, 537–546 (1998).

[221] Asher, S. A.; Mikhonin, A. V.; Bykov, S. UV Raman demonstrates that alpha-helical polyalanine peptides melt to polyproline II conformations. *Journal of the American Chemical Society* **126**, 8433–8440 (2004).

[222] Mikhonin, A. V.; Ahmed, Z.; Ianoul, A.; Asher, S. A. Assignments and conformational dependencies of the amide III peptide backbone UV resonance Raman bands. *Journal of Physical Chemistry B* **108**, 19020–19028 (2004).

[223] Vianello, R.; Maksić, Z. B. Polycyano Derivatives of some Organic Tri- and Hexacyclic Molecules Are Powerful Super- and Hyperacids in the Gas Phase and DMSO: Computational Study by DFT Approach. *The Journal of organic chemistry* **75**, 7670–7681 (2010).

[224] Perrin, D. D.; Dempsey, B.; Serjeant, E. P. *pKa prediction for organic acids and bases* (Chapman and Hall, London, 1981).

[225] Dzwolak, W. et al. The diastereomeric assembly of polylysine is the low-volume pathway for preferential formation of beta-sheet aggregates. *Journal of the American Chemical Society* **126**, 3762–3768 (2004).

[226] Lamm, M. S.; Rajagopal, K.; Schneider, J. P.; Pochan, D. J. Laminated Morphology of Nontwisting β-Sheet Fibrils Constructed via Peptide Self-Assembly. *Journal of the American Chemical Society* **127**, 16692–16700 (2005).

[227] Ma, L.; Ahmed, Z.; Mikhonin, A. V.; Asher, S. A. UV resonance Raman measurements of poly-L-lysine's conformational energy landscapes: dependence on perchlorate concentration and temperature. *Journal of Physical Chemistry B* **111**, 7675–7680 (2007).

[228] Schwieger, C.; Blume, A. Interaction of poly(L-lysines) with negatively charged membranes: an FT-IR and DSC study. *European Biophysics Journal with Biophysics Letters* **36**, 437–450 (2007).

[229] Huang, R. et al. Cross-Strand Coupling of a β-Hairpin Peptide Stabilized with an Aib-Gly Turn Studied Using Isotope-Edited IR Spectroscopy. *Journal of the American Chemical Society* **129**, 13592–13603 (2007).

[230] Baumruk, V.; Pancoska, P.; Keiderling, T. A. Predictions of secondary structure using statistical analyses of electronic and vibrational circular dichroism and Fourier transform infrared spectra of proteins in H2O. *Journal of Molecular Biology* **259**, 774–791 (1996).

[231] Arunkumar, A. I.; Kumar, T. K. S.; Yu, C. Specificity of helix-induction by 2,2,2-trifluoroethanol in polypeptides. *International Journal of Biological Macromolecules* **21**, 223–230 (1997).

[232] Chi, H. et al. Inter-residue Coupling and Equilibrium Unfolding of PM Helical Peptides. Vibrational Spectra Enhanced with C-13 Isotopic Labeling. *Journal of Physical Chemistry B* **114**, 12744–12753 (2010).

[233] Jackson, M.; Mantsch, H. H. Beware of Proteins in Dmso. *Biochimica et Biophysica Acta* **1078**, 231–235 (1991).

[234] Duarte, A. M. S.; van Mierlo, C. P. M.; Hemminga, M. A. Molecular dynamics study of the solvation of an alpha-helical transmembrane peptide by DMSO. *Journal of Physical Chemistry B* **112**, 8664–8671 (2008).

[235] Luo, P.; Baldwin, R. L. Mechanism of Helix Induction by Trifluoroethanol: A Framework for Extrapolating the Helix-Forming Properties of Peptides from Trifluoroethanol/Water Mixtures Back to Water†. *Biochemistry* **36**, 8413–8421 (1997).

[236] Thomas, P. D.; Dill, K. A. Local and Nonlocal Interactions in Globular-Proteins and Mechanisms of Alcohol Denaturation. *Protein Science* **2**, 2050–2065 (1993).

[237] Javid, N. et al. Protein-protein interactions in complex cosolvent solutions. *Chemphyschem* **8**, 679–689 (2007).

[238] Zimm, B. H.; Bragg, J. K. Theory of the Phase Transition between Helix and Random Coil in Polypeptide Chains. *The Journal of Chemical Physics* **31**, 526–535 (1959).

[239] Lifson, S.; Roig, A. On the Theory of Helix—Coil Transition in Polypeptides. *The Journal of Chemical Physics* **34**, 1963 (1961).

[240] Poland, D.; Scheraga, H. A. *Theory of helix-coil transitions in biopolymers: statistical mechanical theory of order-disorder transitions in biological macromolecules* (Academic Press, New York, 1970).

[241] Xiong, K.; Asher, S. A. Circular Dichroism and UV Resonance Raman Study of the Impact of Alcohols on the Gibbs Free Energy Landscape of an alpha-Helical Peptide. *Biochemistry* **49**, 3336–3342 (2010).

[242] Hinderaker, M. P.; Raines, R. T. An electronic effect on protein structure. *Protein Science* **12**, 1188–1194 (2003).

[243] Nerenberg, P. S.; Head-Gordon, T. Optimizing Protein−Solvent Force Fields to Reproduce Intrinsic Conformational Preferences of Model Peptides. *Journal of Chemical Theory and Computation* **7**, 1220–1230 (2011).

[244] Szyc, Ł.; Pilorz, S.; Czarnik-Matusewicz, B. FTIR-ATR investigations of an α-helix to β-sheet conformational transition in poly(l-lysine). *Journal of Molecular Liquids* **141**, 155–159 (2008).

[245] Guex, N.; Peitsch, M. C. SWISS-MODEL and the Swiss-PdbViewer: an environment for comparative protein modeling. *Electrophoresis* **18**, 2714–2723 (1997).

[246] Yamamoto, S.; Watarai, H. Raman optical activity study on insulin amyloid- and prefibril intermediate. *Chirality* **24**, 97–103 (2012).

[247] Surewicz, W. K.; Mantsch, H. H. New insight into protein secondary structure from resolution-enhanced infrared spectra. *Biochimica et Biophysica Acta* **952**, 115–130 (1988).

[248] Dong, A.; Prestrelski, S. J.; Allison, S. D.; Carpenter, J. F. Infrared spectroscopic studies of lyophilization- and temperature-induced protein aggregation. *Journal of Pharmaceutical Sciences* **84**, 415–424 (1995).

[249] Prestrelski, S. J.; Tedeschi, N.; Arakawa, T.; Carpenter, J. F. Dehydration-induced conformational transitions in proteins and their inhibition by stabilizers. *Biophysical Journal* **65**, 661–671 (1993).

[250] Nielsen, L.; Frokjaer, S.; Carpenter, J. F.; Brange, J. Studies of the structure of insulin fibrils by Fourier transform infrared (FTIR) spectroscopy and electron microscopy. *Journal of Pharmaceutical Sciences* **90**, 29–37 (2001).

[251] Lomakin, A. et al. On the nucleation and growth of amyloid beta-protein fibrils: Detection of nuclei and quantitation of rate constants. *Proceedings of the National Academy of Sciences of the United States of America* **93**, 1125–1129 (1996).

[252] Wood, S. J. et al. alpha-synuclein fibrillogenesis is nucleation-dependent - Implications for the pathogenesis of Parkinson's disease. *Journal of Biological Chemistry* **274**, 19509–19512 (1999).

[253] Jarrett, J. T.; Lansbury, P. T. Amyloid fibril formation requires a chemically discriminating nucleation event: studies of an amyloidogenic sequence from the bacterial protein OsmB. *Biochemistry* **31**, 12345–12352 (1992).

[254] Serio, T. R. et al. Nucleated conformational conversion and the replication of conformational information by a prion determinant. *Science* **289**, 1317–1321 (2000).

[255] Dzwolak, W. et al. Conformational Indeterminism in Protein Misfolding: Chiral Amplification on Amyloidogenic Pathway of Insulin. *Journal of the American Chemical Society* **129**, 7517–7522 (2007).

[256] Uversky, V. N.; Fink, A. L. Pathways to Amyloid Fibril Formation: Partially Folded Intermediates in the Fibrillation of Natively Unfolded Proteins. In: *Amyloid Proteins*. 246–273 (Wiley-VCH Verlag GmbH, 2008).

[257] Mansiaux, Y.; Joseph, A. P.; Gelly, J.-C.; de Brevern, A. G. Assignment of PolyProline II Conformation and Analysis of Sequence – Structure Relationship. *PLoS ONE* **6**, e18401 (2011).

[258] Spurlin, T. A.; Gewirth, A. A. Poly-L-lysine-induced morphology changes in mixed anionic/zwitterionic and neat zwitterionic-supported phospholipid bilayers. *Biophysical Journal* **91**, 2919–2927 (2006).

[259] Fandrich, M.; Dobson, C. M. The behaviour of polyamino acids reveals an inverse side chain effect in amyloid structure formation. *EMBO Journal* **21**, 5682–5690 (2002).

[260] Jarvet, J. et al. A left-handed 31 helical conformation in the Alzheimer Aβ(12–28) peptide. *FEBS Letters* **555**, 371–374 (2003).

Index of Figures

Index of Tables

Appendix

Publications

Žerovnik, E.; Stoka, V.; Mirtič, A.; Gunčar, G.; Grdadolnik, J.; Staniforth, R. A.; Turk, D.; Turk, V. Mechanisms of amyloid fibril formation: focus on domain-swapping. *FEBS journal* **278**, 2263–2282 (2011).

Mirtič, A.; Grdadolnik, J. Conformational preferences of alanine tripeptide in water, trifluoroethanol and dimethyl sulfoxide studied by vibrational spectroscopy. In: Proceedings of 4th Jožef Stefan International Postgraduate School Students Conference. 268–274 (2012).

Mirtič, A.; Grdadolnik, J. The Solvent Dependent Structures of Poly-L-Lysine Determined by Infrared Spectroscopy. Biophysical Chemistry **175-176**, 47–53 (2013).

Mirtič, A.; Merzel, F.; Grdadolnik, J. Vibrational Circular Dichroism as a Probe to Detect Conformational Preferences of Alanine Dipeptide in Different Solvents. (manuscript send for publication).

Mirtič, A.; Tušek Žnidarič, M.; Grdadolnik J. Structural characterization of intermediates during fibrillation of poly-L-lysine obtained by infrared spectroscopy (manuscript in preparation).