# Infrared spectroscopic characterization of secondary structure elements of gas-phase biomolecules

Am Fritz-Haber-Institut
der Max-Planck-Gesellschaft
entstandene und

im Fachbereich Physik
der Freien Universität Berlin
eingereichte Dissertation von

## Peter Kupser

Januar 2011

Eidesstattliche Erklärung gemäß §7, Ziffer (4) der Promotionsordnung vom 05.02.2010 des Fachbereichs Physik an der Freien Universität Berlin

Die Disseration habe ich selbstständig angefertigt. Alle Hilfsmittel und Hilfen habe ich angegeben, insbesondere habe ich die wörtlich oder dem Sinne nach anderen Veröffentlichungen entnommenen Stellen kenntlich gemacht.

Die Dissertation hat bisher weder in der gegenwärtigen noch in einer anderen Fassung weder dem Fachbereich Physik der Freien Universität Berlin noch einer anderen Fakultät oder Universität vorgelegt.

Berlin, den 27.01.2011

_____
(Peter Kupser)

# Zusammenfassung

Schwingungsspektroskopie ist eine gebräuchliche Methode zur Strukturaufklärung von Biomolekülen in der kondensierten Phase. Schwingungsmoden, wie zum Beispiel die C=O Streckschwingung (Amid-I Bande) im Bereich von 1700 cm$^{-1}$ oder die N−H Biegeschwingung (Amid-II Bande) bei 1500 cm$^{-1}$ sind empfindlich auf die dreidimensionale Anordnung der Atome in Proteinen. Die Kenntnis der Bandenposition kann es ermöglichen, ein umfassendes Bild der zugrundeliegenden Konformation des Moleküls zu erhalten. Untersuchungen an Biomolekülen in der Gasphase können einen Einblick in die intrinsischen strukturbestimmenden Wechselwirkungen des Moleküls liefern. Vielen Proteinen sind Sekundärstrukturelemente, wie zum Beispiel die $\alpha$-Helix oder das $\beta$-Faltblatt gemein und deren Motive bestimmen häufig die Struktur des ganzen Proteins. Im Rahmen dieser Arbeit wurden Biomoleküle in der Gasphase mittels Infrarotspektroskopie untersucht. Das Ziel war es, strukturcharakteristische Bandenpositionen im Infrarotspektrum, insbesondere die der Amid-I Bande, von verschiedenen Sekundärstrukturelementen zu finden. Hierzu wurden Methoden der Massenspektrometrie mit denen der Infrarotspektroskopie sowie der Dichtefunktionaltheorie verbunden.

Anhand von Untersuchungen an den Anionen der hochsymmetrischen Systeme C$_{60}$ und C$_{70}$ wird die Möglichkeit der Infrarotspektroskopie in Kombination mit theoretischen Rechnungen an Molekülen in der Gasphase zur Strukturaufklärung dargestellt. Cytochrom C, ein Protein mit vorwiegend $\alpha$-helikaler Sekundärstruktur in der kondensierten Phase, wurde in den Ladungszuständen 7+ bis 16+ untersucht. Eine Verschiebung der Amid-I Bande zu höheren Frequenzen mit zunehmendem Ladungszustand wird beobachtet. Die Sekundärstruktur des zweiten untersuchten Proteins, Tendamistat, besteht in der kondensierten Phase vorwiegend aus dem $\beta$-Faltblattmotiv. Als Modellsystem für ein $\alpha$-helikales Sekundärstrukturmotiv in der Gasphase dienten alaninbasierte Peptide, die eine helikale Struktur in der Gasphase ausbilden können. Das zyklische Peptid Gramicidin S wurde als ein mögliches Modellsystem für ein $\beta$-Faltblattmotiv in der Gasphase untersucht. Wechselwirkungen zwischen geladenen Seitenketten und dem Peptidrückgrat wurden durch Komplexierung mit Kronethermolekülen unterbunden. Messungen an dem Molekül VW01, welches so synthetisiert wurde, dass es in der kondensierten Phase als Dimerkomplex eine *coiled coil* Struktur ausbildet, wurden durchgeführt. Das Monomer des Moleküls hat eine *random coil* Struktur und wird nur in der Gasphase beobachtet.

Die Verschiebung der Amid-I Bandenposition der untersuchten Moleküle entspricht den Erwartungen: Mit zunehmender Stärke der Wasserstoffbrückenbindung verschiebt sich diese Bande zu niedrigeren Frequenzen. Die beobachtete Bandenposition variiert für die verschiedenen Moleküle, was auf unterschiedliche Strukturen zurückgeführt werden kann. Eine Unterscheidung anhand der gemessenen Amid-I Bandenposition zwischen der Struktur einer $\alpha$-Helix und der eines *random coil* stellte sich jedoch als schwierig heraus, was so aber auch in der kondensierten Phase gilt. Es zeigte sich, dass eventuell ein anderes System als Modell für eine $\beta$-Faltblatt Struktur in der Gasphase genutzt werden sollte als das zyklische Peptid Gramicidin S. Bei diesem kleinen, starren Molekül sind lediglich vier Wasserstoffbrücken in diesem Sekundärstrukturmotiv involviert. Es wurde weiterhin gezeigt, dass mit dem Dimer von VW01 ein nicht-kovalent gebundener Proteinkomplex intakt in die Gasphase transferiert werden kann. Die beobachtete Amid-I Bandenposition im Infrarotspektrum dieses Komplexes

ist die niederfrequenteste aller untersuchten Moleküle, und der Komplex besitzt damit das stärkste Wasserstoffbrückenbindungsnetzwerk. Im Allgemeinen ist zu erwähnen, dass eine Zuordnung einer gemessenen Bandenposition zu einer exakten Struktur im Moment noch sehr schwierig ist. Immerhin scheinen aber die relativen Amid-I Bandenpositionen für die verschiedenen Strukturelemente in der Gasphase denen der kondensierten Phase ähnlich zu sein. Des Weiteren findet sich eine gute Übereinstimmung im Vergleich der experimentell erhaltenen und mit Theorie bestimmten Spektren und es zeigt sich, dass nicht nur die Amid-I Bandenposition, sondern auch der niederfrequente Bereich im Spektrum Informationen zur Struktur der Moleküle liefern kann. Abschließend sei gesagt, dass strukturrelevante Informationen mittels Infrarotspektroskopie an Biomolekülen in der Gasphase erhalten werden können, aber die physikalischen und chemischen Bedingungen des zu untersuchenden Moleküls genau kontrollierbar sein sollten.

# Abstract

Vibrational spectroscopy is a commonly used method in structural investigations of biomolecules in the condensed phase. Vibrational modes, such as C=O stretching vibrations (amide-I) in the range of 1700 $cm^{-1}$ or N$-$H bending vibrations (amide-II) at around 1500 $cm^{-1}$, are sensitive to the three-dimensional arrangement of the atoms in the proteins. Knowledge of the band positions makes it possible to obtain a global picture of the fundamental conformation of the molecule. Investigations on gas-phase biomolecules allow for an insight into the intrinsic intra-molecular interactions, which determine the molecular structure. Many proteins possess secondary structural elements, including for instance $\alpha$-helix and $\beta$-sheet, and these motifs often determine the structure of the entire protein. For this thesis biomolecules in the gas phase were investigated using infrared spectroscopy, the goal being to define band positions of significant structural characteristics in the infrared spectrum. The amide-I band position for the various secondary structural elements was an area of particular research interest. In the pursuit of these research aims mass spectrometry, infrared spectroscopy, and density functional theory methodologies were employed.

The effectiveness of infrared spectroscopy in combination with theoretical calculations in determining molecular structure of molecules in the gas phase was confirmed in this thesis using the highly symmetrical systems $C_{60}^-$ und $C_{70}^-$. Cytochrome c, a protein with a predominantly $\alpha$-helical secondary structure in the condensed phase, was investigated in the charge states 7+ to 16+. A shift of the amide-I band towards higher frequencies was observed with increasing charge state. The secondary structure of the second investigated protein, tendamistat, is characterized in the condensed phase by a $\beta$-sheet motif. As a model system for $\alpha$-helical secondary structures in the gas phase, alanine based peptides were used due to their demonstrated helical structure in the gas phase. The cyclic peptide gramicidin S was used as a possible model system for a $\beta$-sheet motif in the gas phase. Interactions between charged side chains and the backbone of the peptide were suppressed through complexation with crown ether molecules. Measurements were taken on the molecule VW01, which was synthesized so that in the condensed phase the dimer possesses a *coiled coil* structure. The monomer, which is not observed in the condensed phase, exhibits a *random coil* structure.

The observed shifts in the amide-I band position of the investigated molecules were as expected: With increasing hydrogen bonding strength the band shifts towards lower frequencies. The observed band positions varied for the different molecules, which could be due to their differing structures. It is difficult to differentiate between amide-I band positions for molecules with either an $\alpha$-helical or *random coil* structure in the gas phase, which however also applies for the condensed phase. It became evident that it may be beneficial to use a system other than gramicidin S as a model for the $\beta$-sheet structure in the gas phase. These small, rigid molecules possess namely only four hydrogen bonds in this secondary structural motif. In regards to the dimer of VW01, it was shown that a non-covalently bound protein complex can be transferred intact into the gas phase. The amide-I band position observed for this complex was the lowest of all investigated molecules, leading to the conclusion that it possesses the strongest hydrogen bonding network. In general this research indicates that the definition of band position to determine molecular structure is still problematic. Nevertheless it was suggested that the

relative amide-I band positions for the various structural elements in the gas phase were similar to those previously observed in the condensed phase. Additionally there was good agreement between experimentally observed and by theory calculated spectra and it can be inferred from the results that not only the position of the amide-I band but also the low frequency range in the spectrum can provide further information about molecular structure. In conclusion information relevant to the structure of gas-phase biomolecules can be obtained through the use of infrared spectroscopy but it is critical that the physical and chemical conditions in which the molecule is investigated are stringently controlled.

# Contents

# Chapter 1

# Introduction

In this Chapter, an introduction to the molecular systems investigated and to the experimental methods and techniques used in this thesis is given.

## 1.1 Peptides and proteins

Proteins are a group of macromolecules containing up to thousands of atoms and play a vital role in living organisms. They are directly involved in the chemical processes essential for life, are parts of all different kinds of cells, many biological materials are formed mainly by them such as for example hair, and they play an important role in the biochemical signal pathway and many other biochemical processes. The conformation, or three-dimensional structure, of a protein is correlated to the physiological function of the molecule. It has been discovered that misfolded proteins, where they adapt a different structure compared to the native one, are responsible for diseases such as Parkinson, Alzheimer, or Creutzfeldt Jakob. Thus, an understanding of the structure and of the interactions involved in the folding of those molecules is of importance and a subject of various investigations.

Proteins are polymeric chains formed by a combination of 20 naturally occurring L-$\alpha$-amino acids and a protein is characterized by the sequence of these building blocks. The amino acids consist of a central carbon atom ($C_a$), to which a hydrogen atom, an amino group ($NH_2$), a carboxyl group (COOH), and a characteristic side chain is attached. The molecular structure together with the three and one letter codes of all 20 natural amino acids are shown in Figure 1.1. The amino acids grouped on the left hand side possess hydrophobic side chains and those on the right hydrophilic side chains. The amino acids are connected by peptide bonds (see Figure 1.2). When a small number (2 – 50) are connected, the resulting molecule is designated as a peptide whereas it is referred to as a protein when the number of amino acids is larger. The residue with the free amino group (by definition the furthest left residue) is denoted as the aminoterminus or N-terminus and the residue with the free carboxyl group is denoted as the carboxyterminus or C-terminus. The repeating units of the peptide bonds ($-N-C_\alpha-C(O)-)_n$ in the polymer peptide chain (with $n$ units) is denoted as the main chain or backbone whereby the side chains of the amino acids are neglected.
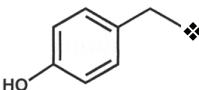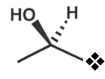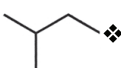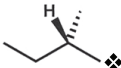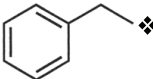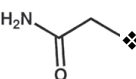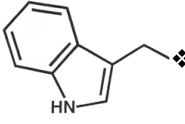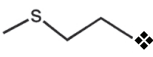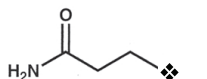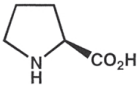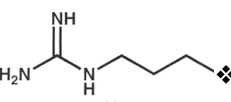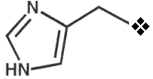
**Figure 1.1:** A list of all 20 naturally occurring amino acids together with the three letter code and one letter code. On the left, amino acids with hydrophobic side chains are grouped and on the right, those with hydrophilic ones. The Figure is taken from [1], Table 1.1 therein. ❖ denotes the main chain.

The term 'protein' was first mentioned by Jöns Jacob Berzelius in the 19th century [2] to describe molecules containing one or more polypeptide chains. The primary structure (amino acid sequence) of a protein is defined by the order of the amino acids in a polypeptide chain going from the N-terminus to the C-terminus. The local spatial arrangement of the backbone is denoted as the secondary structure of the protein and various common motifs are found in many biomolecules which are for example $\alpha$-helices, $\beta$-sheets, and turns; these structures are discussed in further detail below. The tertiary structure of a protein is defined as the higher ordered spatial arrangement of all
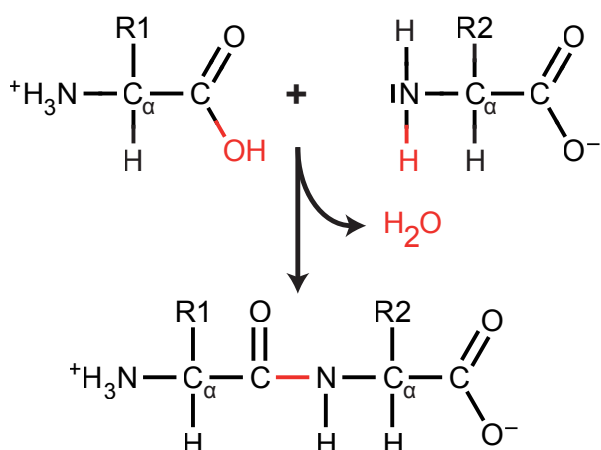
**Figure 1.2:** Two amino acids containing a side chain, R1 and R2 respectively are bound together by eliminating a water molecule and form a dipeptide. The peptide bond is depicted in red.

atoms in one peptide chain (subunit). Interactions including those of the side chains of the amino acids as well as those with adjacent segments of the main chain form this structural element. If there exists more than one subunit in the protein, the spatial arrangement of these peptide chains is described by the quaternary structure. The conformation of a protein is the combination of the secondary, tertiary, and quaternary structure.

Different methods are available for obtaining information about the structure of proteins. Some of those are X-ray crystallography [3], nuclear magnetic resonance (NMR) [4], infrared spectroscopy (IR) [5], and circular dichroism (CD) [6]. The latter three methods can be used to obtain structural information of the protein in the natural environment, the condensed phase.

The structures of peptides and proteins can be visualized in various ways. Pauling developed in 1953 a wooden space filling model representing atoms or groups of atoms where 1 inch in the model corresponds to 1 Å in the molecule [7]. The use of certain colors for different atoms (nowadays known as CPK color) was introduced at this time. With the invention of computer technology, illustration of protein structure became more sophisticated. By using the van der Waals radii of the atoms, space filling representations can be produced. To highlight only the secondary structure of a protein (the backbone), different graphical methods exist. In the New Cartoon representation, a variant of a ribbon representation, a $\beta$-sheet motif, for example, is depicted by an arrow (which points in the direction of the peptide strand) and an $\alpha$-helical motif by a coil. In Figure 1.3, examples of different representations of the structure of the green fluorescent protein (GFP, protein data base entry # 2HGD) [8] obtained by X-ray crystallography are shown. The representation methods and the color codes are: (A) line representation (color code: atom types; H → white, O → red, N → blue, C → cyan, S → yellow, P → brown), (B) CPK (color code: atom types), (C) van der Waals (color code: atom types), (D) bonds (color code: residue type; basic → blue, acidic → red, polar → green, non polar → white, unassigned → cyan), (E) surface (color code: residue type), (F) New Cartoon (color code: secondary structure; $\alpha$-helix → red, $3_{10}$-helix → orange, $\beta$-sheet → yellow, turn → blue, random coil → white).
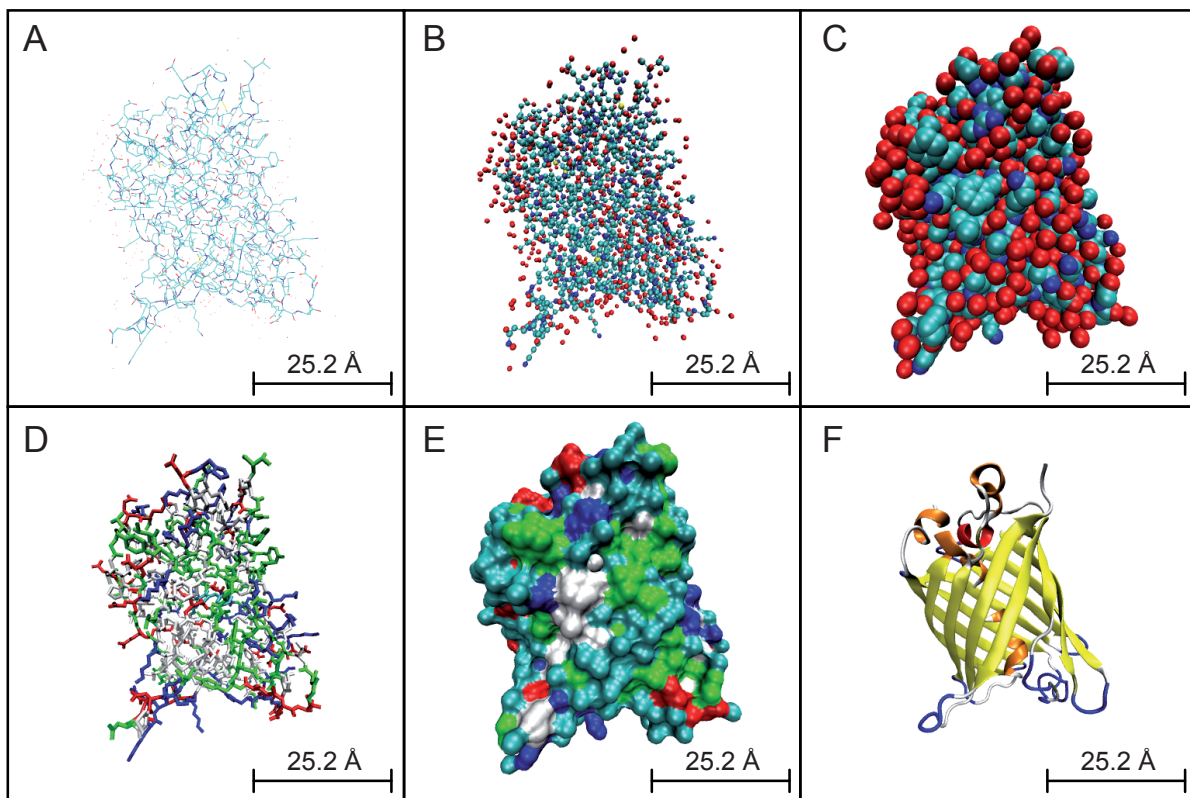
**Figure 1.3:** Different representations of the structure of the GFP molecule. A length scale is given on the bottom of the corresponding images. See text for the representation method and the corresponding color code.

The folding and stability of proteins is influenced by the physical and chemical properties of the amino acids and their interactions among each other as well as with the environment. Bond stretching and bending, repulsive (static) interactions, electrostatic interactions, and van der Waals (dispersive) interactions play a key role in the formation of the conformation.

In the condensed phase, hydrophobic interactions are considered to play a major role in the folding of the native structure of the macromolecules [9]. Non polar side chains aggregate inside the molecule such that their contact to surrounding water molecules is minimal. This arrangement leads to a gain in entropy of the water molecules because otherwise, they would naturally build ordered structures around the hydrophobic residues.

Hydrogen bonding networks between, for example, carboxyl (C=O) and amide (N−H) groups contribute to the stability of protein structures and define the secondary structure elements of the macromolecule [10].

Salt bridges or ion pairs between oppositely charged amino acid side chains are often present inside proteins, however their structural stabilizing role is of minor importance in the condensed phase. On the other hand in the gas phase those interactions

can contribute heavily to the formation of the protein conformation.

Disulfide bonds between thiol groups (S−H) of cysteine residues in one or more polypeptide chains contribute to stabilizing structural motifs of the backbone by cross linking. Further, intramolecular cross links of residues such as cysteine or histidine to metal ions support the formation of stable conformations.

Many textbooks of biochemistry discuss the forces and interactions influencing protein folding in more detail (for example [11]), or see review articles [10, 12–15].

The peptide bond between two amino acids provides a rigid plane due to a partial double bond character of the N−C peptidic bond which can be either in the cis or trans[1] conformation (see Figure 1.4). However, in most peptide groups the bond is found to be in the trans configuration. The backbone conformation can be described by the two dihedral angles $\phi$ (at the N−$C_\alpha$ bond) and $\psi$ (at the $C_\alpha$−C bond) between the peptide groups. Due to the steric hindrance of adjacent peptide groups not all possible angles are allowed. That gives rise to several structural motifs of the protein backbone (secondary structure elements) that are found in many proteins.
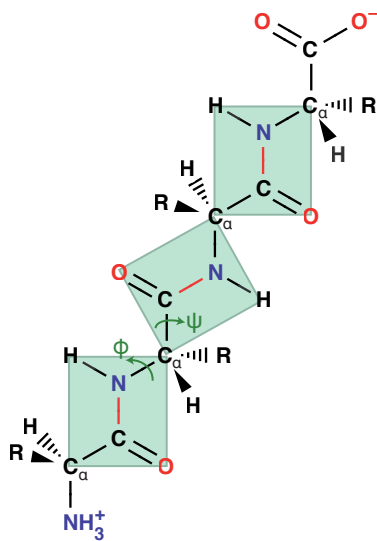


**Figure 1.4:** Shown are four amino acids in a chain. The peptide bond is depicted in red, the rigid planar peptide group plane is presented in light green and the dihedral angles $\phi$ and $\psi$ are shown in dark green.

Based on those angles and the planar nature of the peptide bond Linus Pauling proposed in 1951 the $\alpha$-helix [16] and pleated sheets [17] as possible configurations of polypeptide chains (see Figure 1.5).

The $\alpha$-helix is usually right handed and contains 3.6 amino acids per turn with a pitch of 5.4 Å. The dihedral angles $\phi$ and $\psi$ are both approximately -60°. A hydrogen bond between the C=O group of the $i$th residue and the N−H group of the $i + 4$th residue stabilize this motif (see the red line in Figure 1.6) and the bonding length is $\approx 2.8$ Å. To minimize steric hindrance, the side chains of the residues are pointing away from the helical axis in the direction of the N-terminus.

---

[1]both adjacent $C_\alpha$ atoms are in opposite corners of the plane, see Figure 1.4.
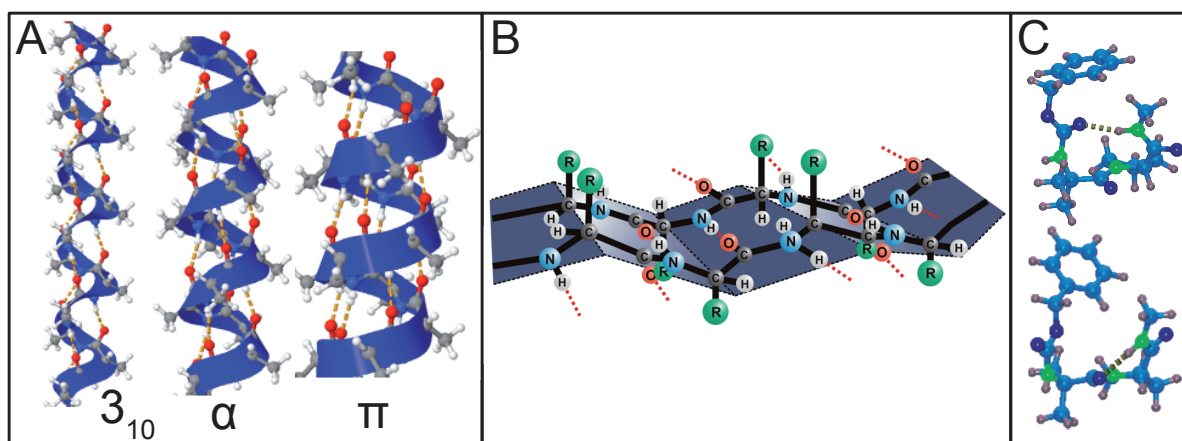
**Figure 1.5:** Different secondary structure elements: (A) helices [18], $\alpha$-helix with a repeating pattern of 3.6 amino acids (spaced by 5.4 Å), (B) antiparallel $\beta$-sheet [19] with a repeating pattern of two amino acids, (C) $\beta$-turn (top) and $\gamma$-turn (bottom) [20]
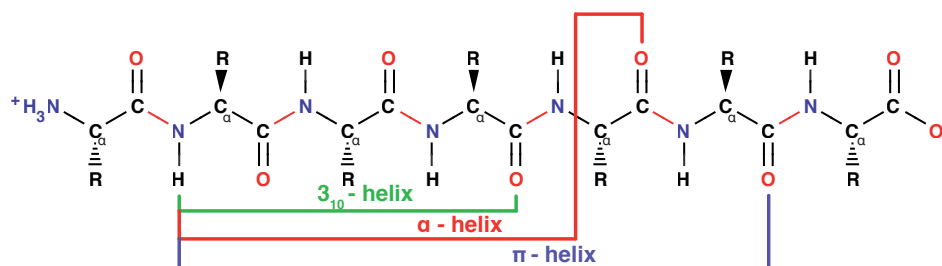


**Figure 1.6:** Polypeptide chain with the hydrogen bonding partners depicting a $3_{10}$-helix in green (between $i$th and $i + 3$th amino acid, ten atoms between the bonding partner leading to a $C_{10}$ ring), an $\alpha$-helix in red (see text, $C_{13}$ ring), and a $\pi$-helix in blue (between $i$th and $i + 5$th amino acid, $C_{16}$ ring).

Each peptide unit contains a dipole moment[2] of 3.46 Debye pointing along the helical axis [21] and as a consequence, a macro-dipole moment is established in the $\alpha$-helical motif, pointing from the C-terminus to the N-terminus (strength: sum of dipole moments of all participating amino acids).

Naturally, proteins occur in aqueous solution or lipid membranes and the stability of the helical conformation depends on the solvent [22, 23]. Spectroscopic investigations on $\alpha$-helical model systems are presented in Chapter 4.

The second very common motif found in the secondary structure of proteins is the $\beta$-sheet. This motif can be found between different adjacent polypeptide chains (in contrast to the $\alpha$-helix) in the quaternary structure or within one strand. The appearance of this structural element is a pleated sheet with the amino acid side chains alternating above and below the peptide group plane and a repetition unit of two residues spaced by $\approx 7$ Å. The dihedral angles $\phi$ and $\psi$ are between -120° and -150° and +120° and +150°, respectively. When the polypeptide strands are oriented in the same direction the motif

---

[2]the N–H and C=O groups have a different polarity

is known as a parallel $\beta$-sheet (see Figure 1.7), upon orientation in the opposite direction it is called an antiparallel $\beta$-sheet (see Figure 1.8). The hydrogen bonds slightly differ in
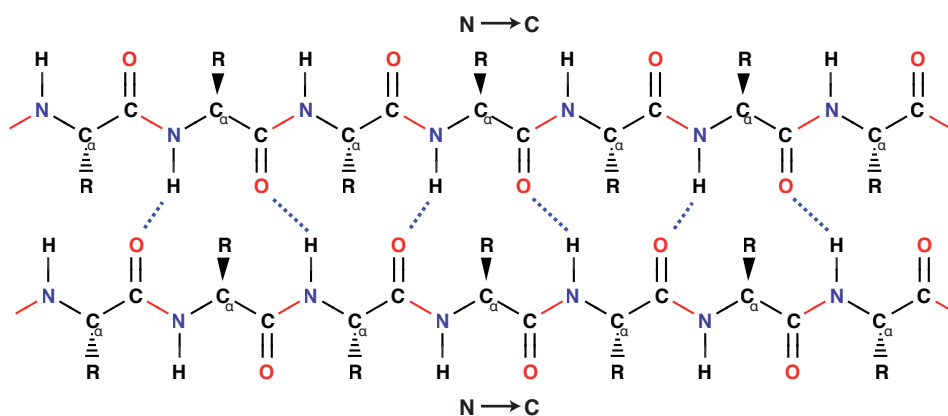


**Figure 1.7:** Parallel $\beta$-sheet motif. Hydrogen bonds are indicated by dotted lines.
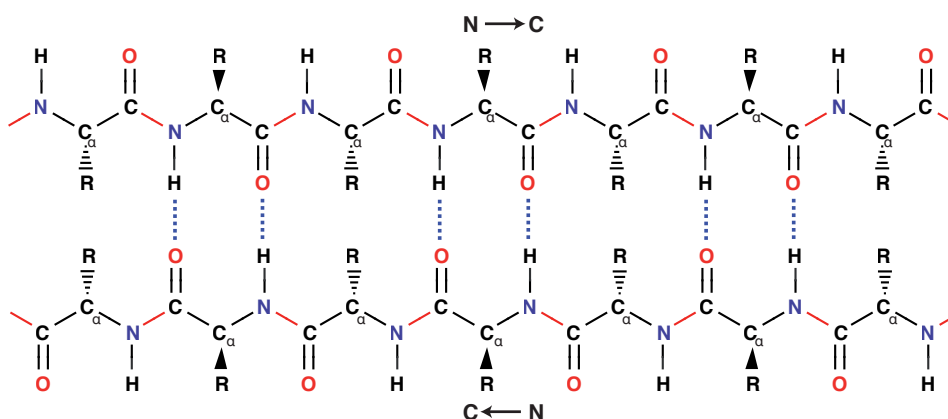


**Figure 1.8:** Antiparallel $\beta$-sheet motif. Hydrogen bonds are indicated by dotted lines.

length and orientation between the parallel and antiparallel $\beta$-sheet motif (see Figure 1.7 and 1.8). The formation of an antiparallel sheet motif in a single strand can be very short and requires only a turn while the formation of a parallel motif requieres a long intermediate strand.

Infrared spectroscopy on a potential antiparallel $\beta$-sheet model system is performed and presented in Chapter 5 of this thesis.

Another frequently occurring structural motif in proteins is the leucine zipper or *coiled coil* arrangement, first demonstrated by Crick in the 1950s [24]. Two or more $\alpha$-helices are wound around each other, thus stabilizing a structural motif and slightly reducing the repetition pattern per turn from 3.6 residues for an $\alpha$-helix to 3.5 residues (if only two helices are involved). A sketch of a *coiled coil* structure is shown in Figure 1.9.

Further discussion and spectroscopical investigations regarding coiled peptides are presented in Chapter 6.
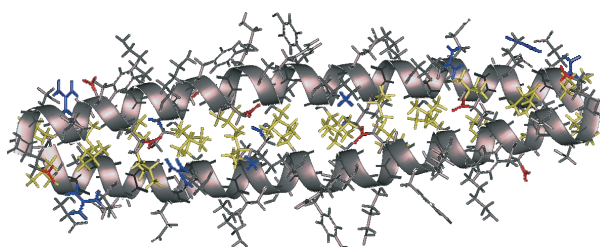
**Figure 1.9:** Sketch of a *coiled coil* structural motif [19].

## 1.2 Peptides and proteins in the gas phase

The gas phase provides a unique environment for studying structures of isolated molecules in which the conditions of the investigated systems such as the charge states and the masses can be controlled very accurately. The absence of solvent molecules allows for the study of intrinsic properties of the molecule. Through investigations on weakly bound systems, such as oligomers of peptides, intermolecular interactions can be probed. Due to the absence of interactions with solvent molecules, the observed structures can be different. However, it has been shown that solution conformations of proteins can be, at least partially, conserved in the gas phase [25–27].

In recent years, soft ionization methods to bring large non-volatile molecular systems intact into the gas phase have been developed; these are widely used in mass spectrometry nowadays. One of these techniques is the electrospray ionization (ESI), developed 1984 by John B. Fenn [29] which is based on experiments performed by S. Chapman [30] in 1937 and by M. Dole [31] in 1968. Fenn was able to produce highly charged biomolecules with mass up to 130 kDa[3] [32] and was awarded the Nobel prize for his work in 2002 [33]. The principles as employed in this thesis are the following: (bio)molecules in solution are sprayed through a capillary (diameter around 125 $\mu$m) to which a high voltage is applied (around 3 kV). The polarity of the applied voltage defines the polarity of the generated ions. Desolvation and nebulizing gas (typically $N_2$, sometimes heated up to around 120° C) helps to stabilize the spray and supports the evaporation of the solvent molecules in the charged droplets. The charge is located at the surface of the droplets and due to the evaporation of solvent molecules, the droplets shrink until the Rayleigh limit is reached where the Coulomb repulsion exceeds the surface tension and the droplets explode (charge $q$ and surface tension $\gamma$ of the droplets are connected to their radius $R$ and the dielectric constant $\epsilon_0$ by $q = 8\pi \sqrt{\epsilon_o \gamma R^3}$ [34]). This is called Coulomb fission. The radius of the split-off droplets is about 1/10th with about 2% mass and 15% charge of the original droplet [35]. The process repeats in cascades and as a result, highly charged single (bio)molecules are produced. However, the last steps to the production of the single charged (bio)molecules is not yet clear. At least two different mechanisms are proposed: first, the charge residue model (CRM) in which individual charged molecules are produced by Coulomb fission[4]. The second is the ion evaporation model (IEM), where the emission of single ions out of the surface of the droplets prior to reaching the Rayleigh limit is suggested. A combination of both effects seems to be plausible but a more detailed discussion of the mechanisms of the

---

[3]1 Da $\hat{=}$ 1 amu
[4]The described process repeats until only one charged molecule is left.

electrospray process can be found in a recent review article by Kebarle [36]. Usually, the ions produced by the electrospray process are highly charged and a broad charge distribution is observed in the mass spectrometer.

One more soft ionization method is mentioned here: the matrix assisted laser desorption/ionization (MALDI) technique developed by K. Tanaka as well as by F. Hillenkamp and M. Karas [37, 38]. (Bio)molecules are embedded in an acidic crystallized matrix capable of absorbing photons of a certain energy (mostly ultraviolet). The absorption of laser light by the matrix molecules leads to a transfer of charges from the matrix to the embedded analyte molecule and to a desorption of the molecules. With the MALDI method it is possible to produce lowly (for example singly) charged gas-phase molecules.

Investigations of molecule structures in the gas phase can provide useful benchmarks for theoretical models and give relevant insight into the physics of structural motifs. The absence of intermolecular interactions with solvent molecules allow for the use of higher level theory: The calculation time of energy optimized molecular structures (with respect to the potential energy surface PES) obtained by density functional theory (DFT) scales normally with $O(N^3)$ with the number of atoms $N$ in the system, but can be optimized to scale linear with $O(N)$. It is possible to include dispersion interactions in the DFT calculations without increasing the calculation time. To obtain infrared spectra of a stable structure found at a minimum of the PES, second derivatives (the Hessian matrix) are formed and the vibrational eigenfunctions and eigenvectors are calculated (by diagonalizing the Hessian matrix) to give the normal vibrations. This technique is used to obtain infrared spectra of fullerenes (see Chapter 2) and gramicidin S (see Chapter 5). Another method of obtaining theoretical infrared spectra is available, referred to as ab initio molecular dynamics (AIMD) and used to calculate systems presented in Chapter 4. A comprehensive introduction to this method is given in the PhD thesis of M. Rossi [28] and is beyond the scope of this work.

## 1.3   Infrared spectroscopy of biomolecules

Vibrational modes in molecules can be excited upon absorption of photons in the infrared (IR) range of the electromagnetic spectrum (2.5 $\mu$m to 250 $\mu$m). Observed IR absorption bands in molecular samples can allow access to information on the nature of present vibrational modes, chemical composition, bonding lengths and strengths, as well as the angles of chemical groups, existing hydrogen bonding networks, and the conformation of the molecule.

A non-linear molecule with $N$ atoms has $3N - 6$ internal degrees of freedom. In the harmonic approximation (abortion of the Taylor expansion of the potential after the quadratic term) this corresponds to $3N - 6$ independent normal modes of vibration[5] and the vibrational energy levels are equidistantly spaced and $E_{vib,i} = h\nu_i(n + 1/2)$ with the Planck's constant being $h$, the vibrational frequency $\nu_i$, and the vibrational quantum number $n$. The vibrational frequencies are properties of the molecular potential (defined by the bonding strength of the chemical groups, the bonding distance, angles, etc.). A

---

[5]Normal mode of vibration: All atoms in the molecule vibrate with the same frequency and phase. The molecule's center of gravity does not move, nor does the molecule rotate.

mode is infrared active when the dipole moment undergoes a change upon vibrational excitation. The intensity of the absorption band is proportional to the square of the change in dipole moment $\partial\mu$ divided by the change in the normal coordinate $\partial Q$. Selection rules obtained in the harmonic approximation by perturbation theory for allowed transitions are $\Delta v = \pm 1$. However, a more accurate description of vibrational modes in real molecules is given by the anharmonic approximation of the vibrational potential. The empirical potential for a diatomic molecule can be described by a Morse function $E_{pot}(r) = D_{min}e^{-2a(r-r_0)} - 2De^{-a(r-r_0)}$ [39] with the equilibrium bond distance $r_0$, the depth of the potential $D$, and $a = \sqrt{k/(2D)}$ ($k$ being the force constant). The spacing between the vibrational modes is then no longer equidistant, and with raising vibrational quantum number it becomes smaller. The vibrational energy levels are given by $E_{vib} = hv\left[(n + 1/2) - x_{anharm}(n + 1/2)^2\right]$ with the anharmonicity constant $x_{anharm}$. The selection rules obtained for allowed vibrational transitions are $\Delta v = \pm 1, \pm 2, \pm 3, \ldots$ For molecules containing more than two atoms, the determination of the vibrational energy levels is more complicated and for example mode coupling can occur. An excellent and detailed introduction to infrared spectroscopy is given in Colthup's book [40].

There are different types of vibration:

- Stretching vibrations (antisymmetric and symmetric): change in the bonding length

- Bending/deformation vibrations (symmetric bending, antisymmetric rocking): change of the bonding angle (within the plane it shares with another bond) while the distance between the atoms is constant

- Wagging vibrations: change of the bonding angle (outside the plane it shares with another bond)

- Twisting vibrations: change of the dihedral angle of a bond

In infrared spectroscopy it is common to use the wavenumber $\tilde{v}$ – which is directly proportional to the energy ($E = hv$) – instead of the wavelength $\lambda$. $\tilde{v}$ is defined as $\tilde{v} = 1/\lambda$ and the unit is usually given in $cm^{-1}$. Around 3000 $cm^{-1}$, X–H stretching vibrations, with X being an atom such as C, occur and below $\approx 2000\ cm^{-1}$, other modes are present. The mid-infrared region between 500 $cm^{-1}$ and 2000 $cm^{-1}$ is called the "fingerprint region". In polypeptide chains various vibrational modes of backbone groups are present:

- Amide-A and -B($\approx 3300\ cm^{-1}$ and $\approx 3170\ cm^{-1}$) modes: The exclusively localized N–H stretching vibration is called the amide-A mode and the band position is insensitive to the conformation of the backbone. Instead, the frequency depends on the strength of the hydrogen bond. The amide-A and amide-B modes correspond to a Fermi resonance doublet. The N–H stretching vibration is resonant with an overtone of the amide-II vibration.

- Amide-I mode ($\approx 1650\ cm^{-1}$): These modes arise mainly from C=O stretching vibrations that are coupled to the out-of-phase C–N stretching vibration, the C–C–N deformation, and the in-plane N–H bending vibration. The frequency

of this vibrational band is highly sensitive to the secondary structure and the hydrogen bonding pattern in the backbone of a polypeptide chain.

- Amide-II mode ($\approx 1550$ cm$^{-1}$): Out-of-phase combination of the in-plane N−H bending vibration and smaller contributions from both the in-plane C=O bending and C−C and N−C stretching vibrations. The correlation of the amide-II band position to the backbone conformation is less straightforward than that of the amide-I band position.

- Amide-II′ mode ($\approx 1450$ cm$^{-1}$): Upon deuteration, the amide-II band changes its character. The signature around 1430 cm$^{-1}$ and 1450 cm$^{-1}$ is commonly called amide-II′ band and is characterized by an increased contribution of C−N stretching vibration in contrast to the amide-II band.

- Amide-III mode ($\approx 1200$ cm$^{-1}$ to $\approx 1400$ cm$^{-1}$): In-phase combination of N−H bending and C−N stretching vibration with minor contributions from in-plane C=O bending and C−C stretching vibration.

The position of the amide-I and amide-II vibrations in the infrared spectrum depends on the conformation of the backbone. When the backbone C=O groups are hydrogen bonded with the backbone N−H groups, shifts of the band positions are observed. With increasing hydrogen bonding strength, the amide-I band position is red-shifted and the amide-II band position is blue-shifted. The bands thus shift in opposite directions. The influence of the anharmonicity of a vibrational mode for the observed spectrum is discussed in Section 1.4.1.

Table 1.1 gives an overview of measured amide-I band positions for different secondary structure elements as observed for proteins in solution ($H_2O$ and $D_2O$).

**Table 1.1:** Experimental values for measured amide-I band positions for different secondary structure elements of proteins in the condensed phase, adapted from a review article by Barth [5], Table 2 therein. The value for the *coiled coil* structure in $D_2O$ is measured by Krimm [41].

| Motif | Band position in $H_2O$ [cm$^{-1}$] | | Band position in $D_2O$ [cm$^{-1}$] | |
|---|---|---|---|---|
| | Average | Extremes | Average | Extremes |
| $\alpha$-helix | 1654 | [1648-1657] | 1652 | [1642-1660] |
| $\beta$-sheet | 1633 | [1623-1641] | 1630 | [1615-1638] |
| | 1684 | [1674-1695] | 1679 | [1672-1694] |
| turns | 1672 | [1662-1686] | 1671 | [1653-1691] |
| | | [1630-1700] | | [1630-1700] |
| random coil | 1654 | [1642-1657] | 1645 | [1639-1654] |
| *coiled coil* | N/A | | 1630 | |

An extensive introduction to vibrational spectroscopy of proteins in solution is given by Barth *et al.* [5, 42] and of gas-phase ionic complexes of biological relevance by Polfer *et al.* [43].

## 1.4   Infrared spectroscopy using free electron lasers

Vibrational spectroscopy together with quantum chemical calculations provide an experimental and theoretical approach to investigating structures of charged and neutral gas-phase molecules [44, 45]. Due to the low density of molecules in the gas phase (for ions: $\leq 10^8$ ions/cm$^3$), direct absorption methods, where the influence of sample molecules to the transmitted light is measured, cannot be applied. However, a technique referred to as 'action spectroscopy' is used, where the influence of the light onto the sample molecules is recorded (see Section 1.4.2). To perform action spectroscopy, a bright and tunable light source is needed and modern table top lasers as well as free electron lasers are available with suitable characteristics.

First spectroscopic measurements on a neutral gas-phase biomolecule, tryptophan, were performed using a supersonic, cold molecular beam and a two-photon ionization scheme by Rizzo *et al.* in 1986 [46]. First experiments to probe local interactions in isolated molecules that stabilize folding were performed in 2006 by de Vries *et al.* [47]. They used double-resonance IR-UV spectroscopy on larger biomolecules, the class of gramicidin molecules, which are brought into the gas phase by laser desorption methods and investigated spectroscopically the N−H stretching region in the infrared spectrum. More recently, highly resolved IR-UV double resonance spectroscopy in this range was performed on gramicidin S ions in a 6 K cold, 22-pole ion trap in the group of Rizzo *et al.* [48]. Mid-infrared spectroscopy using table top laser systems and an infrared resonant two-photon ionisation scheme allows for detailed structural insights, for example into the nature of a $\beta$-sheet structure as performed by Gerhards *et al.* [49]. IR-UV depletion spectroscopy on the DNA base guanine allowed for providing evidence of a existence of a new tautomeric form of the molecule in the gas phase as shown in the work by Mons *et al.* [50].

The first infrared spectroscopic measurements using a free electron laser as a light source in combination with mass spectrometry were performed in 1996 by Meijer *et al.* [51] in the Netherlands. Many experiments on various other systems in the gas phase were performed using the free electron laser in the Netherlands, these included: fullerenes such as $C_{60}$ [52], metal clusters [53], vanadium oxide clusters [54], and amino acids [55]. The coupling of a Fourier transform ion cyclotron mass spectrometer (FT ICR) to a free electron laser, allowing for the combination of several mass selection and ion reaction steps in one apparatus, was performed in the group of Maitre in Orsay, France (FEL CLIO) [56] and in the group of von Helden in the Netherlands (FELIX) [57]. The working principle of a FT ICR mass spectrometer is presented in Section 1.4.4.

To perform infrared spectroscopy in the fingerprint region on big biomolecular systems in the gas phase, some special features are required. Firstly, coherent light in the range from 500 cm$^{-1}$ to 2000 cm$^{-1}$ (corresponding to 20 $\mu$m and 5 $\mu$m respectively) must be available. Secondly, the intensity of the laser must be high enough to allow the absorption of many photons in order to dissociate the molecule for performing action spectroscopy by the IR-MPD process (see next Section). Free electron lasers satisfy these conditions. Their working principles are briefly described in Section 1.4.3, particularly

for the Free Electron Laser for Infrared eXperiments (FELIX), which was used to perform the experiments in this work.

### 1.4.1 The infrared resonance enhanced multiple photon dissociation process (IR-MPD)

The resonant absorption of multiple infrared photons can lead to fragmentation of molecular systems. This process is called infrared resonance enhanced multiple photon dissociation (IR-MPD) and is used to obtain the infrared spectra of the systems investigated in this work. A detailed description of the principles of IR-MPD on gas-phase ions is given by review articles of von Helden and Oomens [43, 58]. A brief introduction of this process is presented here.

The mechanism of IR-MPD can be divided into effects occurring in different energy regions, which are depicted in Figure 1.10:

- Region I is characterized by *coherent interactions*, in which the molecule is excited by resonant absorption of photons over discrete states into the quasi-continuum region. Photons are absorbed in a vibrational ladder in which the spacing of the energy levels is not equidistant and becomes smaller with increasing energy due to the anharmonicity of the vibrational mode. As a consequence, the photon is non-resonant at higher energies of the excited molecule and can no longer be absorbed (anharmonic bottleneck). But often even with the absorption of fewer photons region II is reached in the energy level diagram.

- Region II is characterized by *incoherent* or *anharmonic interactions*, in which the energy of the sequentially absorbed photons is transferred due to anharmonic coupling of the vibrational modes into a quasi-continuum of vibrational modes. This process is called intramolecular vibrational redistribution (IVR). If a high density of states is given (i.e. if the molecule is warm and big enough), the population of the excited state is rapidly distributed via IVR and the molecule can absorb another photon. The absorption lines are broadened due to the short lifetime of the vibrational modes induced by IVR and are less than 1 ns at room temperature for aromatic molecules [58]. For the here investigated systems, it can be safely assumed that the time scales are at least that short. The absorption energy is red-shifted, with increasing internal energy as a consequence of decreased level spacing due to the anharmonicity.

- Region III is the energy range above the dissociation limit of the molecule and therefore the 'real' continuum of levels, leading to the dissociation of the molecule.

### 1.4.2 Action spectroscopy

A technique called vibrational pre-dissociation spectroscopy was developed in 1985 by Lee *et al.* [60]. Ions, produced via expansion and by electron impact, were mass selected and stored in a radio frequency (RF) ion trap. Infrared radiation enters the trap and the contents are then investigated in a quadrupole mass spectrometer. Upon resonant

**Figure 1.10:** Shown is a schematic energy-level diagram for the multiple photon dissociation of $SF_6$. The Roman numerals indicate energy regions in the diagram (see text for details) and $W_{diss}$ is the dissociation energy or the molecule. Reprinted with permissions from [59], Figure 1 therein. Copyright (2011) by the American Physical Society.
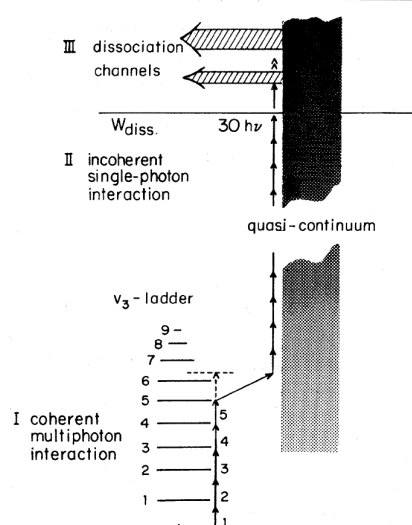
absorption of infrared photons by the ions, photodissociation occurs and fragment masses are observed. An infrared spectrum is recorded by repeating the described steps with varying wavelength of the laser. This new technique has opened up a number of possibilities to record infrared spectra of gas-phase ions. To appreciate the differences of this approach with conventional absorption spectroscopy, some details of this are given.

In direct absorption measurements, light is sent through a sample volume and the intensity of the transmitted light is recorded as a function of the wavelength. If sample molecules absorb light, the measured intensity is decreased. According to the Lambert-Beer law the absorption $A$ is given by:

$$A = \ln \frac{I}{I_0} = -\sigma(\tilde{v})n \cdot l \tag{1.1}$$

The intensity of the transmitted light being $I$, the intensity of the incoming light $I_0$, the absorption cross section $\sigma(\tilde{v})$, the density of sample molecules $n$, and the path-length $l$ of the light in the sample. With typical absorption cross sections, many sample molecules-per-cm$^2$ ($n \cdot l$) are needed and densities attainable for gas phase species are usually not sufficient. Further, this technique is not species specific.

An alternative to obtain a spectrum of low number density samples is to measure the influence of the light onto the sample by optical pumping. An intense light source is required and this technique is referred to as 'action spectroscopy'. The absorption of one or many photons of the sample molecule can, for example, lead to change in mass (due to fragmentation), change in charge (due to dissociation of charge carriers), or change in quantum state. To obtain infrared spectra of the molecules investigated in the present work, the change of mass by fragmentation of the molecules due to the resonant absorption of infrared light is recorded in a FT ICR mass spectrometer:

$$AB^+ + n \cdot h\nu \longrightarrow A^+ + B \tag{1.2}$$

with the parent ion $AB^+$, the Planck constant $h$, the frequency of the photon $\nu$, and the fragments $A^+$ and B respectively.

In a two level system whereby the lower level corresponds to the population of the parent ions $(AB^+)$ being $N_1$, the upper level corresponding to the population of the observed fragment ions $(A^+)$ being $N_2$, and the total population of $AB^+ + A^+$ being $N$, the population of the levels can be described by:

$$N_1(\tilde{\nu}) = Ne^{-\sigma(\tilde{\nu})\cdot F} \tag{1.3}$$

$$N_2(\tilde{\nu}) = N\left(1 - e^{-\sigma(\tilde{\nu})\cdot F}\right) \tag{1.4}$$

The fluence of the photons being $F$ (photons-per-cm$^2$) and the absorption cross section $\sigma(\tilde{\nu})$ (in [cm$^2$]). As can be seen, only a high fluence of the photons is needed to obtain $\sigma(\tilde{\nu})$ (given sufficient detection statistics for the ions) in contrast to the Lambert-Beer law (see Equation 1.1) where the sample density must be high but the fluence can be low.

An infrared spectrum for a given molecule is obtained by plotting the absorption cross section $\sigma(\tilde{\nu})$ as a function of the wavenumber $\tilde{\nu}$. The process is more complicated for multiple photon absorption but valid in first approximation as has been demonstrated with measurements on cationic naphthalene by Oomens *et al.* [61].

To determine $N$ if no fragment ions are measured, alternating recordings of mass spectra with $(N_1(\tilde{\nu}))$ and without infrared photons $(N_0)$ for each wavenumber step can be done (this allows for corrections of intensity fluctuations of the ion source) and the population yield of the parent ion signal is measured. Plotting $N_1(\tilde{\nu})/N_0$ leads to the parent ion depletion spectrum and the absorption spectrum is obtained by converting the depletion spectrum into the absorption cross section $\sigma(\tilde{\nu})$ as well as correcting for the power of the laser over the tuning range $F(\tilde{\nu})$ (see middle part of Figure 1.11):

$$\sigma(\tilde{\nu}) \propto \frac{1}{F(\tilde{\nu})} \ln\left(\frac{N_0}{N_1(\tilde{\nu})}\right) \tag{1.5}$$

A constant linear drop in the ion signal intensity (for example as a result of a change in the solution over time) during the measurement can be corrected by subtraction of a linear baseline, defined by at least two $N_0$ values. The advantage is the reduction of the required measurement time to obtain one infrared spectrum. The disadvantage, on the other hand, is that small sudden temporal fluctuations of the ion source are not corrected and that those can be misinterpreted as real peak features in the obtained infrared spectrum. However, the recording and averaging of multiple scans can minimize the appearance of those 'artifacts' in the infrared spectrum. The settings of the irradiation time and intensity are generally such that the depletion of the signal is not more than 50% to avoid saturation effects.

Secondly, if fragment ion masses due to the resonant absorption of light of the parent ion are observed in the mass spectrum a background-free infrared spectrum can be recorded. $N$ corresponds to the sum of the observed integrated parent ion signal intensity $N_1(\tilde{\nu})$,

together with the integrated intensities of the fragment ion signals $A^+, B^+, ...$ which is $N_2(\tilde{v})$. The depletion spectrum is thus generated by plotting

$$\frac{N_1(\tilde{v})}{N_1(\tilde{v}) + N_2(\tilde{v})}$$

as a function of the wavenumber $\tilde{v}$ and the absorption spectrum is obtained by converting the depletion spectrum to $\sigma(\tilde{v})$ and correcting for the power of the laser $F(\tilde{v})$ over the tuning range approximatively (see right part of Figure 1.11). The advantage of this method is the correction for fluctuations in the ion source, without an increase in the recording time for one infrared spectrum. The settings of the irradiation time and intensity are generally such that the fragmentation yield is not more than 50% to avoid saturation effects. Often, however, it is difficult to record the fragment ion masses in the mass spectrometer due to their low ion signal intensity or a possible distribution over a wide mass range. The described technique is used to record infrared spectra of protonated polyalanine chains, see Chapter 4.

It is also possible, to obtain an IR-MPD spectrum by monitoring the fragment ion yield $Y$ as a function of the wavenumber $\tilde{v}$ and normalizing it with respect to the total ion population yield:

$$Y = N_2(\tilde{v})/\left(N_1(\tilde{v}) + N_2(\tilde{v})\right). \tag{1.6}$$

Due to the highly unstable parent ion signal and an intensity orders of magnitudes higher compared to the fragment ion intensity in the experiment on the fullerene anions (see Chapter 2), the parent ion signal intensity is not taken into account, rather only the fragment yield is monitored, $Cl^-$ in this particular case.

Infrared spectra of gas-phase ions can be additionally obtained with the so-called 'messenger' technique [62]. A rare gas atom X referred to as the messenger is non-covalently bound to the system of interest. The molecular complex $[AB^+ \cdot X]$ can be photodissociated by absorption of a single photon due to the low binding strength of the messenger atom and as a consequence, the bare ion can be measured in the mass spectrometer. The structure of the molecule is not or only weakly affected by the messenger atom due to its low binding energy.

$$[AB^+ \cdot X] + n \cdot hv \longrightarrow [AB^+ \cdot X]^* \longrightarrow AB^+ + X. \tag{1.7}$$
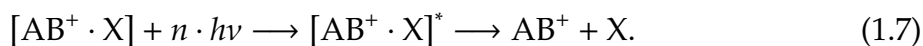
In the experiment on cytochrome c, a similar approach is taken to record infrared spectra. An alkali ion is used as a messenger. As it is more strongly bound compared to a rare gas atom, absorption of multiple photons is needed to dissociate the formed complex. If infrared photons are absorbed by the molecule, the alkali ion can be dissociated, thus leading to a decrease of the charge of the protein complex by one and the observed ion signal intensity of this charge reduced complex can be used to calculate the infrared absorption cross section (see Chapter 3).

To record infrared spectra of anions, it is possible to use an electron scavenger as an indicator for resonant absorption of the parent ion. A background gas, for example, $CCl_4$ or $SF_6$, is injected into the ion trap chamber. The dissociated electron can then be
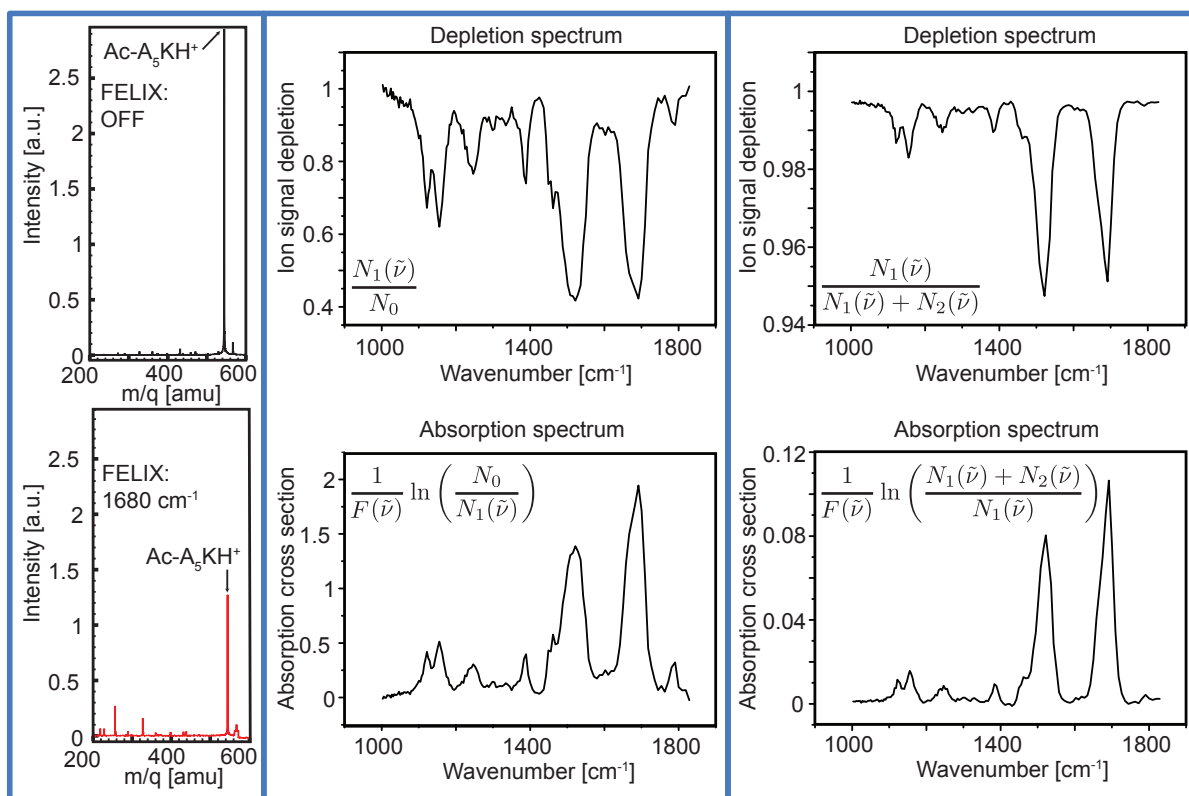
**Figure 1.11:** On the left, mass spectra of $Ac-A_5-K-H^+$ are shown with and without irradiation of FELIX (bottom and top respectively). The equations used to obtain the depicted infrared spectra are given in the corresponding graph, see text for details. The fragment ions used to calculate $N_2(\tilde{\nu})$ in the leftmost Figure are presented in Table 4.3.

caught by the surrounding gas molecules, forming $Cl^-$ and $SF_6^-$ respectively, which can be measured in the mass spectrometer. This technique is used to obtain infrared spectra by anionic fullerenes and the results are described in Chapter 2.

All investigations on the molecular systems shown in this work are performed using the FT ICR mass spectrometer at the free electron laser facility FELIX [57]. Some remarks concerning experimental details are presented here: Firstly, the power of FELIX is not constant over the scanned wavelength range. To correct for differences in power the infrared radiation is guided to a power meter and the power is recorded at different wavelengths. Therefore, these obtained data points are normalized and plotted against the wavelength. A polynomial fitting function (typically third order) is used to fit the data points, and the power corrected values are calculated by dividing the measured ion signal at a certain wavelength $N_1(\lambda)$ by the corresponding fitted relative power intensity. This method is motivated by a near linear correlation of the fragmentation process to the power intensity of the laser beam [58].

Secondly, the set wavelength in the measurement program does not necessarily reflect the actual wavelength value of the laser. Therefore, a wavelength calibration step is performed for each group of measurements and taken into account in the evaluation of

the experimental data. The infrared light is analyzed using a calibrated wavelength spectrometer and the actual wavelength $\lambda_{real}$ is measured as a function of the desired wavelength $\lambda_{set}$, which is also used in the measurement program. Due to the fact that the actual value of the wavelength $\lambda_{real}$ can be either higher or lower than $\lambda_{set}$, the difference $\Delta\lambda$ between the real and set wavelength is taken and plotted versus $\lambda_{set}$. A polynomial function is fitted to the obtained points. Subtracting $\Delta\lambda$ for each corresponding wavelength step from $\lambda_{set}$ of the recorded spectra leads to the corrected wavelength value, which is then converted to the wavenumber.

### 1.4.3   Free electron lasers as powerful high intensity light sources

In table-top lasers, the active medium is a gas, a liquid, or a solid compound, which limits the wavelength due to absorption, characteristic wavelength regimes for a medium, and sometimes the intensity when the damage threshold of the active material is smaller than the deposited energy. In free electron lasers, the active medium are unbound, so-called free electrons traveling with relativistic energies through a periodical magnetic structure. The structure is referred to as undulator or wiggler and can be located between two mirrors that form an optical cavity. In principle it is possible to produce photons at any wavelength with a free electron laser. The physical principles of free electron lasers are extensively described in many books for example by Freund and Antonsen [63]. A short introduction of the working principles of free electron lasers is given here together with characteristic properties of the free electron laser FELIX.

Electrons are produced in an electron gun, bunched together, and accelerated in a linear electron accelerator (LINAC) to relativistic energies. They then enter the undulator of a given length $L$, an alternating periodical magnetic structure (spaced by $\lambda_u$) whose magnetic field is perpendicular to the direction of the electron beam (see Figure 1.12). Due to the Lorentz force, the magnetic field causes a wiggling motion of
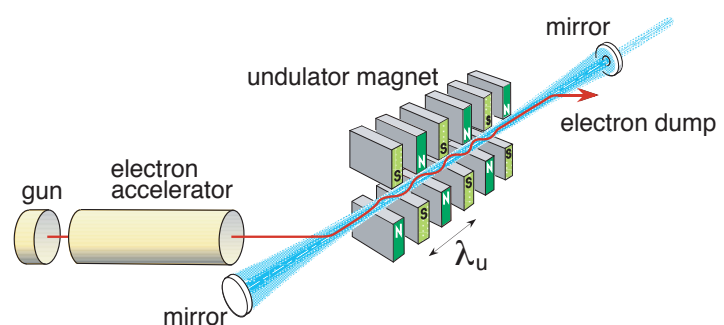


**Figure 1.12:** Shown is a schematic diagram of the working principles of a free electron laser [64]. The details are explained in the text.

the electrons while traversing the undulator and as a consequence of the acceleration[6] radiation is emitted with a frequency of the oscillation period in the direction of the electron beam. The latter property is caused by the relativistic energy electrons, which behave like a Hertz dipole and thus emit radiation perpendicular to their oscillation

---

[6]velocity direction change

direction. The wavelength of the undulator period is relativistically contracted in the electron rest frame to $\lambda_u/\gamma$ with the Lorentz factor $\gamma$ and corresponds to the wavelength of the emitted photons. The Lorentz factor is a function of the kinetic energy of the electrons $mc^2$, the electron rest energy $m_0c^2$, and $\beta$ is defined as the velocity of the electrons $v$ with respect to the speed of light $c$:

$$\gamma = mc^2/m_0c^2 = 1/\sqrt{1-\beta^2}. \tag{1.8}$$

Due to the high velocity (almost the speed of light) of the wiggling electrons, a strong Doppler shift of the irradiation to higher frequencies is experienced in the laboratory frame. This, in turn, leads to the wavelength being reduced by a factor of $\approx 1/(2\gamma)$ as a relativistic effect. The emitted light is concentrated in a narrow cone in the forward direction, referred to as the 'head-light' effect. Due to the fact that the path length of the travelling wiggling electron and the path length of the emitted radiation are different, a dimensionless parameter $K$, which accounts for the undulator period $\lambda_u$ and the amplitude of the magnetic field $B_u$, is introduced and the wavelength $\lambda$ of the emitted light of the laser in the laboratory frame is given by:

$$\lambda = \frac{\lambda_u}{2\gamma^2}\left[1+K^2\right] = \frac{\lambda_u}{2\gamma^2}\left[1+\left(\frac{eB_u\lambda_u}{2\pi m_0c}\right)^2\right]. \tag{1.9}$$

However, the emitted radiation, referred to as spontaneous emission, is often very weak in intensity and scales linearly with the number of electrons $n$. The spectrum of the radiation is determined by the given width of the undulator length $L$ ($L = N\lambda_u$, $N$ being the number of periodic structure elements) with a finite transient time of the electrons as a consequence. Due to a spatial spread of the electrons larger than the emitted wavelength, the radiation is not emitted coherently. To increase the emitted radiation of the electrons, the undulator can be built inside an optical cavity defined by two mirrors and bunches of 'fresh' electrons coming from the LINAC can enter the undulator one after another. The interaction between the magnetic field of the emitted radiation and the magnetic field of the permanent undulater magnets leads to a beat; the so-called ponderomotive wave. The design of the undulator, with respect to the achieved magnetic field and the length, influences the frequency of the ponderomotive wave giving the possibility to adjust it to the emitted radiation. The phase velocity of the beat can be tuned through the velocity of the electron bunches. The energy of the injected electrons into the undulator should be slightly higher than the resonance energy for the desired emitted wavelength, because electrons are decelerated by the ponderomotive force. Their energy is transferred to the electromagnetic field and can amplify the radiation until saturation of the oscillator is reached. The power of the gained laser light is $10^7$ to $10^8$ times higher than of spontaneous radiation alone and thus, free electron lasers provide a very high photon flux.

In Figure 1.13 a scheme of the free electron laser for infrared experiments FELIX [65] is shown.
As described in Equation 1.9, the wavelength of the radiation can be set either by the energy of the electrons or by characteristics of the undulator, for example by varying the magnetic field strength. The latter is used to change the frequency of the infrared
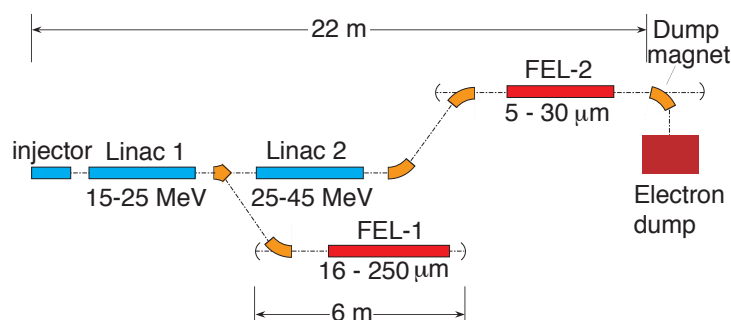
**Figure 1.13:** A scheme of the free electron laser FELIX [64]. The details are explained in the text.

photons by changing the distance between the permanent magnets in the undulator. This allows for a rapid and continuous scanning range of about a factor of three around the central wavelength that is set by the electron beam energy. There are two LINACs present, which are used depending on the desired wavelength of the photons: FEL-1 delivers $16 - 250$ $\mu$m and FEL-2 $5 - 30$ $\mu$m.

FELIX is a pulsed laser, the temporal scheme of the pulses is determined by the electron bunches. Every 1 ns, a new electron bunch enters the undulator, which results in a macro-pulse, composed by micro-pulses. Figure 1.14 shows the pulse structure of the FELIX light.
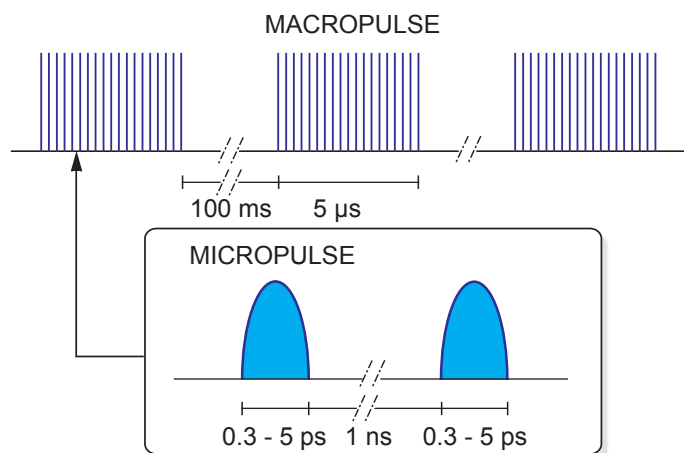


**Figure 1.14:** A scheme of the pulse structure of the free electron laser FE-LIX [64]. The details are explained in the text.

The characteristics of FELIX are presented in Table 1.2. The infrared radiation is guided in a vacuum tube system to different user stations where experiments can be performed. For each beam time shift, the laser is exclusively operated for one experiment. The pulse structure of FELIX is perfectly suitable for IR-MPD experiments (see Section 1.4.1) and many new experiments giving insight into molecular structures are performed at this facility. Unfortunately, the beam time at this facility is very limited. Applications to be allocated beam time need to be made well in advance.

Due to the limited beam time available, it happens that not all interesting experimental questions, which often open up just while performing measurements, can be answered accurately. It would have been useful to record, for example, infrared spectra of cytochrome c under different conditions, see Chapter 3 or to record more infrared spectra of $C_{70}^-$, see Chapter 2. At the moment, a free electron laser with similar properties

**Table 1.2:** Properties of the free electron laser FELIX.

| Characteristics | Value |
| --- | --- |
| Coverable wavelength range | $5 - 250 \, \mu m$ |
| Spectral width (FWHM) | $0.4 - 7\%$ of the central wavelength |
| Linear polarization | $> 99\%$ |
| Macro-pulse length | $5 \, \mu s$ at 10 Hz repetition rate |
| Micro-pulse length | 300 fs to 5 ps at 25 MHz or 1 GHz repetition rate |
| Macro-pulse energy | up to 150 mJ at 1 GHz repetition rate of micro-pulses |
| Micro-pulse energy | 1 to 15 $\mu J$ |
| Electron beam energy | either $15 - 25$ MeV (FEL-1) or $25 - 45$ MeV (FEL-2) |

as FELIX is being built at the Fritz-Haber-Institut in Berlin. The then available light source will allow for a lot of exciting experiments on biomolecular systems in the gas phase by providing additional beam time.

## 1.4.4 Fourier transform ion cyclotron resonance mass spectrometer used to obtain IR spectra of biomolecular ions

To record mass and infrared spectra, the FT ICR at the FELIX facility is used [57]. An excellent review on the principles of FT ICR mass spectrometry is given by Marshall [66] and only a brief introduction to the concepts is presented here.
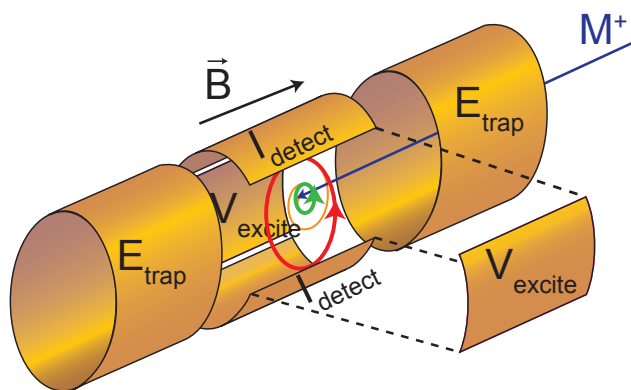


**Figure 1.15:** Schematic drawing of the ICR cell [64]. $\vec{B}$ indicates the direction of the magnetic field. Shown in blue is the direction in which the cations $M^+$ enter the cell. Due to the Lorentz force, the cations are circling perpendicular to the magnetic field (indicated in green). Upon excitation, the ion orbit radius increases (indicated in red), see text for details.

Charged molecules are trapped in the ICR cell, which is mounted inside a magnet, by applying a trapping potential to the electrodes $E_{trap}$. The Lorentz force $\vec{F} = q \left( \vec{v} \times \vec{B} \right)$, which depends on the charge of the ion $q$, the velocity $\vec{v}$, and the magnetic field $\vec{B}$, causes trapping of the ions perpendicular to the magnetic field lines. Parallel to the magnetic field lines, a trapping potential is applied on two opposite electrodes $E_{trap}$ and the ions circulate inside the cell. The angular frequency $\omega_c$ (referred to as the ion cyclotron frequency) of the circulating ions is anti proportional to $m/q$ of the ion and is given by:

$$\omega_c = \frac{qB_0}{m} \tag{1.10}$$

with the amplitude of the magnetic field $B_0$. The ions are spread out on the orbit and need to be coherently excited to a higher orbit and to an oscillation in phase of the ion packet. This can be achieved by applying an excitation pulse $V_{excite}$ on two opposing excitation plates (see the green circle in Figure 1.15). Once the ion packet oscillates in phase, a periodical image charge is induced on the detection plates $I_{detect}$, located perpendicular to the excitation plates, which reflects the ion cyclotron frequencies. This periodical temporal alternating image current is amplified and recorded with a digitizer resulting in a transient. Fourier transformation of this transient leads to the ion cyclotron frequencies and subsequently to the mass-to-charge ratios of the trapped ions.

On the excitation plates, an alternating current (AC) pulse can be applied to resonantly excite the ions to even higher ion cyclotron radii (see red circle in Figure 1.15). When the orbit of the ions is sufficiently high they collide with the electrode walls and are ejected from the trap. This allows for mass isolation by applying a selective excitation pulse sequence such, that only ions with a certain ion cyclotron frequency $\omega_c$ are excited. This pulse sequence is referred to as stored waveform inverse Fourier Transform (SWIFT) ICR excitation pulse [67].

The maximum number of trapped ions is defined by the space charge limit in the FT ICR. The mass resolution of the mass spectrometer depends on the pressure inside the cell as well as the transient length. Collisions of the ions with background gas leads to a damping of the periodical observed ion signal and to a de-phasing of the ion packet such that the ions are distributed over the whole orbit. The number of data points for evaluation by a Fourier transformation is defined by the transient length and the sampling speed of the digitizer. The lower the pressure and/or the longer the recorded transient, the better the mass resolution achieved.

In principle, the upper limits of the mass resolution of FT ICR mass spectrometer can be very high as well as the upper limit of measurable masses. The latter depends linearly on the magnetic field strength, quadratically on the achievable ion cyclotron radius inside the trap, as well as on the trapping voltage on the trapping electrodes. Due to two different reasons, the mass resolution in the presented experiments is not perfect: Firstly, the background pressure is not low enough. Secondly, the recorded transient time is in the order of 32 ms to minimize the time of each scan[7] and should be longer.

Unfortunately, it was not possible to exclude all noise peaks in the mass spectra as presented in this work. Due to the high sensitivity of the amplifier of the image current detector, electromagnetic interferences from, for example, switching power supplies and pressure gauges, are recorded and observed in the mass spectrum. Characteristics of those effects are spurious signals with a width of only a few (two or so) data points in the mass spectra. In mass spectra presented in this work, those noise peaks are indicated by an asterisk.

A schematic drawing of the apparatus is shown in Figure 1.16. Ions are produced by electrospray ionization and guided via a hexapole RF ion guide, a quadrupole bender,

---

[7]The transfer of the data from the apparatus to the computer and the execution of the Fourier transform are also time consuming and on the order of a second.
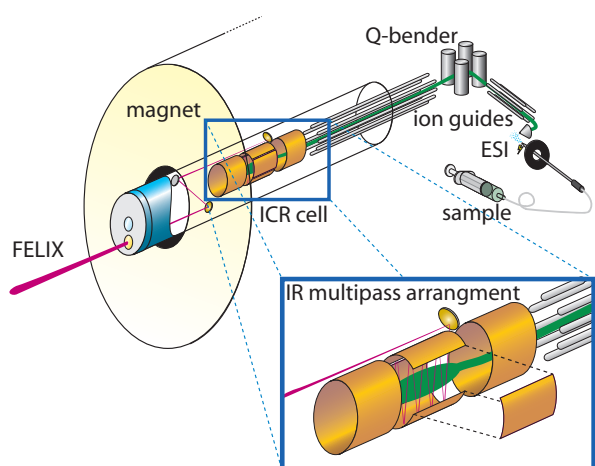
**Figure 1.16:** Principles of a FTICR mass spectrometer [64]. On the right a close up of the ICR cell is shown, details of which are given in the text.

and a one meter long octopole RF ion guide into the ion cyclotron resonance (ICR) cell. Prior to injection of the ions into the cell, a gas pulse (typically $N_2$) can enter the trap region and by collisional cooling the ions are slowed down and can be trapped. To avoid collisions of the ions with gas molecules, a second possibility to decrease the velocity of the ions exists: Once the ions are in the octopole, the dc potential of the guide is lowered below the trapping potential. Thus, when the ions leave the guide their kinetic energy decreases while gaining potential energy (climbing up the potential hill). The ICR cell is mounted inside a 4.7 T actively shielded superconducting magnet and consists of six different elements (see Figure 1.15): two trapping electrodes $E_{trap}$, two opposing image current detection plates $I_{detect}$, and two opposing ion excitation plates $V_{excite}$.

The FELIX laser beam can be introduced into the ICR cell via a window (material depends on the wavelength of the laser) on the backside of the FT ICR apparatus (see Figure 1.16). The beam is reflected on the highly polished copper electrodes to maximize the overlap with the trapped ion cloud. A shutter in front of the window (on the air side) allows for blocking the laser beam.

A typical sequence of an experiment to record an infrared spectrum is as follows:

1. Ions are trapped in the ICR cell and the ions of interest are mass isolated by a SWIFT pulse.

2. Laser light of a certain wavelength is introduced into the cell for about 3 – 6 s (irradiation time depends on the system).

3. A mass spectrum of the trap content is recorded.

4. The procedure (step 1 – 3) is repeated about four times and the obtained mass spectra are averaged.

5. The wavelength of the laser is changed and the procedure starts at step 1. Alternatively, the procedure starts at step 1 with the shutter preventing the introducing of laser light and the next time on step 5, the wavelength is changed.

A Yokogawa oscilloscope as a data acquisition system together with a version of the Modular ICR Data Acquisition System (MIDAS) software [68] to control the apparatus

is used to perform experiments on gas-phase proteins (see Chapter 3) and on gramicidin S (see Chapter 5). For the other experiments (fullerene anions: Chapter 2, polyalanine: Chapter 4, and a *coiled coil* system: Chapter 6) a digitizer and a new FT ICR control software ('awgscan' and 'awe') adapted by a version described by Heeren *et al.* [69] is used.

## 1.5   Outline of the thesis

The methods and techniques described in this Chapter are used throughout the whole thesis and the outline of the thesis is:

**Chapter 2**  Infrared spectra of $C_{60}^-$ and $C_{70}^-$ are presented and structural information is obtained by comparing the results to calculations. With the investigation on those systems the possibilities of structure determination of gas-phase ions by vibrational spectroscopy is shown.

**Chapter 3**  Vibrational spectroscopy is performed on two proteins in the gas phase, namely cytochrome c and tendamistat and the results are shown and discussed. Due to the size of the molecules and thereby the huge conformational freedom it is hard to relate structures to the obtained infrared spectra. As a consequence, smaller systems are investigated in the following Chapters.

**Chapter 4**  Peptides based on alanine residues are investigated in the gas phase by vibrational spectroscopy. They can form stable $\alpha$-helices in the gas phase and thus serve as calibration points for this secondary structure element.

**Chapter 5**  A possible model system for a $\beta$-sheet motif in the gas phase is the peptide gramicidin S. Spectroscopic investigations on this molecule are performed and results are shown.

**Chapter 6**  Vibrational spectroscopy is performed on molecules which are especially designed to possess *coiled coil* conformations if they are complexed to form dimers and results on first investigation of non-covalently bound protein complexes in the gas phase are presented.

**Chapter 7**  The spectroscopically obtained results on the molecules investigated in this thesis are summarized and a proposal on how to improve vibrational spectroscopy in the mid infrared range on gas-phase biomolecular systems to gather information about their structure is given.

# Chapter 2

# IR spectroscopy of gas-phase fullerene anions[*]

The IR spectrum of the gas-phase $C_{60}^-$ and $C_{70}^-$ anions is presented. The anions are mass-selected and trapped in an ion cyclotron mass spectrometer cell. The spectrum is obtained by recording the yield of electrons following infrared multiple photon absorption and subsequent electron detachment of the anions. For $C_{60}^-$ two bands are observed that correlate to two of the four IR allowed transitions in neutral $C_{60}$. The results show that the higher frequency band is strongly shifted from its position in neutral $C_{60}$ and is a sensitive marker for the charge state. The band positions and intensities are compared to results obtained by theory as well as to known bands in solid samples and good agreement is found. An infrared spectrum of gas-phase $C_{70}^-$ is presented and compared to calculated spectra of the anion and the neutral counterpart.

## 2.1   Introduction

Fullerenes and fullerene based materials attract a lot of attention and are the focus of thousands of publications. Especially the solid state properties of doped fullerene materials are of interest. When bulk $C_{60}$ is doped with alkali atoms, intercalated compounds $M_nC_{60}$ (n = 1,3,4 or 6) can form [70–73]. In those materials, the alkali atoms transfer charge to the fullerene, giving rise to $C_{60}^{n-}$ ions surrounded by n alkali ions. It is observed that $M_nC_{60}$ with n = 4 behaves as an insulator while, for example, with n = 3 metallic behavior can be found. Depending on the metal, some even have superconducting properties [71–73] and recently, $Cs_3C_{60}$ was synthesized showing a threshold to superconductivity of 38 K – so far the highest value for a molecular system [74].

It thus comes as no surprise that the properties and the electronic structure of fullerene anions are of special interest. The addition of one or more electrons can cause Jahn-Teller (JT) distortion [75]. For $C_{60}$ and other fullerenes, this distortion has been investigated both theoretically [75–78] and experimentally [79–84]. Especially $C_{60}^-$ serves as a model

system for the JT effect in highly symmetrical molecules and the molecule is expected to experience a distortion from icosahedral ($I_h$) to either $D_{5d}$ or $D_{3d}$ symmetry [75]. In the solid state, $Rb_1C_{60}$ has been investigated using IR reflectivity measurements [85]. These studies show that the position of the highest frequency IR allowed mode of icosahedral $C_{60}$ is sensitive to the charge state. However, questions have been raised as to the role of the counter ions, and, in order to minimize their interactions, special compounds have been synthesized and investigated [84, 86, 87].

$C_{60}^-$ has also been studied in the gas phase. Using photoelectron spectroscopy, the electron affinity of neutral $C_{60}$ is determined to be 2.70 eV [80]. $C_{60}^{2-}$ can be observed in the gas phase as well [88]. It is, however, only metastable and can decay to $C_{60}^-$ and a free electron [89]. $C_{60}^-$ is also investigated spectroscopically in an ion storage ring. From these measurements, it is concluded that $C_{60}^-$ is dynamically distorted to presumably $D_{3d}$ symmetry [82].

Previously, it was shown that neutral gas-phase $C_{60}$ can be efficiently excited using tunable IR light from a free electron laser [52, 90, 91]. When the radiation is resonant with an IR allowed transition in the molecule, the sequential absorption of hundreds of photons can take place, resulting in very hot molecules. These hot molecules can cool by the emission of an electron and the positively charged $C_{60}^+$ can be detected as a function of IR wavelength. Plotting the ion yield as a function of wavelength gives a spectrum which was shown to be similar (but not equal to) the linear absorption spectrum. Here we show that gas-phase $C_{60}^-$ can be excited using IR light and that the hot molecules can detach an electron. The detachment yield can be monitored as a function of wavelength, yielding an IR spectrum of $C_{60}^-$.

As example for another molecule with many atoms, yet high symmetry, is $C_{70}$. While the symmetry of $C_{60}$ is icosahedral, the symmetry of the larger neutral molecule $C_{70}$ is found to be $D_{5h}$. Solid $C_{70}$ has been investigated using a multitude of methods, among them electronic absorption spectroscopy [92], IR spectroscopy [93] (recently also a temperature dependent study ranging from 93 K to 523 K [94]), Raman spectroscopy [95, 96], and as a thin film on metal substrates in the UV-near IR region [97]. Further, neutral $C_{70}$ molecules from solution are investigated on ZnSe slides in the infrared range from 650 cm$^{-1}$ to 2500 cm$^{-1}$ by Fourier transform infrared (FT-IR) spectroscopy [98] and six bands are observed at 1462 cm$^{-1}$, 1431 cm$^{-1}$, 1415 cm$^{-1}$, 1134 cm$^{-1}$, 795 cm$^{-1}$, and 674 cm$^{-1}$ and a $D_{5h}$ symmetry is assigned to the structure. Electronic spectra of $C_{70}$, $C_{70}^-$, and $C_{70}^+$ embedded in a 5 K neon matrix [99] as well as of $C_{70}$ anions at 77 K (produced in $\gamma$-irradiated organic glasse) are recorded [100]. Resonance enhanced multiphoton electron detachment (REMPED) spectra for $C_{70}^-$ in the range of 7000 cm$^{-1}$ to 10000 cm$^{-1}$ are taken and Jahn-Teller distortion of the anion is observed [101]. Recently, Hampe *et al.* investigated gas-phase $C_{76}^{2-}$ by infrared multiphoton electron detachment spectroscopy using a free electron laser. The overall $D_2$ structure of neutral $C_{76}$ seems to be retained for the di-anion.

To the best of my knowledge, only one investigation using mid-IR spectroscopy on the anionic $C_{70}^-$ molecule, obtained by embedding neutral $C_{70}$ in a salt structure

$([(C_6H_5)_4P]_2C_{70}Br)$ is performed [102]. The extra electron in the anion distorts the $D_{5h}$ symmetry of the neutral molecule due to the Jahn-Teller effect. As a consequence, the symmetry of the anion is reduced to a $C_{2v}$ symmetry [103] or even to a $C_s$ symmetry [101]. Here, an infrared spectrum of gas-phase $C_{70}^-$ is presented and is compared to spectra observed of $C_{70}$ embedded in solid state matrices as well as to calculated spectra of the anion (using density functional theory).

## 2.2 Experimental

The experiments are performed at the FOM-Institute for Plasma Physics "Rijnhuizen" in the Netherlands using the Free Electron Laser for Infrared eXperiments (FELIX) [65]. $C_{60}$ or $C_{70}$ anions are generated via electrospray ionisation (ESI) and transported to and accumulated in a hexapole ion trap. The ions are then transfered into the fourier transform ion cyclotron resonance (FT ICR) mass spectrometer, see Chapter 1, Section 1.4.4. Trapping, mass-selective isolation, and detection of the ions, as well as irradiation with IR photons is done inside the ICR cell. A leak valve allows for the admission of reactant gases.

$C_{60}$ and $C_{70}$ was obtained from MER company, Texas, whereas all other chemicals were purchased from Sigma-Aldrich. Both fullerene anions are generated using the ESI source in a procedure developed by Kappes and coworkers [104]. $C_{60}$ and $C_{70}$ respectively is dissolved in 1,2-dichlorobenzene at a concentration of 500 $\mu$M. As electron donor for the fullerenes, a solution of 10 mM (tetrakis)-dimethylaminoethylene (TDAE) in toluene is used. Due to its sensitivity to oxygen, the TDAE solution is prepared under nitrogen atmosphere. In order to avoid blocking of the 125 $\mu$m diameter stainless steel ESI capillary with aggregates produced in the solution and to improve the signal stability, the two solutions are mixed just before injection into the capillary. For this purpose, each solution is supplied by its own syringe pump (Harvard Apparatus 11plus). The solutions are combined, passed through a 5 $\mu$m filter and enter the ESI capillary. The flow rates are set to 1 $\mu$l/min for the TDAE solution and 2 $\mu$l/min for the fullerene solution. The ESI needle voltage is adjusted to $-3$ kV and the cone voltage to around $-60$ V. The aggregation process for the $C_{70}$ solution is much faster compared to the $C_{60}$ solution. This is indicated by the growth of black deposits in the solution and as a consequence, the ESI needle is blocked rapidly and the solution needed to be changed regularly.

The anions are transferred to and trapped in the ICR cell, which is optically accessible via a KRS-5 window at the back end. To maximize the overlap of the IR beam from FELIX with the ion cloud, a multipass arrangement is created using the polished inside surfaces of the excite plates of the cylindrical ICR cell [105].

FELIX is tunable over the range from 40 cm$^{-1}$ to 2000 cm$^{-1}$, although here we investigate only the 400-1700 cm$^{-1}$ range. The light output comes in macropulses of 5 $\mu$s duration containing 0.3 - 5 ps long micropulses at a repetition rate of 1 GHz. The bandwidth is adjustable and transform limited by the micropulse duration. In the

current experiments, the bandwidth amounts to approximately 0.5% of the central wavelength and at these settings, the pulse energy in a macropulse is about 35 mJ at the experiment.

In the present experiment, the irradiation time is set to 7.2 s with a macropulse repetition rate of 5 Hz. When the IR light is resonant with an IR active mode of the ion, the absorption of many photons can occur [58], which can lead to the detachment of an electron. Although these electrons remain trapped in the ICR cell, their cyclotron frequency is too high to be detected by the data acquisition system. Nonetheless, they can be detected indirectly by monitoring reaction products of the electrons with electron capture agents, admitted to the ICR cell [106, 107]. In the present experiments, two different electron capture agents are used: $SF_6$ yielding $SF_6^-$ as reaction product, and $CCl_4$, which undergoes dissociative electron attachment with $Cl^-$ as product. Adding the neutral reaction partners raises the base pressure in the apparatus from around $1 \times 10^{-8}$ Torr to $1.4 \times 10^{-7}$ Torr. IR spectra can then be obtained by either monitoring the depletion of the parent ion or the appearance of a reaction product as a function of IR wavelength.

A typical mass spectrum of the $C_{70}^-$ is shown in Figure 2.1. On the left, IR pho-
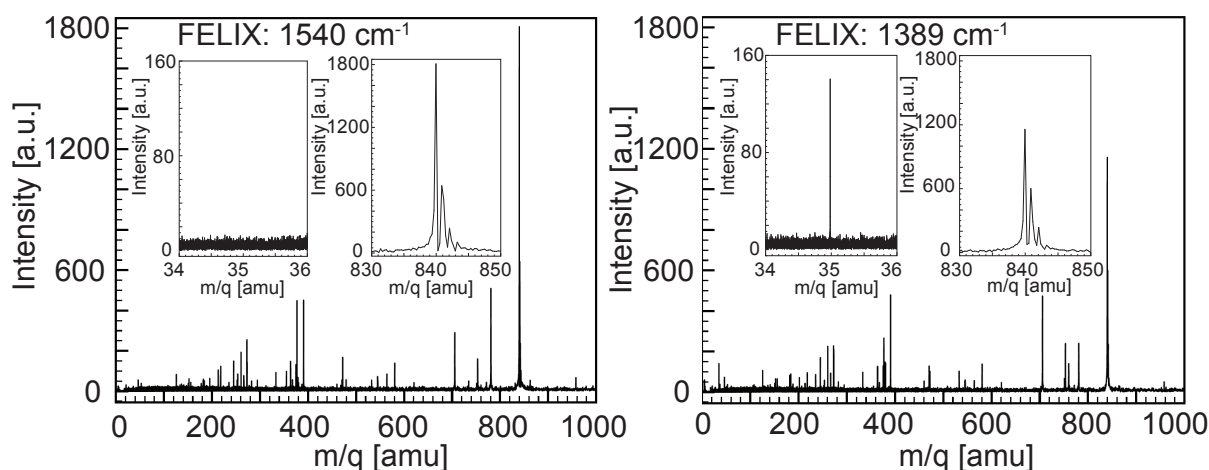


**Figure 2.1:** Mass spectra of $C_{70}^-$. On the left, the infrared irradiation inside the ICR cell prior to recording the spectrum is set to 1540 cm$^{-1}$. On the right, the infrared irradiation inside the ICR cell prior to recording the spectrum is set to 1389 cm$^{-1}$. An infrared transition of the molecule is observed and additional chloride ions appear in the mass spectrum. The insets are magnifications in the region of the chloride ion mass and of the parent ion mass.

tons are off resonant to any vibrational mode of the molecule. On the right, photons at a wavelength of 7.2 $\mu$m (1389 cm$^{-1}$) are absorbed by the anion and a decrease in the parent ion signal together with an appearance of $Cl^-$ ions is observed. Other ions (unassigned) due to impurities in the solution are observed in the mass spectrum of $C_{70}^-$. The appearance of the chloride ions is used as a marker for photon absorption of the molecule.

Density functional theory (DFT) is used to calculate theoretical IR spectra of both fullerene molecules. The calculations are performed with Turbomole V5.10 [108] us-

ing the B3LYP functional and the def2-SVP [109–111] basis set. To generate starting geometries, icosahedral $C_{60}$ is compressed or elongated along a $C_3$, $C_5$ or $C_2$ symmetry axis to result in structures of $D_{3d}$, $D_{5d}$ or $D_{2h}$ symmetry, respectively. Geometries are optimized in their respective symmetries and vibrational frequencies and IR intensities are calculated using analytical second derivatives.
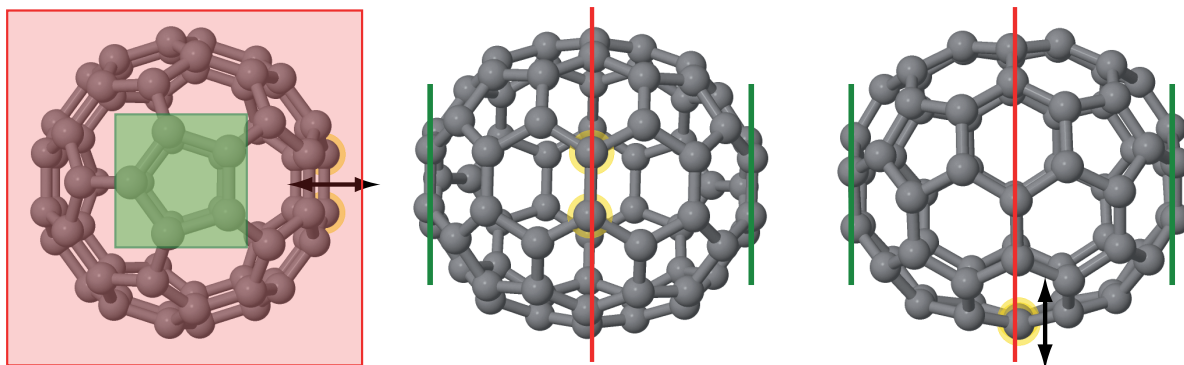


**Figure 2.2:** Three different views of the structure of $C_{70}^{-}$. The yellow marked atoms are displaced to generate the distorted structure (see text for details). Displacement is done along the direction of the arrows in the plane marked in red.

The situation for the bigger molecules is slightly different and $C_{70}$ in $D_{5h}$ symmetry is distorted such that a $C_{2v}$ or $C_s$ symmetry of the molecule is obtained. This is achieved by shifting two C-atoms (yellow marked in Figure 2.2, arrows indicate the direction of the shift) in the plane as marked in Figure 2.2: One group of geometries is generated by shifting the atoms in the plane about 0.5 Å towards the center of the molecule, another group is generated by shifting the atoms in the plane about 0.5 Å away from the center. The obtained geometries are optimized in their respective symmetries and in Figure 2.3 the optimized calculated structure of the anion in $C_s$ symmetry is compared with the calculated structure of neutral $C_{70}$ in $D_{5h}$ symmetry. For clarity, only the carbon atoms present in the three planes (marked in red and green in Figure 2.2) are shown.

Vibrational frequencies and IR intensities are calculated using analytical second derivatives. If the atoms are shifted towards the center, only transition states of the molecule are found in the calculations. If the atoms are shifted away from the center of the molecule structures are converged and energy minima in the potential energy landscape are found.

Different to the neutral molecule with $D_{5h}$ symmetry with only 52 infrared active vibrational modes[1] out of possible 204 normal vibrations, in lower symmetry more vibrational modes become infrared active. For $C_{70}^{-}$ in $C_{2v}$ symmetry 156 infrared active vibrational modes are expected and for the anion in $C_s$ symmetry, all 204 modes become infrared active.

---

[1] 21 doubly generated modes of symmetry type $E_1'$ and ten modes of symmetry type $A_2''$ [112]
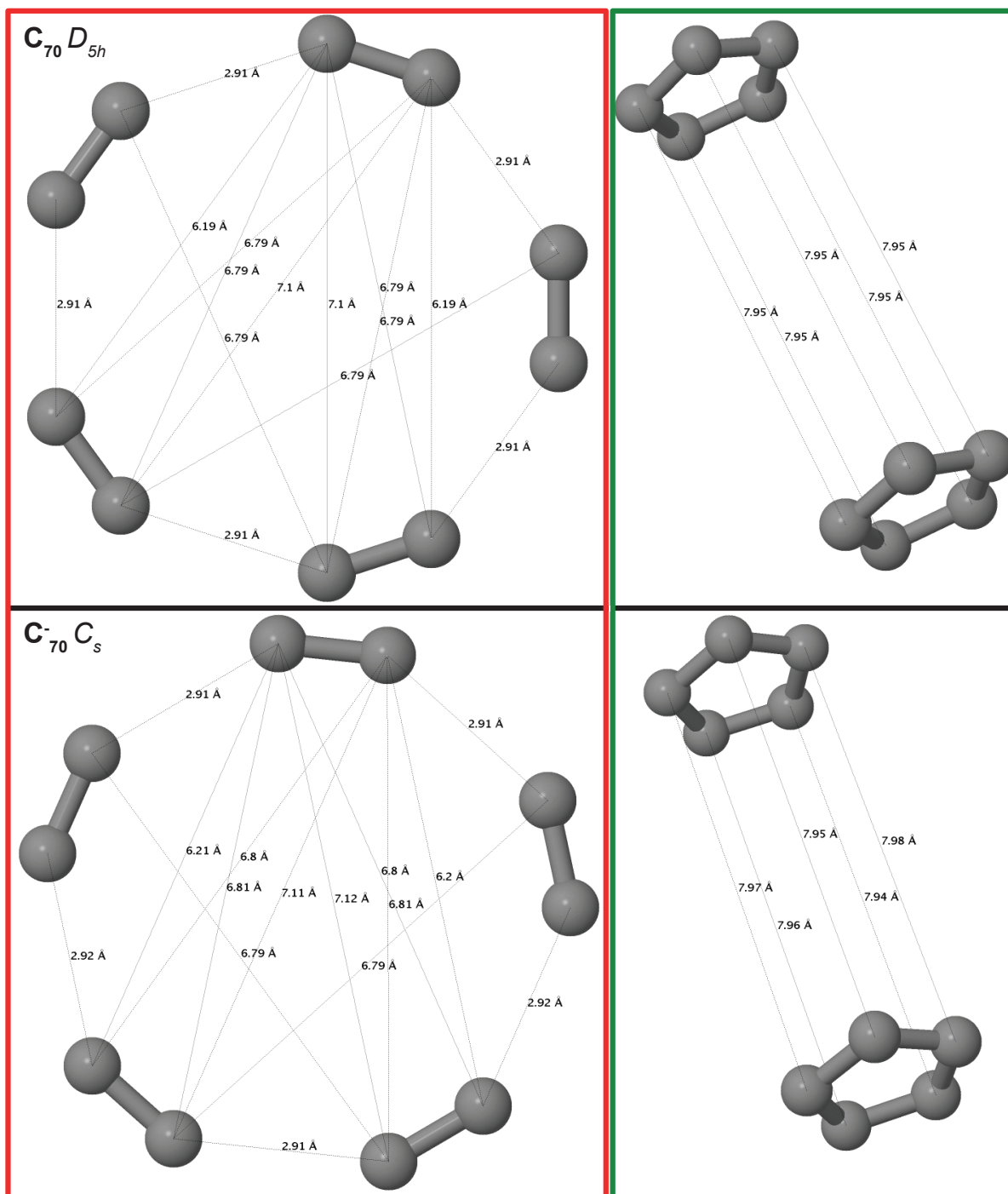
**Figure 2.3:** Top: neutral $C_{70}$ in $D_{5h}$ symmetry, bottom: $C_{70}^-$ in $C_{2v}$ symmetry. Distances are given in Å and measured between the carbon atoms as indicated by the dotted lines.

## 2.3 Results and discussion

Mass spectra of isolated and subsequently irradiated $C_{60}^-$ are recorded at different wavelengths of FELIX. If the IR photons are in resonance with an IR transition in the molecule, several hundreds of photons can be absorbed. In general, several cooling

pathways are possible for the hot molecules. The excited molecules may undergo fragmentation, detach an electron or emit radiation. At the excitation levels expected under our conditions, the latter process occurs only on a very long time scale and is not expected to be important here. All other processes cause a depletion of the parent ion signal. If the molecules undergo dissociation, this would have been observed as the appearance of additional peaks in the mass spectrum, which is not the case.

As mentioned above, detached electrons can not be detected directly, however, they can be monitored by admitting, for example, $SF_6$ or $CCl_4$ to the ICR cell. In the experiments presented here, reaction products of detached electrons are observed for both gases. However, $Cl^-$ as a reaction product has the advantage that it cannot absorb IR radiation and hence cannot introduce artifacts caused by IR induced dissociation of the product ion.

The appearance of $Cl^-$ ions accompanied by a decrease in the parent ion signal in the mass spectrum is an indicator for an IR transition in the fullerene anion.
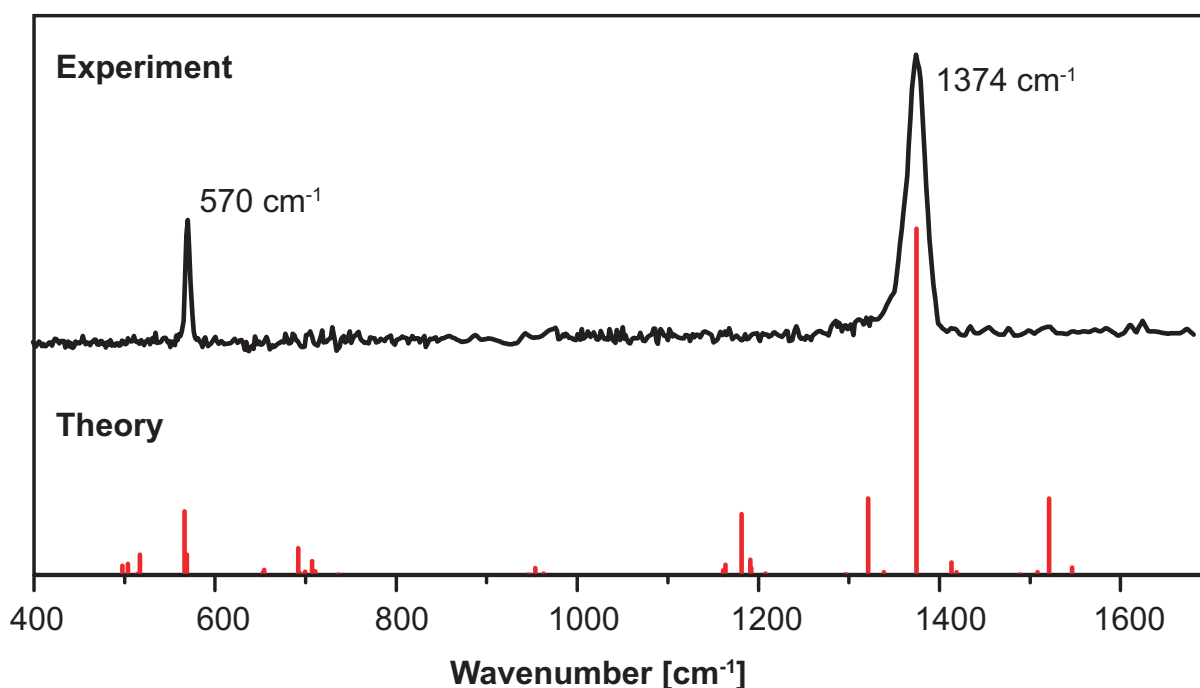


**Figure 2.4:** Experimental IR spectrum of $C_{60}^-$, recorded via the measurement of the yield of $Cl^-$ atoms as a function of IR frequency. The scaling of the ordinate is linear. No signal of chloride ions is obtained on the baseline. A theoretical spectrum for $C_{60}^-$ in $D_{3d}$ symmetry is shown as a stick spectrum. The calculations are performed using the B3LYP method and the def2-SVP basis set. Frequencies are scaled by 0.9648. See text for details.

Figure 2.4 shows the appearance of $Cl^-$ as a function of IR frequency together with a theoretical stick spectrum for $C_{60}^-$. In the experimental spectrum, the intensity is corrected for drift in the parent ion signal during a scan. Several spectra have been recorded and are averaged. At least three spectra were taken in every wavelength region, except between 850 $cm^{-1}$ and 970 $cm^{-1}$, where only one scan is performed. The spectra clearly show two peaks at 570 $cm^{-1}$ and 1374 $cm^{-1}$. The full width at half maximum of

these bands is 6 cm$^{-1}$ and 27 cm$^{-1}$, respectively. On both resonances, the appearance of chloride ions is accompanied by a depletion of the $C_{60}^-$ parent ion signal. In order to avoid saturation effects, the FELIX fluence is adjusted such as to keep the amount of depletion below 50%. The accuracy of the wavelength position is estimated to be ± 0.25%.

IR spectra of neutral $C_{60}$ in thin films [113] as well as for neutral $C_{60}$ in the gas phase are known [114]. For $C_{60}$ in icosahedral symmetry, only the four $T_{1u}$ out of 46 symmetry unique modes are IR active. In solid $C_{60}$, in addition weak resonances caused by anharmonic couplings and/or symmetry reduction due to the environment can be observed [115] and the fundamental transitions are found at 527 cm$^{-1}$, 576 cm$^{-1}$, 1183 cm$^{-1}$ and 1429 cm$^{-1}$.

The IR spectrum of neutral $C_{60}$ in the gas phase has been measured via IR emission spectroscopy [114] and IR resonance enhanced multiple photon ionization (IR-REMPI) spectroscopy of hot (875 K) $C_{60}$. There, the four $T_{1u}$ modes are found at 518 cm$^{-1}$, 557 cm$^{-1}$, 1145 cm$^{-1}$ and 1397 cm$^{-1}$. In addition a combination band is observed at 1525 cm$^{-1}$. The positions of these modes shift to lower energies with increasing temperature of the molecules. For example, the $T_{1u}(4)$ mode is red-shifted by about 40 cm$^{-1}$ at a temperature of 1830 K, compared to its value at 875 K [91].

Negatively charged $C_{60}$ has been investigated intensively in solid state compounds such as $M_nC_{60}$ with M = K or Rb and n ranging from one to six. In these compounds, each alkali atom donates one electron to $C_{60}$. In particular, infrared transmission as well as infrared reflectivity studies were performed on these compounds [85, 116, 117]. In solid RbC$_{60}$ four modes are found at 525 cm$^{-1}$ ($T_{1u}(1)$), 574 cm$^{-1}$ ($T_{1u}(2)$), 1182 cm$^{-1}$ ($T_{1u}(3)$) and 1392 cm$^{-1}$ ($T_{1u}(4)$) [85]. Compared to neutral $C_{60}$, the highest frequency $T_{1u}(4)$ band is observed to have the largest shift. The relative intensities also show a dramatic change with the $T_{1u}(1)$ and $T_{1u}(3)$ modes being very weak and barely observable [85]. For $M_nC_{60}$, n = 3 and 4, the $T_{1u}(1)$ and $T_{1u}(3)$ modes are not observed at all. With increasing charge state, the $T_{1u}(4)$ band shows a pronounced shift to lower frequencies while the $T_{1u}(2)$ hardly shifts (see Table 2.1).

In $M_nC_{60}$, the $C_{60}$ anion may be significantly perturbed by the alkali cations and this may have a significant effect on the electronic structure as well as on the position of the IR absorption bands. As a model system for the isolated $C_{60}$ mono-anion in a crystalline environment, $C_{60}$-tetraphenylphosphonium- and arsonium halides have been investigated. The fullerenes in these compounds are well separated from each other and surrounded by PhX$^+$ (Ph = phenyl = $C_6H_5$, X = P or As as counterions) [84]. In their IR spectra, two resonances at positions similar to those for RbC$_{60}$ are observed, with the low frequency mode showing a splitting of about 3 cm$^{-1}$ (see Table 2.1).

Adding a single electron to neutral $C_{60}$ is expected to cause Jahn-Teller distortion of the molecule. This lowers its symmetry from icosahedral to $D_{3d}$, $D_{5d}$ or $D_{2h}$ [75], and experimental [81] as well as theoretical [78] evidence indicates that the ground state is of $D_{3d}$ symmetry. As a result of this symmetry lowering, more transitions in the molecule are IR allowed, potentially leading to additional bands in the spectrum. In order to assign the experimental spectrum, density functional calculations were performed using the B3LYP methods with the def2-SVP basis set. For $C_{60}^-$ in $D_{3d}$, $D_{5d}$ and $D_{2h}$ symmetry, the geometries are optimized and the vibrational spectra calculated. The $D_{3d}$ structure is

found to be the lowest energy structure. Vibrational analysis reveals that this structure corresponds to a real minimum while the $D_{5d}$ and $D_{2h}$ structures are saddle points on the potential energy surface, indicated by the presence of imaginary frequencies in their vibrational spectra. The $D_{3d}$ structure is found to have a slightly oblate structure which is 2.38 eV below the optimized structure of neutral $C_{60}$ in $I_h$ symmetry. The position (scaled by 0.9648 [118]) and intensities of the calculated resonances are shown in Fig. 2.4. Comparing experiment and theory (see Table 2.1) shows that two strong transitions predicted by theory are observed in the experiment. The calculated low frequency

**Table 2.1:** Positions of the $T_{1u}(2)$ and $T_{1u}(4)$ vibrational bands of neutral and negatively charged $C_{60}$ molecules. Shown are band positions obtained by IR transmission, reflectivity and absorption spectra on solid samples and IR-REMPI, as well as emission spectra for gas phase molecules. Also shown are results from density functional calculations. All values are given in $cm^{-1}$.

| experiment (gas phase) | $T_{1u}(2)$ [$cm^{-1}$] | $T_{1u}(4)$ [$cm^{-1}$] | experiment (solid state) | $T_{1u}(2)$ [$cm^{-1}$] | $T_{1u}(4)$ [$cm^{-1}$] |
|---|---|---|---|---|---|
| $C_{60}$@875 K (IR-REMPI)[2] | 557 | 1397 | $C_{60}$ (absorption)[3] | 576 | 1429 |
| $C_{60}$ (emission)[4] | 570 | 1411 | | | |
| $C_{60}(I_h)$ (theory)[5] | 569 | 1418 | | | |
| $C_{60}^{-}$ (e$^{-}$detach.)[6] | 570 | 1374 | $(Ph_4As)_2ClC_{60}$ (transm.)[7] | 575/579 | 1387 |
| $C_{60}^{-}(D_{3d})$ (theory)[5] | 566 | 1375 | $Rb_1C_{60}$ (refl.)[8] | 574 | 1392 |
| | | | $K_3C_{60}^{-}$ (refl.)[8] | 572 | 1363 |
| | | | $Rb_3/K_3C_{60}$ (transm.)[9] | 573 | 1393 |

band near 566 $cm^{-1}$ actually consists of a mode of $E_u$ symmetry and a weaker mode of $A_{2u}$ symmetry shifted about 2 $cm^{-1}$ to the blue. This splitting is small compared to the 6 $cm^{-1}$ width of the measured band and is therefore not expected to be observable in our experiment. The calculated band at 1374.6 $cm^{-1}$ is of $E_u$ symmetry and stands

---

[2]Ref. [91].
[3]Ref. [115].
[4]Ref. [114].
[5]calculations scaled by 0.9648
[6]presented in this work
[7]Ref. [84].
[8]Ref. [85].
[9]Ref. [117].

alone. When comparing the experimental and theoretical peak positions, one can notice that theory fits experiment very well and that the deviation is only about 0.7% (see Table 2.1).

In the experiment, two bands are observed while theory predicts the presence of additional IR allowed modes in this frequency range. This discrepancy is possibly due to the excitation mechanism. $C_{60}$ has an electron affinity of 2.70 eV [80]. However, for a thermal detachment of the electron to occur on our experimental timescale, the internal energy needs to be several times higher. Thus, at around 500 cm$^{-1}$, more than 100 photons need to be absorbed for electron detachment to take place. The experimental spectrum is therefore not to be considered as a linear absorption spectrum. IR radiation from FELIX has previously been used to resonantly excite neutral $C_{60}$ followed by thermal ionization [52]. For that process to occur, several hundreds of photons need to be absorbed by a single molecule. The resulting IR-REMPI spectrum shows the presence of all four IR allowed modes, plus one combination band. It is observed, however, that the measured relative intensities do not correspond to the calculated ones. Those observations could be explained by the excitation mechanism and the anharmonicities of the vibrational modes [90, 91]. The absorption of a single IR photon is followed by fast intramolecular vibrational redistribution (IVR), depositing the energy in the bath of vibrational states and thereby effectively de-exciting the absorbing mode. This process can occur many times during a FELIX macropulse. During that process, the internal energy rises, causing anharmonicities to induce a red-shift of the absorbing mode. When the anharmonicity is large, the red-shift can limit the number of absorbed photons, as the mode shifts out of the bandwidth of the IR radiation. A very similar excitation mechanism is applicable here as well. The relative intensities in the experimental spectrum are influenced and it is possible that the modes that are not observed have a low oscillator strength and/or a large anharmonicity.

Another related effect is important too. For the low frequency mode, theory predicts two closely spaced bands. During the excitation process, while the lower frequency mode shifts out of the laser bandwidth, the mode to the blue side can shift into resonance [58]. Thus, while theory predicts the $T_{1u}(2)$ mode to be not much more intense than some of the modes that are not observed, the weaker mode to the blue can result in an enhancement of the excitation efficiency. In addition, in the IR-REMPI experiment, FELIX was tightly focussed leading to a very high IR intensity. In the present experiment, due to constraints imposed by the experimental setup, FELIX is not focused as tightly, resulting in an at least one order of magnitude lower radiation intensity. We expect that, under higher intensity conditions, more bands will be observable.

The anharmonicities, together with the spectral profile of the IR laser, determine not only the excitation efficiencies but also influence the peak positions and shapes in the spectrum. Modeling shows that peaks can have an asymmetry and a tail on their red side [90, 91]. In agreement with that, one can notice that the most intense peak at 1374 cm$^{-1}$ is slightly asymmetric. More important here is the possibility of peak shifts. However, it is not likely that this effect is dramatic, as this shift is expected to be less than the spectral width of the excitation laser, which is 0.5 - 1% FWHM of the center frequency in the present experiment.

For $C_{60}$ anions in the solid, it is observed that the vibrational modes shift to the red with increasing negative charge (see Table 2.1; references are given there). This shift is small for the $T_{1u}(2)$ mode around 570 cm$^{-1}$, but large for the $T_{1u}(4)$ mode, which shifts from 1392 cm$^{-1}$ for $Rb_1C_{60}$ to 1340 cm$^{-1}$ for $Rb_6C_{60}$. The value of 1374 cm$^{-1}$ reported here for the $T_{1u}(4)$ mode is slightly lower than the values reported for $C_{60}$ with nominally one negative charge. In experiments on solid state materials, the charge transfer to $C_{60}$ might be incomplete and moreover, $C_{60}$ anions might be significantly perturbed by the surrounding cations. In the present experiments, the charge on $C_{60}$ is $-1$ and the ions are not perturbed by any surroundings. In agreement with that, the here presented value for the $T_{1u}(4)$ of $C_{60}^-$ is closest to the value reported for solid state materials which are designed to limit the interactions of $C_{60}$ anions with the surroundings [84].

The mid-IR spectrum of gas-phase $C_{70}^-$ together with a calculated spectrum ($C_{70}^-$ in $C_s$ symmetry, scaled by a factor of 0.9648 [118]) are shown in Figure 2.5. Most parts of the spectrum are only recorded once which are the red marked regions: [395 - 512] cm$^{-1}$, [555 - 1344] cm$^{-1}$, and [1440 - 1556] cm$^{-1}$. Between [512 - 555] cm$^{-1}$ and [1344 - 1440] cm$^{-1}$ several scans are recorded and averaged.
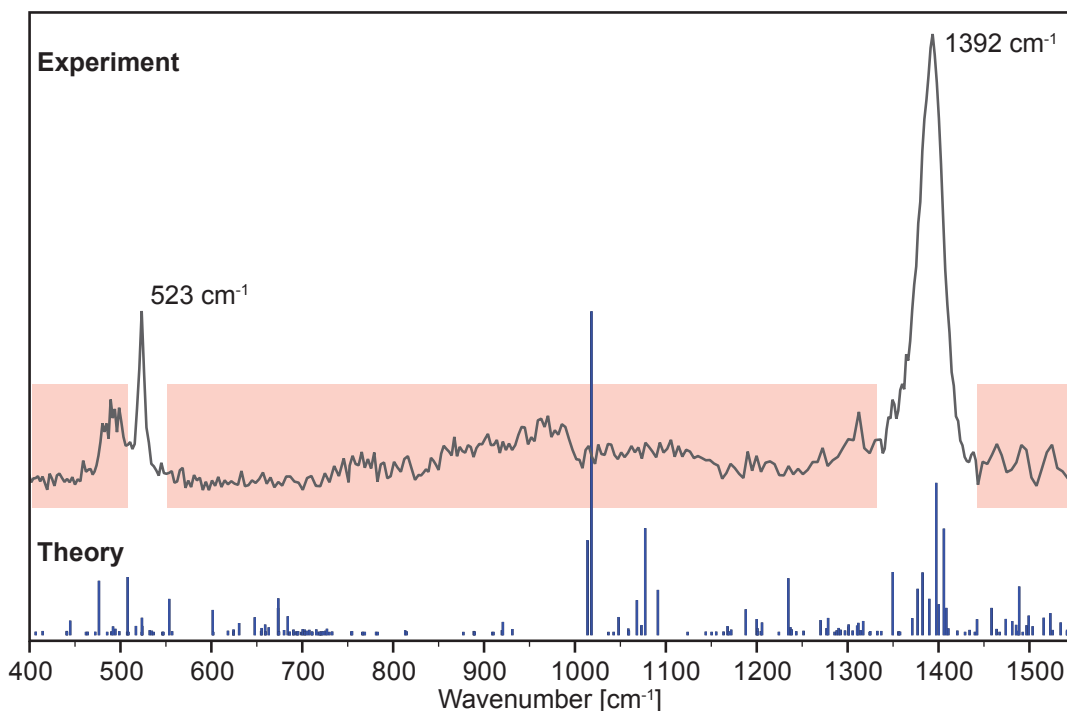


**Figure 2.5:** Experimental IR spectrum of $C_{70}^-$, recorded by measuring the Cl$^-$ ion yield as a function of the wavelength. The scaling of the ordinate is linear. No signal of chloride ions is obtained on the baseline. Only one scan is taken in the red marked regions. A theoretical spectrum for $C_{70}^-$ in $C_s$ symmetry is shown as stick spectrum.

When being in $C_{2v}$ symmetry, only saddle point structures were obtained in the calculation but in $C_s$ symmetry minima in the potential energy landscape of the molecule were found. The calculated total energy of the anion in $C_s$ symmetry is 2.38 eV below the energy calculated for the neutral molecule.

At least two peaks are clearly present in the experimental infrared spectrum. One is observed at 1392 cm$^{-1}$ with a full width at half maximum (FWHM) of 31 cm$^{-1}$ and the second peak is observed at 523 cm$^{-1}$ (FWHM of 6 cm$^{-1}$). The peak structure at the lower frequency side of the latter peak seems to be a real feature, nevertheless an absolute statement cannot be given here because only one scan is available in this region. The peak at higher frequency found in the experimental spectrum is in good agreement with peaks present in the calculated spectrum. Unfortunately, some infrared transitions (for example the strong transition at 1017 cm$^{-1}$) observed in the theoretical spectrum are not found in the experimental one. This might be due to a strong anharmonicity of these bands (similar observations are found in the infrared spectrum of $C_{60}^-$, see above). The observed peak at 523 cm$^{-1}$ is slightly blue-shifted with respect to peaks found in the calculations. Possibly, a different scaling factor of the calculated infrared transitions better suitable for low frequency modes should be used. In theoretical studies on the evaluation of harmonic vibrational frequency scaling factors, a factor of 1.0117 for calculated vibrational frequencies in the low frequency range is found to being more accurate (for the used functional and basis set) [118].

In neutral $C_{70}$, a band found at 1430 cm$^{-1}$ is assigned to an $E_1'$ mode [119]. When $C_{70}$ is embedded in a salt structure and expected to be negatively charged, this band is shifted down to 1423 cm$^{-1}$ [102]. The here observed band of isolated $C_{70}^-$ in the gas phase at 1392 cm$^{-1}$ can therefore be assigned to the $E_1'$ mode. According to a force-constant model this mode should be further shifted down to 1387 cm$^{-1}$ for $C_{70}^{6-}$ in solid $M_6C_{70}$, where M represents an alkali-metal atom [119]. The here observed red-shift when going from $C_{70}$ to $C_{70}^-$ is thus in general agreement with those calculations. The observed stronger red-shift of the band for gas-phase $C_{70}^-$ with respect to the salt embedded molecule might be due to the differences in the environment and further, the charge of the molecule in the gas phase is well defined whereby the charge state of the salt embedded molecule is more undefined.

Unfortunately, the published mid-IR spectrum of the fullerene embedded in salt does not cover the range of the other vibrational mode which is found in the spectrum of gas-phase $C_{70}^-$. Comparing the observed low frequency mode to calculated and observed modes for neutral $C_{70}$, the vibration observed in the gas-phase infrared spectrum of $C_{70}^-$ at 523 cm$^{-1}$ is red-shifted. The mode is found in the spectrum of the neutral molecule at 534 cm$^{-1}$, in good agreement with a calculated vibration of symmetry type $E_1'$ at 533 cm$^{-1}$ [112]. Another mode of symmetry type $E_1'$ is found at 509 cm$^{-1}$ (the calculated position is 507 cm$^{-1}$). However, it seems reasonable that the origin of the observed mode in the anion comes from the higher frequency mode. The red-shift with respect to the mode for $C_{70}$ might be due to the anionic character of $C_{70}^-$ and the solvent free environment as well as potential anharmonicities of this band.

## 2.4   Conclusions

The infrared spectrum of isolated room temperature $C_{60}^-$ and $C_{70}^-$ are presented. For $C_{60}^-$, two vibrational modes are found at 570 cm$^{-1}$ and 1374 cm$^{-1}$. The bands correlate

to two of the four IR allowed transitions in neutral $C_{60}$. The results show that the higher frequency band is strongly shifted from its position in neutral $C_{60}$ and is a sensitive marker for the charge state. The measured spectrum is in good agreement with calculated spectra for Jahn-Teller distorted $C_{60}^-$ with $D_{3d}$ symmetry. The band positions and intensities are compared to results obtained by theory as well as to known bands in solid samples and good agreement is found.

In the mid-infrared spectrum of $C_{70}^-$, two vibrational modes at 523 $cm^{-1}$ and 1392 $cm^{-1}$ are found. The measured spectrum is in agreement with calculated spectra of Jahn-Teller distorted $C_{70}^-$ with $C_s$ symmetry. Not all calculated vibrational modes are found in the experimental infrared spectrum. However, it should be noted that most parts in the experimental spectrum are only measured once. To improve the accuracy of the experimental spectrum, more measurements in a wide spectra range need to be performed and the results need to be compared to the calculated infrared vibrations. The shift of the higher frequency mode compared to its position found in neutral $C_{70}$ is in agreement with published results based on theory (force-constant model) and spectroscopic results on $C_{70}$ embedded in salt.

To conclude this chapter, infrared spectroscopy on isolated systems in the gas phase can be a powerful tool to investigate structures of molecules. In this chapter through the investigation of well defined systems with high symmetry, the efficiancy of infrared spectroscopy has been demonstrated. In the following chapters, however, this technique will be used to investigate the structures of biologically relevant systems. Naturally, these more complicated systems do not possess the same high level of symmetry and therefore, will present a greater challenge to the method of infrared spectroscopy in investigation of biological molecules.

# Chapter 3

# Gasphase IR spectra of proteins: Cytochrome c and tendamistat

In this Chapter, spectra in the mid-infrared range of real biological macromolecules are shown and interpreted. Two different proteins are selected: cytochrome c from equine heart and tendamistat. In solution, cytochrome c exists predominantly in an $\alpha$-helical conformation while tendamistat is found to have a $\beta$-sheet rich secondary structure. The question arises whether these structural elements can be maintained upon the transition from the solution phase to the gas phase. If these structures are preserved in the solvent molecule free environment, is it possible to measure and assign structural elements to the molecule? To that end, mass spectra and infrared spectra of gas-phase cytochrome c – protonated and deuterated – in various charge states are recorded and analyzed. Tendamistat is investigated in one single charge state in the gas phase and the observed spectrum is analyzed.

## 3.1   Introduction

### 3.1.1   Cytochrome c: a perfect model system for a protein in mostly $\alpha$-helical conformation?

Cytochromes are 'haemoproteins whose principle biological function is electron and/or hydrogen transport by virtue of a reversible valency change of their haem iron' [120, 121]. Their major role is to carry electrons of molecular oxygen in the respiratory chain [11]. Further, it recently has been discovered that cytochrome c plays an important role as a mediator in apoptosis – the programmed cell death [122, 123]. There are four groups of cytochromes (a to d), of which the group c, examined in this experiment, is characterized by covalent linkage between the heme group and the protein [120, 121]. All cytochromes with prosthetic groups[1] linked this way belong to group c. The exact chemical composition of cytochrome c (especially the type of the amino acids in the primary structure) varies between different species, sometimes only by one amino acid. Those differences however, can be significant, as will be discussed below.

---

[1]tightly-bound cofactors

Equine heart cytochrome c consists of 104 amino acids and a single heme group covalently linked to $Cys_{14}$ and $Cys_{17}$. The average isotopic mass of the protein is 12384 amu. The complete amino acid sequence was derived in 1961 [124]. In 1967 the tertiary structure of a mammal cytochrome c was obtained using X-ray diffraction [125] supporting the postulated position of the heme iron [126]. These results were confirmed by nuclear magnetic resonance (NMR) studies in 1969 [127].

Cytochrome c is a molecule that has been the subject of intense investigation in solution phase (for example resonance raman spectroscopy [128]) and in the gas phase (for example ion mobility studies [129], see Section 3.4.2 for details). The primary structure[2] is found in one letter code in the footnote[3]. 42 amino acids in equine heart cytochrome c form $\alpha$-helices, four amino acids are involved in a turn and the remaining amino acids are orientated randomly [130]. As a result, the protein adapts a mostly $\alpha$-helical conformation and hence serves as a good model system for a helical secondary structured protein. In Figure 3.1 two different representations of equine cytochrome c derived from NMR studies and X-ray diffraction are shown.
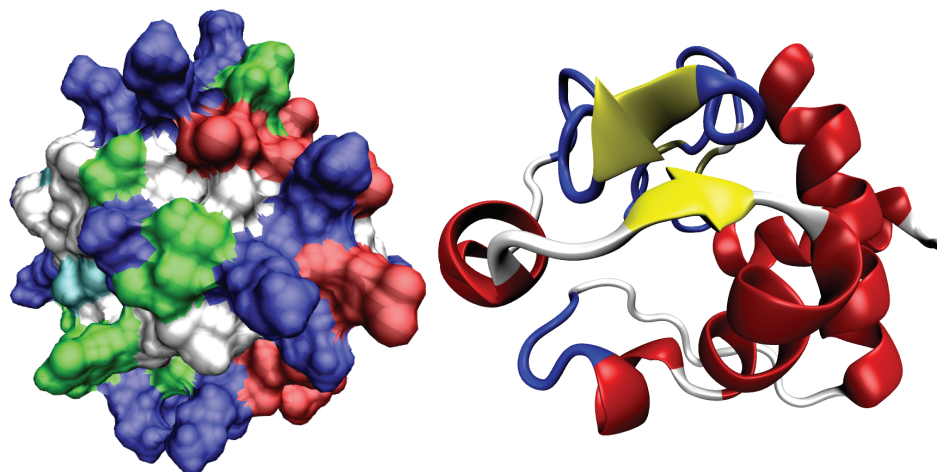


**Figure 3.1:** The Figure on the left shows an NMR structure of cytochrome c in solution as a surface plot representation. Basic side chains are shown in blue, acidic side chains in red, polar side chains in green, nonpolar side chains in white and unassigned side chains in cyan. This structure is taken from the protein data base (pdb), #1OCD. The Figure on the right is derived from a crystallized structure via X-ray diffraction (#1HRC at pdb) with a resolution of 1.90 Å. Shown is only the peptidic backbone with secondary structure elements in a New Cartoon representation. The $\alpha$-helical conformation is presented in red, the yellow arrows, pointing from the N- to the C-terminus, represent a $\beta$-sheet conformation, turns are marked in blue, and random coil motifs are shown in white.

---

[2]the amino acid sequence
[3]GDVEK GKKIF VQKCA QCHTV EKGGK HKTGP NLHGL FGRKT GQAPG FTYTD ANKNK GITWK EETLM EYLEN PKKYI PGTKM IFAGI KKKTE REDLI AYLKK ATNE

### 3.1.2   Tendamistat: a good example for investigation of a $\beta$-sheet conformation in the gas-phase?

The second protein studied here, tendamistat, is an $\alpha$-amylase inhibitor, characterized by its tight binding to mammalian $\alpha$-amylase [131]. $\alpha$-Amylase is an enzyme in the pancreas that hydrolyses dietary starch into di– or trisaccharides [11]. The primary structure of the 74 amino acid long protein with an average isotopic molecular mass of 7958 amu was discovered in 1981 [132] and the secondary structure of a truncated version in solution was first measured in 1985 using NMR [133]. The complete three-dimensional structure was measured by NMR in solution in 1988 [134] (protein data base entry #2AIT), a low resolution (2.0 Å) crystallized structure at room temperature was observed via X-ray diffraction in 1986 [135] (protein data base entry #1HOE) and a high resolution (0.93 Å) crystal structure at 100 K was obtained in 2003 [136] (protein data base entry #1OK0). Figure 3.2 depicts the basic crystal structure corresponding to the protein data base entry #1OK0. Tendamistat is of interest in this work due to its large $\beta$-sheet proportion and the complete absence of an $\alpha$-helical secondary structural motif in the condensed phase. Depending on the algorithm used to characterize and assign structural elements, the amount of amino acids occupying a $\beta$-sheet structure ranges from 33 residues in six strands (corresponding to 44% of the protein [137]) to 36 residues in eight strands (corresponding to 48% of the protein [138]). The question arises, whether some amount of a $\beta$-sheet fraction in the backbone of the molecule is also conserved in the gas phase and if infrared spectroscopy is capable to characterize this structural element.
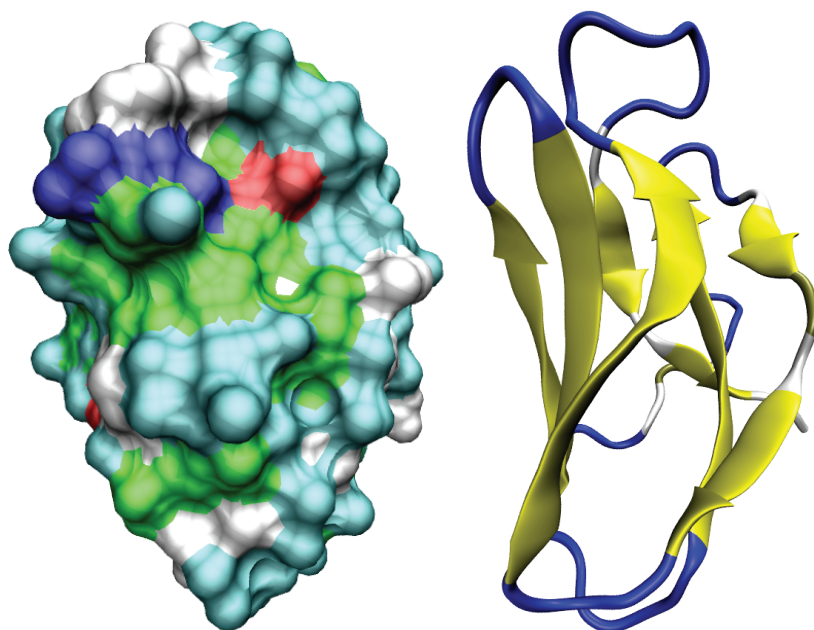


**Figure 3.2:** The Figure on the left shows the structure in a surface representation of crystallized tendamistat (pdb #1OK0). The Figure on the right shows only the peptidic backbone in a New Cartoon representation with the secondary structure elements. The color code and symbols are described in the caption of Figure 3.1.

## 3.2 Experimental

In this Section, the experimental conditions to produce gas-phase ions from a solution of the investigated proteins are described. The parameters and specific settings of the FT ICR apparatus to obtain mass spectra of the molecules are presented and the mechanism to record infrared spectra of those proteins is explained.

### 3.2.1 Experimental conditions to obtain mass- and infrared spectra of cytochrome c

50 $\mu$M cytochrome c in a solution containing 30% water, 70% methanol (MeOH) and 1% acetic acid (AcOH) is used to obtain charge states ranging from 9+ to 16+. In addition, KCl in a concentration of 1 mM is added to allow for binding of potassium ions to the charged proteins, thus forming potassiated clusters. All chemicals are purchased from Sigma Aldrich and used without further purification. The high amount of MeOH could lead to a higher $\alpha$-helical structural fraction in cytochrome c in the solution compared to lower MeOH concentrations. This is called the helical denatured state of cytochrome c and a tertiary structure is absent as observed by Kamatari *et al.* [139]. In order to get mass spectra of low charge states (7+ to 10+) with an intensity allowing for recording infrared spectra, cytochrome c is dissolved to a concentration of 50 $\mu$M in a solution containing 1 mM ammonium acetate and 1 mM KCl with 3% MeOH following specifications by Konermann *et al.* [140]. To optimize the ion signal intensity in the mass spectrometer of charge states higher than 8+, a small amount of AcOH is added. It is found that, as a consequence, the signal intensity of the protonated mass peaks is increased but the potassiated mass peaks disappear. This might be due to the donation of protons to the molecule by the acid, which suppresses the binding of a potassium ion. Adding more KCl to increase its concentration in the solution shows no effect but by raising the amount of MeOH up to a quantity of approximately 50% of the solution volume, the ion signal intensity enhances dramatically and charge states from 9+ to 15+ are observed in the mass spectrometer.

To record mass spectra of highly charged molecules a method following Iavarone and Williams is employed [141]: Cytochrome c (50 $\mu$M) is dissolved in a solution containing 50% H$_2$O, 49% MeOH, 1% m-(3)-nitrobenzyl alcohol (m-NBA), 0.2% AcOH, and KCl (1 mM). In the supercharging mechanism postulated by Iavarone and Williams, the charge state of cytochrome c (or other proteins) increases upon adding a low volatile, high surface tension solvent – in this case m-NBA is used – to denaturing solvents with high volatility as observed in their experiment. The resulting surface tension of the solvent system is increased and therefore, the higher charge density on the surface of those droplets makes more charges available to the analyte ions [141]. Cytochrome c ions up to charge state 19+ are observed. However, when too much water is present and the solvent system is not in a denaturing condition for the protein, adding m-NBA lowers the surface tension. Therefore, higher charge states are not expected in the supercharging mechanism. The solvent conditions used here are slightly different than in [141] (in this study, only 0.2% AcOH is used compared to 3% in the published article) and the degree of the denaturing condition for the protein is not

clear. However, in experiments done by other groups, where m-NBA is added to non denaturing protein solvents [142, 143], an enhancement in the charging of the ions is also observed. These results suggest an additional mechanism involving direct interaction of the protein with the reagent in addition to factors such as the surface tension.

Whilst a thorough presentation of the apparatus used in this experiment is given in Chapter 1, Section 1.4.4, the specific settings and important mechanisms relevant to the systems investigated here are described below.
The protein samples are electrosprayed by a syringe pump at a flow rate of 30 $\mu$l/min. The desolvation gas temperature ($N_2$) is set to 100° C with a flow rate of 200 l/min and the voltage at the ESI needle with respect to ground is set to 3 kV. The cone voltage is applied at the entrance cone with respect to the differentially pumped part of the apparatus. Fragmentation of the protein can occur in the high pressure region of the FT ICR mass spectrometer and should be avoided. The cone voltage effects the fragmentation[4] and also influences the charge state of sprayed molecules in the FT ICR. Therefore, it is essential that the cone voltage is optimized for each system separately. The cone voltage is set to 198 V for measuring the low charge states and is decreased to about 90 V for the observation of higher charge states. Once in the apparatus, the ions are guided in radio frequency (RF) ion guides and accumulated via a small electric longitudinal field applied between the entrance and the exit in a hexapole. The accumulation voltage and time are also important parameters because fragmentation can occur in the trap and optimal parameters vary again for different molecules. The 'Gert-box'[5] is set on a potentiometer to the value 4.0 and the accumulation time is set around 500 ms for obtaining the highest cytochrome c ion signal intensity in the mass spectrometer. Alternatively, a 5.2 ms long $N_2$ gas-pulse is introduced by a leak valve into the ICR cell to collisional cool and thereby stop the ions. One of these techniques to inject the ions into the ICR cell is used depending on the highest achievable ion signal intensity. Mass spectra of cytochrome c are recorded after a trapping time of the molecules of 4 s.

Figure 3.3 shows a measured mass spectrum of protonated cytochrome c possessing various charge states. The mass distribution of the cytochrome c clustered with potassium is clearly visible. The most intense peak for each charge state is due to the purely protonated cytochrome c ion in this charge state. The decaying mass distribution on the higher mass side of those peaks is due to the potassium ions that are (non-covalently) bound to the molecule. The spacing of this underlying distribution is exactly the mass of a potassium ion (39.1 amu) divided by the charge state.

To record mass and IR spectra of deuterated cytochrome c, the same methods are applied, however this time using only deuterated solvents. The cytochrome c solution is allowed to incubate 48 h prior to the recording of infrared spectra to ensure complete H/D exchange of all possible hydrogen atoms. Table 3.1 gives an overview of the observed masses and the mass shift between the protonated and deuterated cytochrome c ion in the different charge states. This shift indicates a complete deuteration of the

---

[4]and therefore the ion signal intensity of the molecule, which is investigated
[5]the setting corresponds to the trapping depth
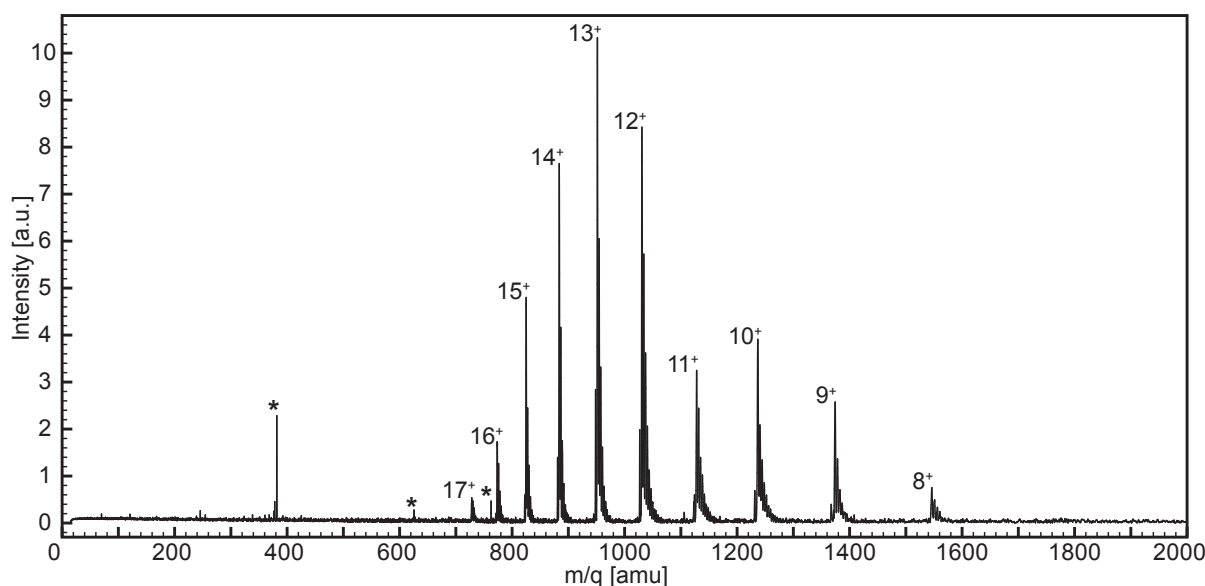
**Figure 3.3:** Mass spectrum of protonated cytochrome c ions in charge state 8+ (1546 amu) to 17+ (728 amu). Noise peaks are marked by an asterisk (see Section 1.4.4).

100 backbone N−H groups (this group is exchanged in all amino acids except in the four proline residues that are present) and of exchangeable side chain groups. In the charge state 12+ of the protein, where all basic and all acidic side chains are expected to be fully protonated (deuterated), three side chain protons are exchanged in Lys (19), two in His (3), six in Arg (2), one in Asp (3), one in Glu (9), two in Gln (3), and two in Asn (5). The number in brackets is the amount of the residue present in cytochrome c. Including a protonated C-terminus (the only proton of the protected N-terminus is already included) a total number of 204 exchanged protons upon deuteration in the charge state 12+ of the protein is possible.

**Table 3.1:** Observed masses and mass shifts between protonated and deuterated cytochrome c ions in different charge states.

| Charge state | Protonated m/q [amu] | Deuterated m/q [amu] | Mass shift [amu] |
|---|---|---|---|
| 8+ | 1545.96 | 1570.57 | 197 |
| 9+ | 1374.04 | 1395.92 | 197 |
| 10+ | 1236.81 | 1256.81 | 200 |
| 11+ | 1124.22 | 1142.37 | 200 |
| 12+ | 1030.76 | 1047.68 | 203 |
| 13+ | 951.56 | 967.21 | 203 |
| 14+ | 883.75 | 898.24 | 203 |
| 15+ | 824.81 | 838.49 | 205 |

In Figure 3.4 two different mass spectra of deuterated cytochrome c in different charge states are shown. The spectrum presented in black shows the low charge states of

cytochrome c and is recorded using an ammonium acetate solution as described earlier in this Section. The spectrum shown in red is taken by adding m-NBA to the solvent (see page 42) and higher charge states of cytochrome c molecules are present. The peak at 425 amu corresponds to an unidentified ion signal whereas the peaks at lower masses mostly correspond to electromagnetic interferences caused by the apparatus (see Section 1.4.4).
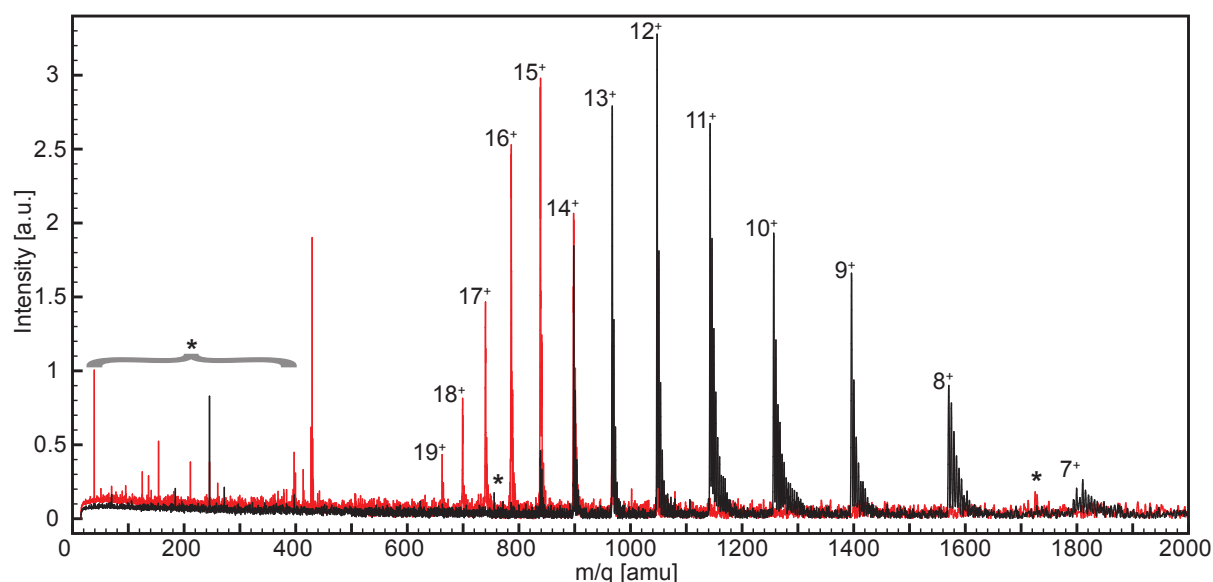


**Figure 3.4:** Mass spectrum of deuterated cytochrome c ions in charge state 7+ (1807 amu) to 19+ (666 amu). Noise peaks are marked by an asterisk (see Section 1.4.4).

High-intensity IR radiation from FELIX can enter the ICR cell via a ZnSe window and is reflected inside the cell by its highly polished copper walls to maximize the overlap with the ion cloud. IR spectra of cytochrome c are recorded by taking advantage of IR-MPD (see Section 1.4.1). The non-covalently bound potassium ion(s) can be dissociated from the protonated (or deuterated) cytochrome c cluster at resonant IR radiation. As a consequence, mass peaks of protonated (or deuterated) cytochrome c at lower charge states (typically lowered by one charge carrier) appear in the mass spectrum. The ion signal intensity of the parent ion decreases.

In Figure 3.5, two mass spectra of charge state 13+ of protonated cytochrome c with non-covalently bound potassium ions are shown. To record the spectrum on the left, the wavelength of FELIX is set to 5.5 $\mu$m and after an irradiation time of about six seconds, a mass spectrum is recorded. Only the parent ion signal is present (here [cytochrome c + 13 H]$^{13+}$, [cytochrome c + 12 H + 1 K]$^{13+}$, etc.). The right side of the Figure shows a mass spectrum, in which the protein is irradiated about six seconds with resonant IR radiation (FELIX is set to $\approx$ 6$\mu$m) before taking the mass spectrum. The parent ion signal is depleted by about one third and additional mass peaks are observed in the mass spectrum. The peak at 1030 amu corresponds to charge state 12+ ([cytochrome c + 12 H]$^{12+}$, [cytochrome c + 11 H + 1 K]$^{12+}$, etc.). The signal appearing at 1125 amu is due to charge state 11+ ([cytochrome c + 11 H]$^{11+}$,

[cytochrome c + 10 H + 1 K]$^{11+}$, etc.). This is an indication that either one or two potassium ions are cleaved from the parent molecule complex. The ion signal intensity of the double potassium loss is low compared to the signal corresponding to the loss of one charge carrier. The mass spectra in Figure 3.5 are recorded using excellent experimental conditions (long transient time and many times averaged). However, to obtain infrared spectra of the charged protein compromises concerning the measurement time need to be made and not as many mass spectra are taken and averaged for each wavelength step. As a consequence, in the mass spectra of the ions recorded after irradiation the mass peak corresponding to the loss of two charge carriers of the parent ion is not observed. Every
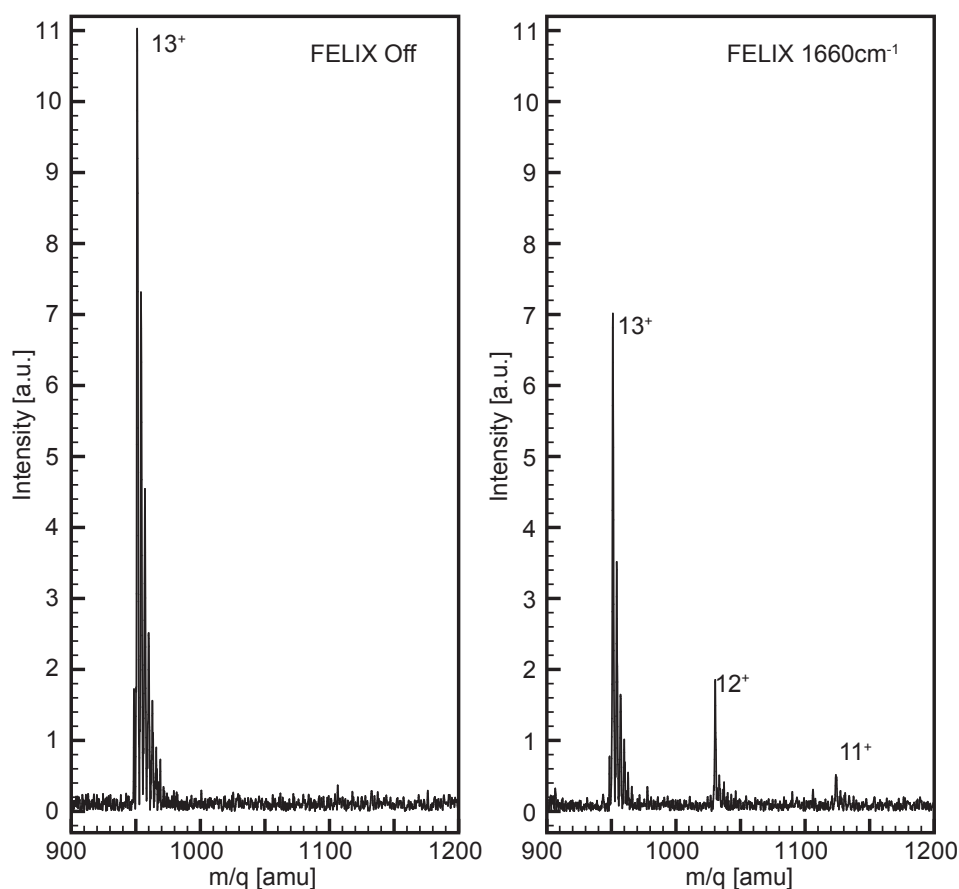


**Figure 3.5:** Mass spectrum of isolated protonated cytochrome c ions in charge state 13+. On the left, the wavelength of FELIX is set to 5.5 $\mu$m prior to taking the mass spcectrum and on the right, the wavelength is 6 $\mu$m. Upon radiation with light resonant to an IR transition of the molecule, one or two potassium ions are dissociated from the complex and new charge states of cytochrome c appear in the mass spectrum. Also the signal intensity of the parent ion in the mass spectrum on the right is decreased about one third compared to the previous signal strength.

other charge state is isolated in the ICR cell to allow for recording of more charge states simultaneously. To record infrared spectra for example of evenly charged cytochrome c ions, all oddly charged ions are ejected from the mass spectrometer. Upon resonant irradiation, one K$^+$ ion is detached from the cytochrome c complex and the intensity of the mass signal in the mass spectrometer for this complex decreases. Thereby, the charge

state of the resulting cytochrome c complex is lowered by one and an ion signal appears at the position of the previously ejected molecules in the mass spectrum. By tuning the wavelength and recording the corresponding mass spectra, IR spectra can be generated. A general introduction on recording IR spectra of gas-phase ions is given in Section 1.4.2.

To record the IR spectra, four mass spectra are taken and averaged for each wavelength step of FELIX. The irradiation time is set to four seconds. At least two IR scans for each species are recorded over the whole wavelength range and averaged to obtain the infrared spectra (see Figure 3.8 and 3.9). Table 3.2 gives an overview of simultaneously recorded infrared spectra with different charge states of protonated/deuterated cytochrome c. The symbol # indicates the number of the scans. If the index is set to two, two scans are taken one after another. The index following the day represents the chronological order of scans and x indicates the recorded charge state. The upper table displays the scan sequence of the protonated species and the lower that of deuterated ions.

**Table 3.2:** Overview of simultaneously recorded IR spectra for different charge states of protonated (top) and deuterated (bottom) cytochrome c molecules. In the first column, the day together with an index is recorded. The index marks the chronological order of recording the infrared spectra on that day. With x, (simultaneous) recorded charge states are marked. The number of recorded scans with the 'set' of charge states is marked with #. A two corresponds to recording two scans directly one after another.

| Date | 7+ | 8+ | 9+ | 10+ | 11+ | 12+ | 13+ | 14+ | 15+ | 16+ | # |
|------|----|----|----|-----|-----|-----|-----|-----|-----|-----|---|
| Day1/1 |   |   |   | x |   | x |   | x |   | x | 2 |
| Day1/2 |   | x |   |   | x |   | x |   | x |   | 2 |
| Day1/3 |   | x |   |   |   |   |   |   |   |   | 2 |
| Day2/1 |   | x |   |   |   |   |   |   |   |   | 1 |
| Day2/2 | x |   |   |   |   |   |   |   |   |   | 1 |
| Day2/3 | x |   | x |   |   |   |   |   |   |   | 1 |
| Day2/4 |   | x |   |   |   |   |   |   |   |   | 1 |
| Day2/5 | x |   | x |   |   |   |   |   |   |   | 1 |

| Date | 8+ | 9+ | 10+ | 11+ | 12+ | 13+ | 14+ | 15+ | 16+ | 18+ | 19+ | # |
|------|----|----|-----|-----|-----|-----|-----|-----|-----|-----|-----|---|
| Day2/6 |   |   | x |   | x |   | x |   |   |   |   | 2 |
| Day2/7 | x |   | x |   | x |   | x |   | x |   |   | 2 |
| Day3/1 | x |   |   |   |   |   |   |   |   |   |   | 2 |
| Day3/2 |   | x |   | x |   | x |   |   |   |   |   | 2 |
| Day4/1 |   |   |   |   | x |   | x |   | x | x |   | 2 |
| Day4/2 |   |   |   | x |   | x |   | x |   |   | x | 2 |

### 3.2.2 Experimental conditions to record mass- and infrared spectra of tendamistat

Tendamistat is kindly provided by Dr. Jürgen Pünter (sanofi aventis) and used without further purification. All other chemicals are purchased from Sigma Aldrich. To obtain mass spectra in the FT ICR mass spectrometer, a solution containing 100 $\mu$M protein with 10% MeOH and 0.1% AcOH is electrosprayed. The needle voltage is set to 3 kV above the ground potential, the cone voltage to 160 V, and a desolvation gas ($N_2$) temperature of 80° C is used.

Out of the 74 amino acids in tendamistat, there are eight acidic (five aspartic acid and
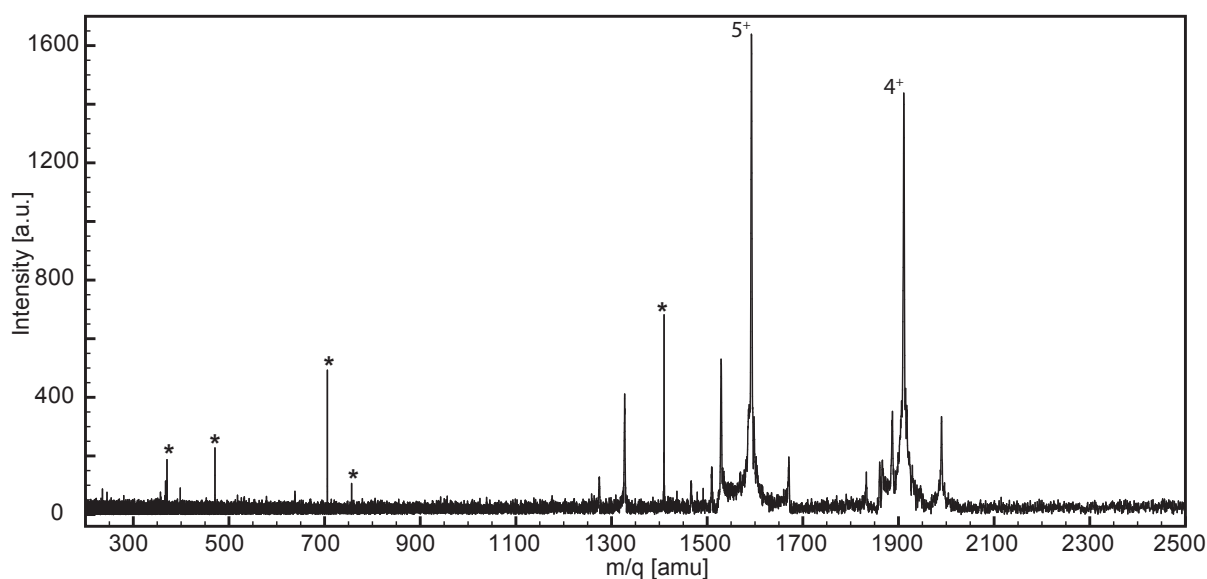


**Figure 3.6:** Mass spectra of tendamistat in the range from 200 amu to 2500 amu. Noise peaks are marked by an asterisk (see Section 1.4.4).

three glutamic acid) and six basic (three arginine, one lysine, and two histidine) residues. Assuming that the charge distribution is explained by the conformation-dependent neutralization theory (CDNT), the expected upper limit of the net charge of the protein would be minus two (corresponding to the difference between basic and acidic residues in the primary structure [144]). However, this does not seem to be the case. In Figure 3.6 a mass spectrum of tendamistat is shown and the charge states 4+ and 5+ are observed. The presence of these charge states can be explained by the Rayleigh limiting charge theory (RLCT) [34], where the maximal charge carried by a spherical ion (which is a good approximation for folded proteins) corresponds to the upper limiting charge from spherical droplets. Table 1 published in reference [145] gives an overview of expected charge states for different proteins by comparing the values estimated using CDNT and RLCT to measured results. For tendamistat, CDNT would lead to a charge state of -2 and RLCT to +6.

To record an infrared spectrum of tendamistat, the charge state 5+ of the protein is isolated in the mass spectrometer (SWIFT on charge state 4+) and the molecules are
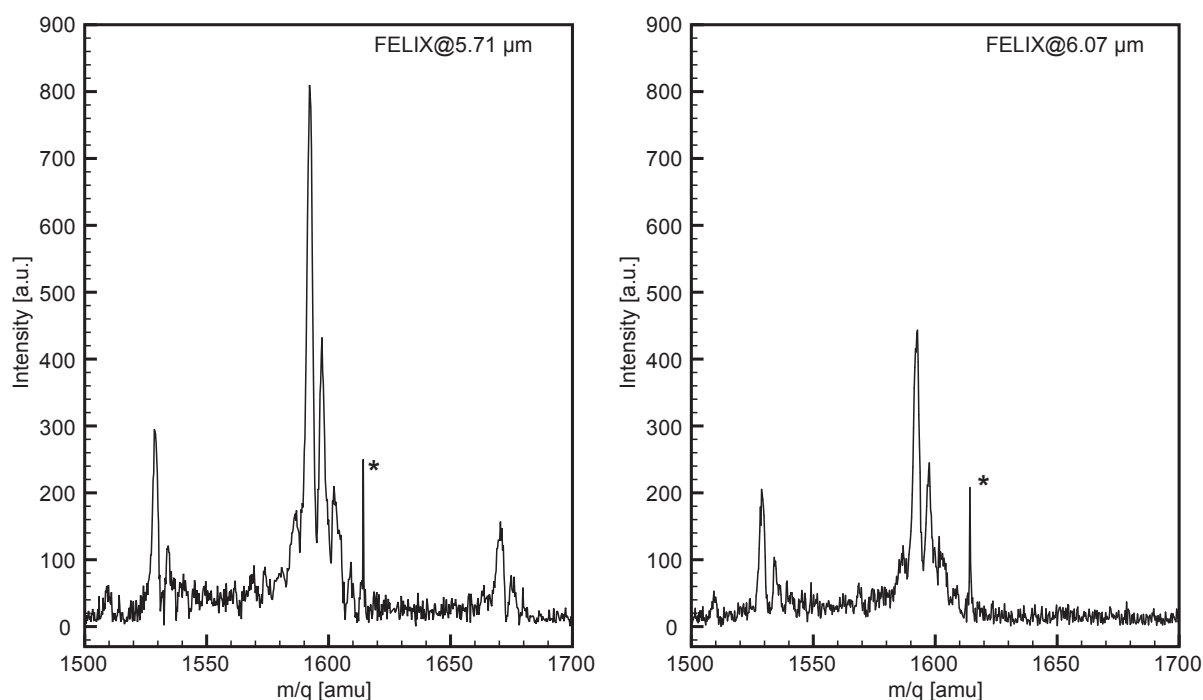
**Figure 3.7:** Mass spectra of tendamistat in the range from 1500 amu to 1700 amu. The left Figure is recorded after three seconds of irradiation with FELIX set to a wavelength of 5.71 $\mu$m (corresponding to 1750 cm$^{-1}$). No IR resonant transitions in the protein at this wavelength are observed. The right Figure is taken after a irradiation time of three seconds with FELIX at 6.07 $\mu$m (corresponding to 1647 cm$^{-1}$). An infrared resonant transition in the molecule at this energy is present – indicated by the decreased intensity of the parent ion signal. Noise peaks are marked by an asterisk (see Section 1.4.4). Further details of the presented spectra are given in the text.

stored there and irradiated with FELIX light for three seconds. Figure 3.7 shows two mass spectra after three seconds of illumination with infrared light at two different wavelengths. The observed mass distribution is due to sodium ions that are non-covalently bound to the protein. The sodium ions are present in the mass spectrum due to impurities on or in the glass containers used to prepare the solution. It was not possible to isolate the desired mass peak solely (1592 amu with its isotopic/sodiated distribution) from masses observed at 1529 amu and 1670 amu due to an otherwise loss in the signal intensity. The molecules corresponding to the latter two peaks in the mass spectrum are not assigned and might be due to impurities of the tendamistat sample. The peak in the mass spectrum at mass 1597 amu corresponds to [tendamistat + 4 H + 1 Na]$^{5+}$ and at 1602 amu [tendamistat + 3 H + 2 Na]$^{5+}$, respectively. Upon infrared radiation resonant with a transition in the charged protein a decrease in signal intensity is observed. The mass peak at 1670 amu completely vanishes and the intensity of the mass peak located at 1529 amu is decreased by one third. The five fold protonated parent ion signal in the mass spectrum is decreased by a factor of two. An interesting observation is that the relative intensity of the mass peak corresponding to the sodiated complex with respect to the protonated one is the same in both mass spectra. Unfortunately, no fragment

ions are observed at IR resonant conditions[6]. It is possible that the produced fragments are distributed over a wide range in the mass spectrum with the consequence that the ion signal intensity of the resulting fragments is below the threshold of the FT ICR ion signal detection sensitivity. Further optimization and increasing the parent ion signal intensity along with recording and averaging more mass spectra should allow for detection of resulting fragments upon a resonant transition in the mass spectrometer but was beyond the time schedule for this project.

## 3.3 Results

Infrared spectra of protonated and deuterated cytochrome c ions in various charge states and of tendamistat in one charge state are presented in this Section. Observations in the spectra are pointed out and vibrational bands are assigned.

### 3.3.1 Infrared spectra of protonated and deuterated cytochrome c in various charge states

Figure 3.8 shows infrared spectra of the protonated protein cytochrome c from equine heart in several charge states. The number x (x = n + m) on the left side of each individual spectrum indicates the charge state and is read as [cytochrome c + nH + mK]$^{x+}$. The presented spectra represent significant improvement in the signal-to-noise ratio in comparison to previous work done on cytochrome c from bovine heart in the mid-infrared region [146]. A detailed discussion comparing the previously published results with those presented here follows later in the text (see Section 3.4.3). The band assignment is not done by fitting Gaussian functions (explanations see Section 3.4, page 57)

In the infrared spectra ranging from charge state 7+ to 11+ two broad features are clearly present. The peak around 1660 cm$^{-1}$ can be assigned to the C=O stretching vibration (amide-I band). This band is found in solution usually between 1625 cm$^{-1}$ and 1690 cm$^{-1}$ (see Section 1.3) depending on the secondary structure of the protein. The peak around 1540 cm$^{-1}$ can be assigned to the N−H bending vibration (amide-II band), which is observed in solution between 1540 cm$^{-1}$ and 1550 cm$^{-1}$. The peaks are relatively broad (FWHM around 50 cm$^{-1}$) and the amide-II band seems to be slightly asymmetric with the steeper side to higher wavenumbers. The intensity of the amide-I band is generally higher compared to the amide-II band (except in the spectrum of the charge state 16+). There is a common trend for a small blue-shift (around 10 cm$^{-1}$) of the position of the amide-I band with increasing charge state. This is an indication of weaker hydrogen bonding to the C=O group (see Section 1.3). The amide-II band is slightly red-shifted with increasing charge state. This observation is a further indication of reduced hydrogen bonding [7] to N−H groups (mainly present in the backbone of the

---

[6]It is noteworthy that no signals of sodium ions and/or charge state 4+ of the protein at 1911 amu are observed.

[7]Generally involved in the hydrogen bonding network of the protein are surrounding solvent molecules (absent in the gas phase), carboxyl groups and amide groups in the backbone (responsible for secondary structural motifs) and interactions between the side chains of the amino acids (either with backbone
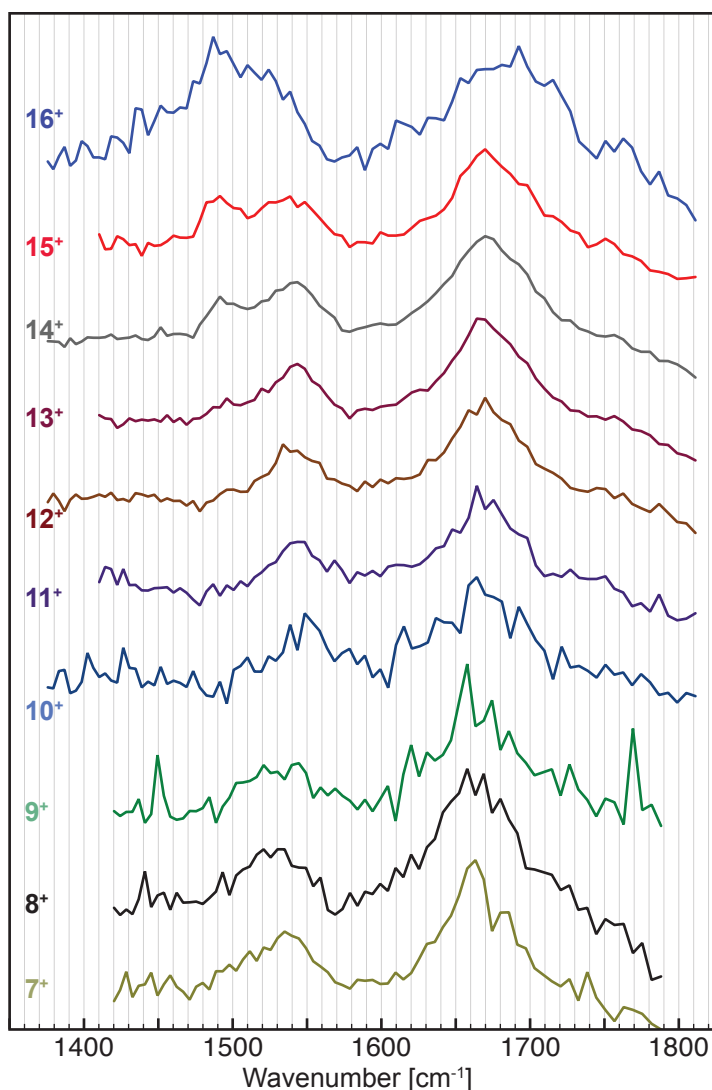
**Figure 3.8:** Experimental infrared spectrum of protonated and potassiated cytochrome c in different charge states ranging from 7+ (bottom) to 16+ (top). The charge state is given on the left side for each spectrum in the corresponding color. In the top part of Table 3.2 a list of the simultaneously recorded charge states of protonated cytochrome c is given.

protein) upon higher charge states. One exception are the observations in the IR spectra of the charge state 7+ to 9+. The spectra look very similar and, between the two, no shift of the amide-I and amide-II band is observed. The position of the amide-II band in those spectra is red-shifted around 10 cm$^{-1}$ with respect to the spectra for the higher charge states. This indicates an increased hydrogen bonding involving N−H groups for those species compared to the higher charged ions. Unfortunately, the signal-to-noise ratio for the charge states 8+ to 10+ is not very high due to a low parent ion signal in the corresponding mass spectra (not shown in this work). The features in the IR spectra for charge state 8+ at 1440 cm$^{-1}$ and those present in the spectrum of the charge state 9+ around 1450 cm$^{-1}$ and 1760 cm$^{-1}$ consist of only a few data points and are due to fluctuations in the parent ion signal. Also the region in the IR spectrum of charge state 10+ between 1370 cm$^{-1}$ and 1500 cm$^{-1}$ appears to be very noisy and is not considered suitable in the interpretation of the data.

In the IR spectra from charge state 12+ to 15+, a new band at around 1483 cm$^{-1}$ appears

groups or with each other).

(henceforth called 'mystery band'). This band has also been previously observed by Oomens *et al.* [146] and its origin is not clear. Different possibilities concerning the origin of this band will be given in Section 3.4.3.

In the infrared spectrum of charge state 16+ of the molecule the intensity of the 'mystery band' is higher than the amide-I and the amide-II band. Note the significantly broadened amide-I band in this spectrum (FWHM around 90 cm$^{-1}$) and the strong red-shift of the peak position (located at 1690 cm$^{-1}$) compared to the IR spectra of the lower charged ions. Due to the fact that the even charge states (10+, 12+, 14+, and 16+) are recorded simultaneously (see Table 3.2) errors resulting from the wavelength or power calibration of the IR photons can be excluded as the reason for this observation. However, the ion signal intensity of charge state 16+ of cytochrome c in the mass spectrum is the lowest compared to the other charge states.



**Figure 3.9:** Experimental infrared spectra of deuterated cytochrome c in different charge states ranging from 8+ (bottom) to 19+ (top). No IR spectrum of charge state 17+ is shown. The charge state is given on the left side of each spectrum in the corresponding color. In the bottom part of Table 3.2 a list of the simultaneously recorded charge states of deuterated cytochrome c is given.
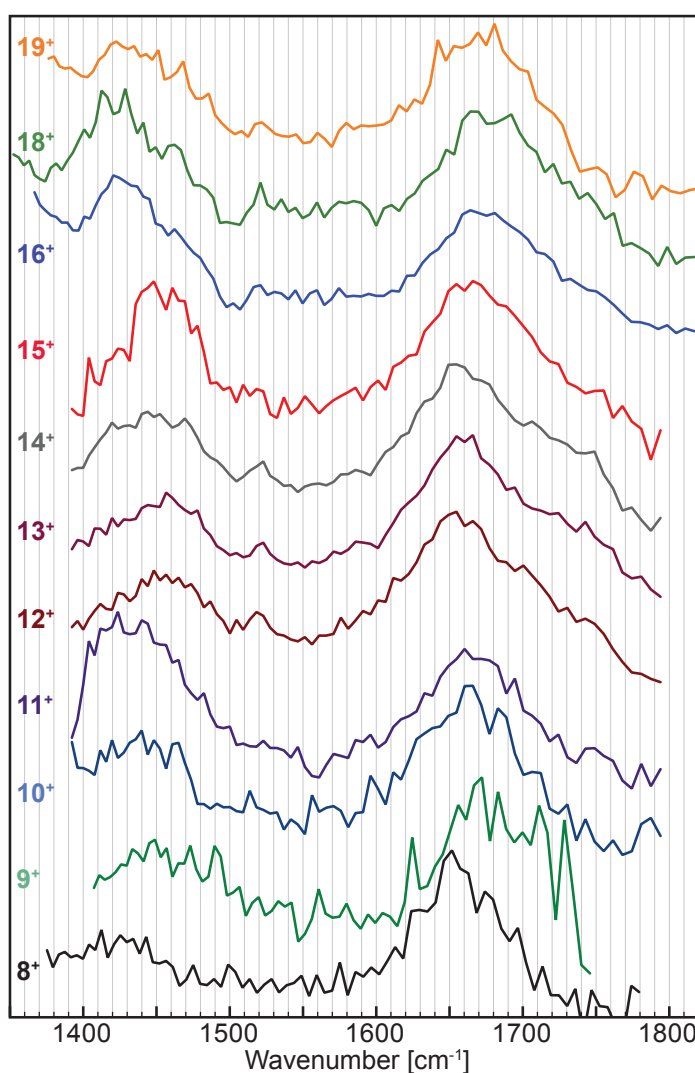
Figure 3.9 shows infrared spectra of deuterated cytochrome c ions in different charge states ranging from 8+ to 19+. No spectrum of the charge state 17+ is shown due to a low parent ion signal intensity in the experiment leading to a very poor signal-to-noise ratio. Two broad peaks are present in all spectra. The amide-I band is located at around

1655 cm$^{-1}$. For deuterated samples, the signature around 1430 cm$^{-1}$ to 1450 cm$^{-1}$ is commonly called amide-II'. Compared to non-deuterated samples, the amide-II' band in deuterated samples is characterized by an increased contribution of C$-$N stretching vibration and therefore an altered character of the band (see Section 1.3). The shift to lower wavenumbers upon exchange of the proton with a deuteron for a solely N$-$H bending vibration would be bigger. The group changes to N$-$D and therefore the position of this vibration would be located at the position of the N$-$H bending vibration divided by the factor $\sqrt{2}$ – resulting from the mass difference between a proton and a deuteron. The position of the amide-II' band varies in the spectra between different charge states. Two different 'families' of charge states can be observed with respect to the amide-II' band position: From charge state 8+ to 11+ and 16+ to 19+, the position is centered at around 1430 cm$^{-1}$. In the range of charge states 12+ to 15+, the position is blue-shifted and is centered at around 1450 cm$^{-1}$. Errors in the wavelength calibration of FELIX can be excluded since the amide-I band position does not follow this trend. The FWHM of the amide-I band is in the order of 80 to 90 cm$^{-1}$ and the FWHM of the amide-II' band, although not completely resolved in all spectra, is in the order of 70 cm$^{-1}$. With the increasing charge state of the protein the amide-I band shifts about 30 cm$^{-1}$ to the blue. One exception here is found for the band in the spectrum of charge state 9+. Here the band is shifted around 40 cm$^{-1}$ to higher energies compared to its position in the spectrum of charge state 8+. This spectrum anyway exhibits the worst signal-to-noise ratio of the deuterated spectra and there might be a problem with the reproducibility and stability of the parent ion signal. Due to the fact that this charge state is recorded together with charge states 11+ and 13+ a problem with the infrared photon energy calibration for this scan can be excluded. Interestingly, a smaller peak appears in the spectra of the charge states 12+ to 19+ at 1520 cm$^{-1}$ and does not shift with increasing charge state.

Consistent in both the recorded IR spectra of the protonated cytochrome c ions (see Figure 3.8) and the deuterated cytochrome c ions (see Figure 3.9) is the observed isotope shift of the amide-II band upon deuteration (see Chapter 1, Section 1.3). In Figure 3.10 the IR spectra of charge state 12+ and 14+ of deuterated and protonated cytochrome c are compared. These spectra are selected because the 'mystery band' is clearly present in the IR spectrum of protonated cytochrome c in charge state 14+ (intensity comparable with amide-II band) and begins to appear in the spectrum of charge state 12+ of the ion. Additionally, the infrared spectra for both charge states of protonated/deuterated cytochrome c are recorded at the same conditions, thus allowing a discussion of the difference between the protonated and deuterated spectra using those charge states. The amide-I band observed in the infrared spectra of the deuterated molecules is broader compared to the band found for the protonated ions. Also, the amide-I band peak position is shifted about 20 cm$^{-1}$ to lower energies upon deuteration, slightly larger than shifts that are observed in solution (depending on the secondary structure element involving this group, the average shift is between two and nine wavenumbers [42]). The character of the amide-I band is also not only given by a C=O stretching vibration: minor contributions from the out-of-phase C$-$N stretching vibration, the C$-$C$-$N deformation and the N$-$H in-plane bending vibration are involved (see Section 1.3). As expected, the amide-II' band is clearly shifted about 90 cm$^{-1}$ to the red side of the spectrum
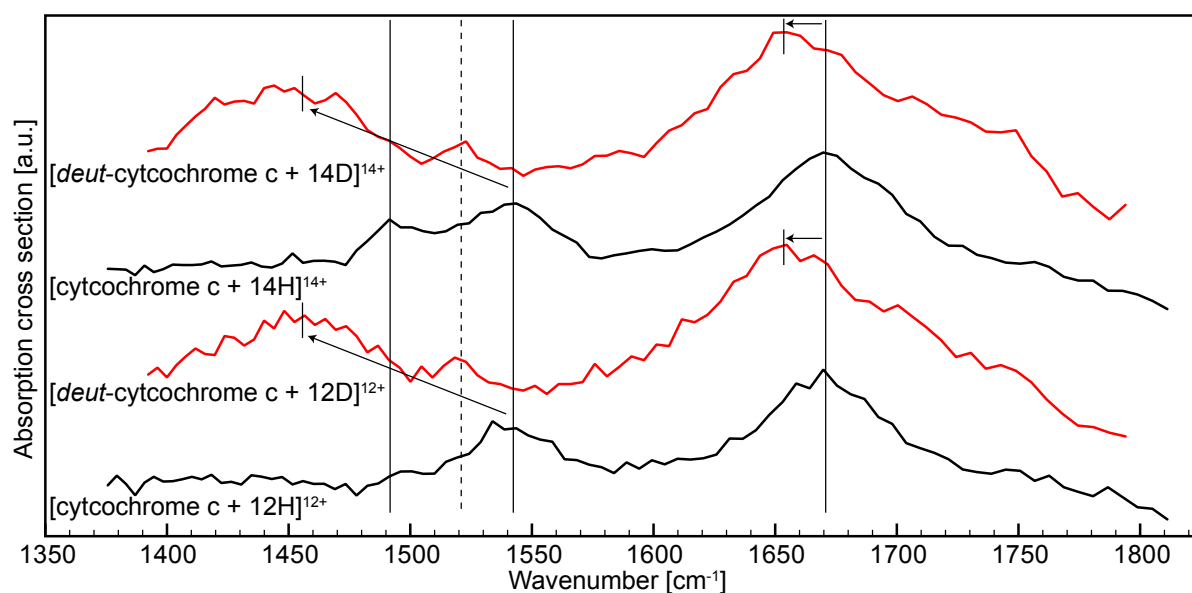
**Figure 3.10:** Comparison of the infrared spectra between protonated (black) and deuterated (red) cytochrome c ions in the charge states 12+ and 14+. The long horizontal lines mark the amide-I band, the amide-II band and the 'mystery-band' for the protonated species. The short lines indicate the position of the amide-I band and the amide-II' band for the deuterated species, with the arrows clarifying the shift upon deuteration of the amide bands. The dotted line indicates the small peak at 1520 cm$^{-1}$ as observed in the deuterated spectra.

upon deuteration (see Section 1.3). The 'mystery band' is not present in the infrared spectra of the deuterated species, instead at its former position, only the rising slope of the amide-II' band is observed. The small peak clearly appearing at 1520 cm$^{-1}$ in the deuterated spectra might be observed as a shoulder on the red side of the amide-II band in the protonated spectra.

The observed positions of the amide-I and amide-II band together with the corresponding FWHM for all recorded infrared spectra are shown in Table 3.3. The boundary values for the estimation of the FWHM are also presented.

### 3.3.2    Infrared spectrum of tendamistat in charge state 5+

Figure 3.11 shows an mid-infrared spectrum of tendamistat in charge state 5+ measured by monitoring the depletion of the parent ion mass peak as a function of the infrared wavelength. No appearance of sodium or of lower charge states of the protein are observed in the mass spectrum. In the range from 1580 cm$^{-1}$ to 1800 cm$^{-1}$, three IR spectra are recorded and averaged while from 1400 cm$^{-1}$ to 1580 cm$^{-1}$ only two infrared spectra are averaged. The amide-I band position is found at 1664 cm$^{-1}$ with a FWHM of 57 cm$^{-1}$. This peak exhibits a slightly asymmetric shape. This asymmetry in the peak structure is likely caused by two shoulders, which are present in all three independently

**Table 3.3:** Peak positions of the amide-I and amide-II band (amide-II' band respectively) in the IR spectra of cytochrome c (top: protonated, bottom: deuterated) for different charge states. The numbers in the column 'Start' and 'End' respectively describe the values used to determine the FWHM of the corresponding peaks. All units are given in [cm$^{-1}$].

|        | Amide-I | FWHM | Start | End  | Amide-II | FWHM | Start | End  |
|--------|---------|------|-------|------|----------|------|-------|------|
| Cyto c |         |      |       |      |          |      |       |      |
| 7H+    | 1663    | 49   | 1640  | 1689 | 1535     | 54   | 1504  | 1558 |
| 8H+    | 1663    | 62   | 1631  | 1693 | 1529     | 52   | 1501  | 1553 |
| 9H+    | 1657    | 49   | 1642  | 1691 | 1534     | 53   | 1500  | 1553 |
| 10H+   | 1664    | 68   | 1632  | 1704 | 1548     | 34   | 1532  | 1566 |
| 11H+   | 1670    | 60   | 1640  | 1700 | 1544     | 30   | 1529  | 1559 |
| 12H+   | 1670    | 55   | 1644  | 1699 | 1543     | 35   | 1526  | 1561 |
| 13H+   | 1670    | 55   | 1644  | 1699 | 1544     | 34   | 1526  | 1561 |
| 14H+   | 1671    | 62   | 1640  | 1702 | 1542     | 40   | 1518  | 1558 |
| 15H+   | 1671    | 56   | 1647  | 1703 | 1543     | 47   | 1511  | 1558 |
| 16H+   | 1688    | 100  |       |      | (1524)   | N/A  |       |      |

|          | Amide-I | FWHM | Start | End  | Amide-II' | FWHM | Start | End  |
|----------|---------|------|-------|------|-----------|------|-------|------|
| *d*-Cyto c |         |      |       |      |           |      |       |      |
| 8D+      | 1656    | 71   | 1621  | 1692 | 1419      | 79   | 1380  | 1459 |
| 9D+      | 1681    | 84   | 1636  | 1720 | 1450      | 77   | 1416  | 1493 |
| 10D+     | 1666    | 95   | 1618  | 1713 | 1437      | 65   | 1408  | 1473 |
| 11D+     | 1668    | 79   | 1624  | 1705 | 1433      | 79   | 1399  | 1478 |
| 12D+     | 1657    | 110  | 1608  | 1718 | 1455      | 66   | 1420  | 1486 |
| 13D+     | 1661    | 93   | 1617  | 1710 | 1456      | 71   | 1414  | 1485 |
| 14D+     | 1656    | 93   | 1622  | 1715 | 1448      | 78   | 1407  | 1481 |
| 15D+     | 1665    | 79   | 1629  | 1710 | 1456      | 53   | 1429  | 1482 |
| 16D+     | 1673    | 85   | 1634  | 1719 | 1429      | 63   | 1396  | 1459 |
| 18D+     | 1676    | 84   | 1639  | 1723 | 1423      | 94   | 1377  | 1471 |
| 19D+     | 1676    | 82   | 1630  | 1712 | 1425      | 70   | 1402  | 1472 |

recorded IR spectra[8]. The shoulder on the red side of the peak can be identified at 1619 cm$^{-1}$ and the one present at higher energies is located at 1692 cm$^{-1}$. At the position of the amide-II band in the infrared spectrum an (unresolved) double peak structure is observed. This feature is present in both independently taken scans and is hence not to be ignored. The maximum of the peak is found to be at 1519 cm$^{-1}$ and a plateau is located at around 1530 cm$^{-1}$. The FWHM of the measured peak is 53 cm$^{-1}$.

In Figure 3.12 infrared spectra of cytochrome c and tendamistat are shown together to

---

[8]which are the basis of the spectra shown in Figure 3.11
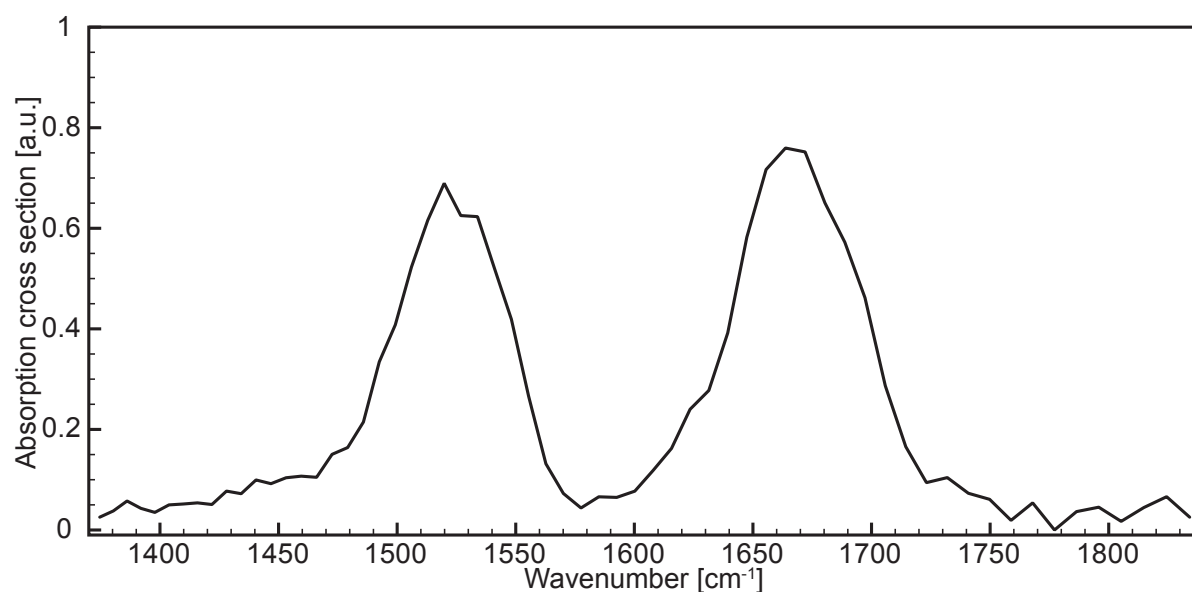
**Figure 3.11:** Infrared spectrum of tendamistat in charge state 5+.

emphasize differences. Because of its good signal-to-noise ratio and the presence of the 'mystery band', the infrared spectrum of charge state 14+ is selected for cytochrome c, both in protonated (blue) and deuterated (red) form. This allows for direct comparison of the observed spectral features and the region where the 'mystery band' is observed is of interest and might help to clarify the nature of the band.

The amide-I band of protonated cytochrome c in charge state 14+ is slightly blue-shifted with respect to the position of the band in the spectrum of tendamistat. On the other hand, the amide-I band position is red-shifted in the spectrum of deuterated cytochrome c compared to tendamistat. There is a small peak (mentioned on page 53) at $1520 \, \text{cm}^{-1}$ visible in the IR spectrum of deuterated cytochrome c at the position of the amide-II band of tendamistat. At this wavenumber, a plateau between the amide-II band and the 'mystery band' in the protonated cytochrome c spectrum is present. The width of the amide-I band in the infrared spectrum of tendamistat is smaller than in the spectrum of cytochrome c. It must be denoted that the amide-I band position observed in the spectra of protonated cytochrome c for the low charge states (7+ to 10+, see Figure 3.8) is the same as for tendamistat.

## 3.4 Discussion

Unfortunately, for the molecules discussed here, it is difficult to perform quantum chemical calculations – in order to find meaningful theoretical structures of the proteins or to obtain infrared frequencies and intensities – due to the large size and high number of internal quantum states of the investigated systems.

In Table 3.3, the observed positions of the amide-I and amide-II bands in the infrared spectra for protonated and deuterated cytochrome c ions in different charge
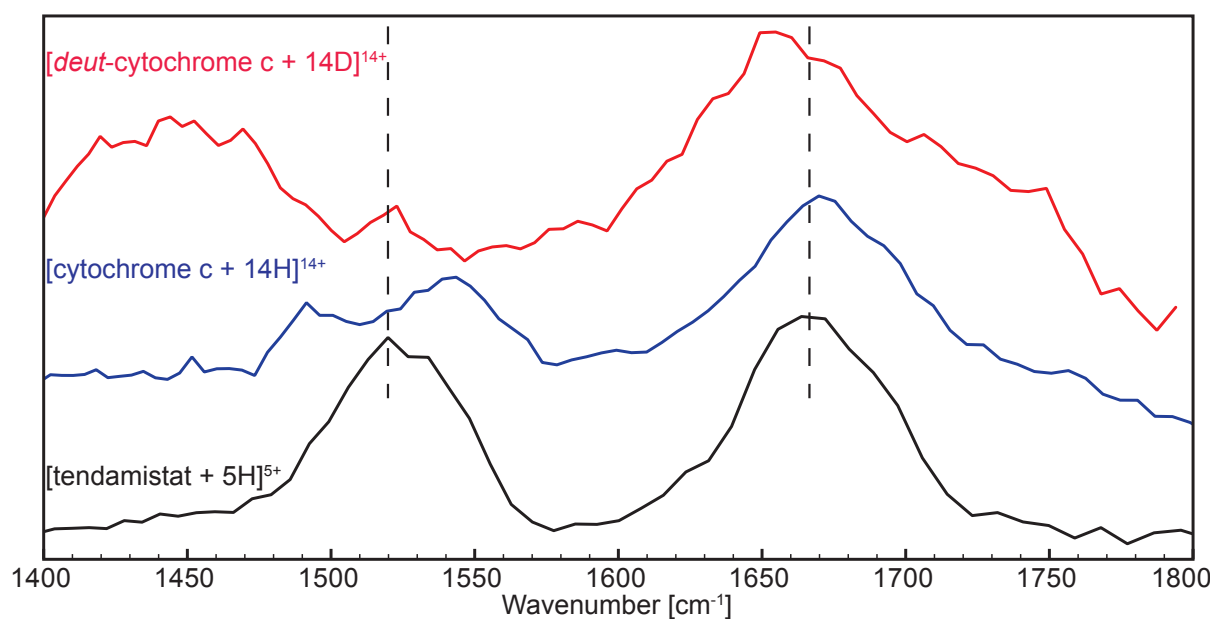
**Figure 3.12:** Comparison of the infrared spectra of tendamistat in charge state 5+ (black) with the protonated (blue) and deuterated (red) cytochrome c in charge state 14+. The dotted line indicates the position of the amide-I and amide-II band of tendamistat.

states are shown. In the condensed phase, the position of the amide-I and amide-II band is sensitive to the three-dimensional structure of a protein. Comparing the position of those bands in the infrared spectra for different molecules and charge states allows for some qualitative interpretation of possible structural changes involving the charges on the protein. The blue-shift of the amide-I band position in the IR spectra of protonated and deuterated cytochrome c for high charge states is an indicator for reduced strength in the hydrogen bonding pattern of the protein backbone. In the condensed phase the amide-I band positioned between 1648 cm$^{-1}$ and 1657 cm$^{-1}$ is an indicator for an $\alpha$-helical secondary structure of the backbone in proteins [42]. Upon the transitions from solution phase to the gas phase, calculations on small model systems predict a shift of this band between 10 cm$^{-1}$ and 80 cm$^{-1}$ towards higher energies [147]. However, if the molecule is big enough, self-solvation can occur and the shift of the band might be smaller than the one estimated for small molecules.

In this study the peak positions are not determined by fitting Gaussian functions to the spectrum. There are too many degrees of freedom – such as the number of Gaussians, the width, the mean position(s) of the Gaussian functions – to assign the bands accurately by fitting. Therefore, the absolute precision of the band assignment is not especially high but nevertheless it is possible to compare the infrared spectra of the different charge states and assign relative shifts. Especially at the amide-I band position a blue-shift of the band with increasing charge state is clearly observed. In the accurate assignment of the amide-II band, a problem occurs: the 'mystery band' appears in that region in the infrared spectra of the protein in higher charge states and a clear distinction between those bands is difficult.
The amide-I band position in the infrared spectra, however is blue-shifted with in-

creasing charge state ranging from 7+ to 16+. This is an indicator for a decrease in the backbone hydrogen bonding network strength with increasing charge state. The amide-I peak position in the lower charge states (7+ to 10+) of cytochrome c indicate an increased hydrogen bonding strength and the band position is similar for that group. The increased hydrogen bonding strength might indicate a more globular conformation of cytochrome c for the low charge states. The amide-I band position in the IR spectra of the charge states 11+ to 15+ of the molecule is located at 1670 cm$^{-1}$, blue-shifted by around 20 cm$^{-1}$ with respect to its expected position for an $\alpha$-helical motif in the condensed phase. Taking a blue-shift of the band upon the transition from solution to the gas phase into account this observed band for those charge states might be an indicator for an $\alpha$-helical motif in the backbone of the protein in the gas phase. The further blue-shifted band position in the spectrum of charge state 16+ might be correlated with a decrease in the backbone hydrogen bonding network and could be a result of an elongation of the denatured protein due to Coulombic charge repulsion.

The amide-I band in the infrared spectra of the different charge states of deuterated cytochrome c behaves similar to that observed for the protonated ions. Compared to the position in the protonated molecule, the amide-I mode of the corresponding charge states is red-shifted upon deuteration, which is in line with observations of proteins in the condensed phase (see page 53). As the charge state increases the position of the band is shifted to higher energies. In the IR spectrum of the lowest charge state 8+, the amide-I band position is located at 1656 cm$^{-1}$, the largest red-shift observed in all deuterated spectra, thus indicating a stronger hydrogen bonding network in the backbone. On the other hand the band position in charge state 9+ is the most blue-shifted of all spectra but the band is also relatively broad (FWHM 84 cm$^{-1}$) and therefore not very significant. In the infrared spectra of charge state 10+ to 15+ of the molecule the amide-I band position is observed at around 1665 cm$^{-1}$, which is consistent with the observed results for the protonated molecules. That position of the amide-I band might be indicative for an $\alpha$-helical structure of the peptidic backbone when the protein is deuterated. However, there are two exceptions: The band position in the spectra of charge state 12+ and 14+ is slightly red-shifted with respect to the position in the other spectra of neighboring charge states but the FWHM value for those charge states is also exceptionally large. This makes it difficult to assign an absolute value for the peak position. In the infrared spectra of the proteins in charge state 16+ to 19+, the amide-I band is located at around 1775 cm$^{-1}$ and indicates a weaker hydrogen bonding network strength with respect to the lower charge states, similar to the protonated counterparts. An elongation of the backbone due to charge repulsion is a plausible reason for that. It needs to be pointed out that the amide-I bands in the spectra of deuterated cytochrome c appear about 50% broader compared to the width observed for the protonated counterparts.

In solution-phase IR spectroscopy of proteins, the amide-II mode is less employed in assigning secondary structural motifs. On the one hand, an overlap with a bending mode of $D_2O$ is present at this wavelength, making a determination of this protein vibrational mode difficult. On the other hand, it seems that this mode is less sensitive to conformational changes in the secondary structure [148, 149]. Obviously, there are

no problems with bands originating from solvent molecules in the gas phase. Upon decrease in hydrogen bonding of the backbone involving the amide N−H and N−D groups respectively, a red-shift is expected to occur for the amide-II band position that is in the opposite direction compared to that of the amide-I band (see Section 1.3).

In the IR spectra of the lower charge states (7+ to 9+) of protonated cytochrome c the amide-II band position is found at about 1533 cm$^{-1}$. Interestingly, the assigned position for the amide-II band in the IR spectrum of cytochrome c in charge state 10+ is the most blue-shifted. This behavior is in contrast to the observations of the amide-I peak position in this charge state. Those observations suggest a decrease in the backbone hydrogen network bonding strength, but the amide-II band shift suggests an increase of the strength. With increasing charge state, the assigned amide-II band position in the IR spectra shifts slightly to the red, indicating a decreasing strength in the hydrogen bonding strength. However, as stated, a new band appears and assignment of the amide-II band position is very difficult. The clear assignment of the position of the N−H bending vibration is also difficult because vibrations originating from amino acid side chain groups are found in this region [150]. This might also lead to an incorrect assignment of the absolute band position.

As expected upon deuteration of the molecule the amide-II′ band position in the infrared spectra is shifted to lower energies due to the change in character of the band as has been described previously.

In the low energy region the infrared spectra of the deuterated molecules appear noisier, compared to the protonated ones. This is due to a lower parent ion signal intensity for those species.

For the low charge states (8+ to 11+), the amide-II′ band position is located at around 1430 cm$^{-1}$ and at around 1455 cm$^{-1}$ in the infrared spectra of charge state 12+ to 15+ of the protein. With increasing charge state of the protein, the amide-II′ band position might be blue-shifted thus indicating an increased strength of the backbone hydrogen (deuteron) bonding network. A stronger red-shift of this band is observed in the spectra of the highest charged molecules (16+ to 19+), indicating a decrease in the hydrogen (deuteron) bonding network strength. The latter observation is in agreement with the shift of the amide-I band position as found in the IR spectra for those charge states.

In the infrared spectra of deuterated cytochrome c a band appears at 1520 cm$^{-1}$. This might be due to amino acide side chain vibrations. In solution a relatively strong vibrational band of tyrosine is located at around 1515 cm$^{-1}$. This band is found in a number of proteins and sometimes appears as a slight shoulder on the low-frequency side of the amide-II band [151]. The band arises from a stretching vibration ($\nu$(CC)ring) and a bending vibration ($\delta$(CH)) of side chain groups in tyrosine (Tyr–OH). Upon deuteration this band is shifted from [1516 - 1518][9] cm$^{-1}$ to [1513 - 1517] cm$^{-1}$. The oscillator strength of this modes is quite strong. However, it needs to be pointed out that only about 4% (exact: 4 residues) of the amino acids in cytochrome c are tyrosine. Two other bands of tyrosine resulting from stretching vibration of C−C ($\nu$(CC)ring) are found in the condensed-phase IR spectra at [1614 - 1621] cm$^{-1}$ (deuterated: [1612 - 1618] cm$^{-1}$) and [1594 - 1602] cm$^{-1}$ (deuterated: [1590 - 1591]). In the higher frequency band

---

[9]the square brackets indicate a wavenumber region, in which the corresponding mode is found in the condensed phase according to [5]

the bending vibration of C−H ($\delta$(CH)) is weakly coupled. The intensity of the latter two bands however is low compared to the vibration found at around 1517 cm$^{-1}$ [5]. As one can see from the numbers above, these bands do not shift dramatically when protons in exchangeable chemical groups are replaced by deuterons (the OH hydrogen of tyrosine is not expected to exchange). In the infrared spectra of protonated cytochrome c the amide-II band is relatively broad and, especially for higher charge states where the 1483 cm$^{-1}$ band appears, a band located at 1520 cm$^{-1}$ cannot be independently resolved. Thus, taking a small blue-shift of the vibrational band due to the solvent-free environment into account an unresolved band could be located at around 1520 cm$^{-1}$ in the infrared spectra. IR-MPD studies of protonated bare tyrosine also suggest a band at around 1520 cm$^{-1}$ in the IR spectrum [152]. Upon deuteration, interfering vibrational modes due to the amide-II vibration with the 1483 cm$^{-1}$ band, which is not present in the spectra of deuterated cytochrome c, are shifted. If the shift of the Tyr–OH band upon deuteration in the spectrum is small, the origin of the peak at 1520 cm$^{-1}$ might be a result of this amino acide side chain IR absorption. The other two previously described bands corresponding to Tyr–OH might be also present in all IR spectra. However, as they are weak they are not visible. It has been found that in IR-MPD studies of protonated bare tyrosine, those higher frequency bands are absent as well. In the infrared spectra of tendamistat the vibrational modes of amino acid side chains might be visible. The results are described in Section 3.4.4.

The deuteration experiment gives an insight into the origin of the 'mystery band'. Due to the absence of this band in the deuterated spectra, an involvement of a proton(s) in the oscillating group is expected. The origin of the band is not connected to the non-covalent binding of potassium ions to the protein. Further discussions concerning the 1483 cm$^{-1}$ band are presented in Section 3.4.3. In Section 3.4.4 the differences between the IR spectra of tendamistat and cytochrome c are discussed and interpretations are given.

### 3.4.1 Gas-phase IR spectra of cytochrome c compared to solution IR spectra

An infrared spectrum of cytochrome c in aqueous solution to assign positions of the amide-I band to secondary structure elements was recorded in 1990 by Caughey *et al.* [153]. Those results have been compared to structural assignments obtained by X-ray diffraction methods and it is found that the amount of secondary structure elements obtained by both methods are in very good agreement. Cytochrome c contains 42% $\alpha$-helical, 21% $\beta$-sheet, 25% turn, and 12% random structure elements in solution. The band position of the C=O vibration is found for an $\alpha$-helical motif at 1656 ± 2 cm$^{-1}$. In the infrared spectrum of cytochrome c in solution, the FWHM of the amide-I band is around 50 cm$^{-1}$. In a more recent infrared study on cytochrome c in aqueous solution by Rush *et al.* [154], similar results concerning the structure are observed. In that work, IR spectra of cytochrome c with and without guanidine hydrochloride[10] (GuHCl) are taken. Due to infrared absorbance of a C=N group at 1671 cm$^{-1}$ of GuHCl (stretching

---

[10]a denaturant

vibration), the deuterated species (DGuHCl) is used to record an IR spectrum in the amide-I band region. As a consequence, the vibrational mode is shifted to a lower energy at 1598 cm$^{-1}$ and is not overlapping with normal modes of the amide-I band. Cytochrome c is also deuterated in the presence of the deuterated denaturant. In contrast, infrared spectra of protonated and deuterated cytochrome c in solution under native conditions are presented in the work by Rush. With denaturing conditions (6.6 M DGuHCl), the peak becomes broader and is shifted to a lower energy indicating a higher $\beta$-sheet contribution compared to the native form. The width of the amide-I band found in the gas-phase infrared spectra of protonated cytochrome c is in the order of 55 cm$^{-1}$ and therefore more related to the width found in IR spectra recorded under native solution conditions. This might be an indicator for a more similar structure of the molecule in the gas phase to its native solution form, rather than to a completely denatured form. Also, it might indicate that the structural diversity of cytochrome c in the gas phase is not much greater than in solution, as suggested previously in the work published by Oomens *et al*.

One observation however is not yet clear: The width of the amide-I peak of deuterated cytochrome c in the gas-phase is broader compared to its protonated counterpart. This is not the case in solution.

### 3.4.2 IR spectra of cytochrome c compared to results obtained by ion mobility studies

In ion mobility studies (IM), the drift times for an ion through a buffer gas filled drift tube in the presence of an electrical field is measured. This time can be assigned to the collision cross section and thereby to a global structure of the molecule [155, 156]. According to their mobilities, different conformers such as compact or extended structures of proteins can be separated. Compact molecules have a shorter drift time in the tube compared to extended ones due to their smaller cross section and thus a smaller net decelerating force by molecular collisions. Upon inelastic collisions it can happen in the drift tube that a part of the kinetic energy of the injected ion is converted into internal energy, which then leads to a structural change. Therefore, experiments at low injection energies are often used. Nevertheless, conformers observed at high injection energies can overcome possible barriers in the potential energy landscape as a result of unfolding and refolding in the drift tube. They are believed to reflect the most stable gas-phase structures [157]. Ion mobility experiments allow for insights into the global average structure of a protein in contrast to infrared spectroscopy, where informations on local structures are obtained. A detailed explanation on IM on biomolecular systems is given for example in review articles by Clemmer *et al.* [158, 159].

In IM experiments on cytochrome c, it is observed that for charge states up to 7+ with low injection energies, collapsed[11] native solution phase conformations are present. On the other hand, if the low charge states 6+ and 7+ are generated via charge stripping from higher charged ions, the conformation is unfolded. For the charge states 3+ to

---

[11]high density of van der Waals contacts and hydrogen bonds

5+, the conformation is independent from the method of generating the gas-phase ions and the observed structure is very similar to that in the native solution. For the charge states 8+ to 10+, the compact geometries are metastable and can gradually unfold upon collisional heating through a series of well-defined intermediates. The conformations of highly charged ions in the gas phase is extended and the degree of extension increases with the charge state of the molecule [157, 160]. IM studies on cytochrome c in charge state 9+ after different storage times in a Paul trap show a rapid transition from a compact solution structure to unfolding into several more open structures in only 30 - 60 ms. However, a refolding of these unfolded states into folded structures appears to happen on a very long timescale in the order of 250 ms to 10 s [161]. Furthermore, FT ICR studies emphasize that distinct conformations do not interconvert even after storage times in the order of minutes and that at long time scales gas-phase ion structures have not reached the equilibrium [162, 163]. This is interesting because in the experiment, the transition of the molecules from solution to the ICR cell and the subsequent recording of an infrared spectrum is on a time scale of a few seconds. The macro-pulses of the infrared irradiation is on a $\mu$s long time scale and if conversion of different structures occurs at those times, a superposition of different conformers is measured leading to the observed infrared spectrum (remember, the ions are usually irradiated for four seconds in the cell corresponding to about 40 infrared macro-pulses).

Comparing the results obtained by ion mobility studies with those from infrared spectroscopy (see Figure 3.8) can potentially give some additional insights in the protein structure. The band position of the C=O vibration is found for an $\alpha$-helical motif at $1656 \pm 2$ cm$^{-1}$. In the infrared spectrum of cytochrome c in solution, the charge state 8+ is broader compared to higher charge states such as 12+ (except 16+). This might be an indicator for the presence of multiple stable and long lasting conformations for the intermediate charge states because the measured infrared spectrum is a superposition of the spectra of all present conformations in the ICR cell. No IR spectra in the low charge state range from 3+ to 5+ are recorded to probe the presence of globular conformations by IR spectroscopy as it has been done with IM experiments. It is expected that narrow bands (for example amide-I) together with the band position would be an indicator for globular structures of the molecule. An abrupt change of the conformation between charge state 11+ and 12 + is not observed in IM experiments.

In conclusion, it is difficult to compare interpretations obtained by IM with the IR spectra measured in this work. One possible explanation is the different time scales between the two experiments. Due to the harsh, possibly denaturing conditions of the protein solution, it is not unlikely that unfolded cytochrome c molecules are investigated spectroscopically. Another difference is that IR spectroscopy probes structural features more locally while IM experiments provide insight into the global structure of the molecule. Nevertheless, the combination of both methods is a powerful tool to investigate gas-phase structures of proteins. This will be further discussed in Chapter 7.

### 3.4.3   A new perspective on previously published infrared spectra of gas-phase cytochrome c

In 2005, mid-IR spectroscopy was performed on bovine cytochrome c in different charge states in the gas phase at the FELIX facility in Nieuwegein [146]. In this work, a band at 1483 cm$^{-1}$ appeared from charge state 13+ upwards in the infrared spectra. The position of the band in the IR spectra is in excellent agreement with the so-called 'mystery band' observed in this data presented here. As mentioned earlier, the origin of this band is not yet clear. Several possible explanations for the appearance in the spectra are given in the previous publication: 1) COH in-plane bending mode of COOH side chain groups, 2) umbrella mode of the $NH_3^+$ moiety, 3) changes in secondary structure of the protein with the charge state.

Before discussing those conclusions and comparing them with the results obtained here, it is important to point out a specific difference between both investigated systems. As mentioned in Section 3.1, there are distinctions between the protein cytochrome c from different species. Compared to the primary structure of equine heart cytochrome c, bovine cytochrome c has three point mutations[12] (both amino acid sequence informations are taken from the UniProt database [164]; equine heart: P00004; bovine heart: P62894). As a consequence, equine cytochrome c possesses 24 basic and 12 acidic amino acids meanwhile the bovine species has 23 basic and 12 acidic amino acids.

This is of consequence for the first interpretation regarding the appearance of the 'mystery band' drawn in the work done by Oomens *et. al.*: There, it has been observed that the appearance of the band could coincide with a COH in-plane bending (ipb) mode of the COOH side chain groups in the acidic residues Glu and Asp. It is based on the protein being in a zwitterionic protonation state (bovine cytochrome c: all 23 basic residues are protonated and all 12 acidic residues are de-protonated), leading to an 11+ charge state. Together with the potassium ion and the singly positive charged heme group [165], the charge state 14+ would be the first where protonation of the $COO^-$ groups of the acidic residues is possible to form COOH. The ipb mode of COH of $CH_3COOH$ in solution is – if hydrogen bonded – observed at 1058 cm$^{-1}$ and 1450 cm$^{-1}$ [150]. Therefore, a value of 1483 cm$^{-1}$ for this mode in the gas phase seems reasonable. The higher the charge state, the more COOH groups are formed, explaining an enhancement of the intensity with increasing protonation of the molecule. Upon deuteration, the ipb mode of COD is shifted to lower frequencies compared to the ipb mode of the COH group. The additional basic residue in equine cytochrome c would increase the charge state to 15+, where the first protonated COOH group (from an acidic residue) would exist. Since the band starts to appear in the infrared spectrum of charge state 12+ presented here and the intensity of this band increases for higher charge states, the simple mechanism described earlier cannot be applicable.

Another argument against the assignment of the 'mystery band' to an ipb mode of a COH group is the very low number of possible oscillators absorbing IR light. In protonation state 15+ of equine cytochrome c, only one COH group would be present.

---

[12]amino acid in equine → amino acid in bovine: $Thr_{47}$ → $Ser_{47}$, $Lys_{60}$ → $Gly_{60}$, $Thr_{89}$ → $Gly_{89}$, the number corresponds to the position of the amino acid in the sequence

But the intensity of the 'mystery band' in the infrared spectrum of this charge state is comparable to the intensity of the amide-II band, where around 100 oscillators absorb the infrared light. It is hard to imagine how only one oscillator can lead to such a strong band. Nevertheless, due to the absence of the band at 1483 $cm^{-1}$ in the IR spectra for deuterated cytochrome c an involvement of a hydrogen atom is plausible.

The second possibility for assigning the 1483 $cm^{-1}$ band is based on a quite intense absorption of IR light close to the observed energy of the umbrella mode of a $NH_3^+$ moiety in the gas phase [166], which can arise in cytochrome c by protonation of the lysine side chains. In contrast to bovine cytochrome c – where 18 lysine residues are present – 19 lysine residues are part of the primary structure of equine cytochrome c. The other five basic residues (two Arg, three His) in the amino acid sequence of cytochrome c (both equine and bovine species) rather form $NH_2$ (Arg) or NH (His) groups upon protonation. One argument supporting the role of $NH_3^+$ umbrella vibrations in this region is the fact that one more possible oscillator is present in the equine cytochrome c with respect to the bovine one. This could explain the earlier appearance and the higher observed intensity for the 'mystery band' in the IR spectrum of equine compared to the bovine cytochrome c molecule. All lysine residues should already be protonated in charge state 12+ in both cases implicating a protonation of lysine prior to histidine under the solution conditions used in this study due to the lower isoelectric point of the latter molecule. Therefore, it is not expected that the intensity of the band increases, when all possible lysine residues are already protonated, which is definitely the case at charge state 14+ (charge state 13+ respectively for bovine cytochrome c).
The interpretation of IR spectra of deuterated cytochrome c does not lead to a definite conclusion as to whether an umbrella mode of $NH_3^+$ is the cause for the appearance of the 1483 $cm^{-1}$ band or not. Upon deuteration, the mode would shift to lower energies. In the case of deuterated ammonia ($NH_3 \rightarrow ND_3$), a shift of about 500 $cm^{-1}$ is expected [167]. As well for lysine in solution, there is a shift upon deuteration of about 400 $cm^{-1}$ observed [150]. Unfortunately, this range is not covered by the here presented measurements.

A third possible explanation for the appearance of the 1483 $cm^{-1}$ band is that changes in the secondary structure occur with increasing charge state of the protein. Along with a blue-shift of the amide-I band with increasing charge state of bovine cytochrome c, the intensity of the 'mystery band' increases. As described above, a blue-shift of the amide-I band indicates a possible reduction in the intramolecular hydrogen-bonding strength in the backbone of the protein. Upon higher charging of the protein, Coulomb repulsions between the charged residues in the ion increase. This can cause structural rearrangements (such as unfolding or an elongation of the protein leading to maximizing the distances between the same charged group). Lysine residues can serve as possible chromophores that are influenced by the charge state. The increase of the 1483 $cm^{-1}$ band intensity is correlated with the charge state and therefore, might be related to the same phenomenon. In the IR spectra of equine cytochrome c, the observation is similar. There is a slight blue-shift of the amide-I band with increasing charge state of about 10 $cm^{-1}$ in the spectra of cytochrome c between charge state 7+ and 15+. The amide-I band position in the spectrum of charge state 16+ of the molecule is strongly

blue-shifted (about 25 cm$^{-1}$ with respect to the position in charge state 7+). Along with a large increase in intensity of the 'mystery band' in the spectrum of the latter charge state, a correlation between structural changes and the increase in intensity of this band seems plausible. Taking into account the low signal-to-noise ratio and fluctuations of the baseline in the infrared spectrum of protonated cytochrome c in charge state 11+, there might also be a band present at 1483 cm$^{-1}$. This in combination with the observation of a jump of 6 cm$^{-1}$ in the amide-I band position in the spectra between charge state 10+ and 11+ might be an indicator for a structural transition of the protein between those charge states.

The most important point mutation with respect to the 'native charge state' between equine and bovine cytochrome c is the exchange of the Gly (bovine) with Lys (equine) residue at the position 60 of the protein sequence. In the solution and crystal structure obtained via X-ray diffraction (see Figure 3.1) this lysine residue is located at the end of an $\alpha$-helical motif. Geometrically the closest amino acids are the residues at the position 61 to 64 and 59 in the primary sequence. Residue 61 and 62 are both glutamic acids. In the transition from solution to the gas phase, the triplet Lys-Glu-Glu might form somehow a zwitterionic motif. This motif might be disturbed upon further protonation of the molecule. Obviously, this is not possible in bovine cytochrome c due to the lack of that basic residue. As a consequence, there might be a local slightly different conformation present of bovine and equine cytochrome c in the same charge state, for example 12+. Thus, if the 'mystery band' is sensitive to the conformation, a difference in intensity and the appearing of this band at different charge states for the different cytochrome c molecules might occur.
There are also motifs present, where two lysine residues are adjacent to each other in the primary sequence (residue 7 and 8, 72 and 73, 99 and 100), three lysine residues (86, 87, and 88) and the motif Lys-His-Lys (25, 26, and 27). This might influence the protonation[13] of the basic residues with increasing charge state due to Coulombic charge repulsion of adjacent amino acids because not necessarily all lysine residues are protonated before the other basic residues[14]. Therefore, studies on mutant forms of cytochrome c would be interesting and the influence of adjacent charged amino acids could be investigated when, for example, those basic adjacent amino acids are exchanged by neutral residues.

Another possible explanation for the origin of the 1483 cm$^{-1}$ band might be given by infrared absorption of the side chains of the amino acids. One amino acid that has a side chain with an infrared active transition in this region is proline – the only amino acid in which the side chain is locked to the backbone and hence possesses a very rigid structure. A strong band originating from a bending vibration of CH$_2$ groups in the pyrroline ring ($\delta$(CH$_2$)) is found between 1425 cm$^{-1}$ and 1475 cm$^{-1}$ in solution [150]. The intensity of this band is increased when the oscillating group is next to a C=O group. The stretching vibration of CN ($\nu$(CN)) is found in the range of 1400 cm$^{-1}$ to 1465 cm$^{-1}$ in solution and is highly sensitive to the backbone conformation. In aqueous solutions of poly-L-proline, a strong infrared absorption at around 1450

---

[13]more specifically, the order of the protonation
[14]as indicated by the isoelectric point

**Table 3.4:** Distances of amino acids to proline in cytochrome c (structure: #1HRC in pdb). The distance is measured between the $C_\alpha$ atoms of the amino acids. The subscript number at the amino acid corresponds to its position in the primary sequence of the molecule.

| Amino acid | Distance [Å] | Amino acid | Distance [Å] |
|---|---|---|---|
| $Pro_{30}$ | | $Pro_{44}$ | |
| $His_{26}$ | 4.41 | $Asn_{31}$ | 9.29 |
| $Pro_{71}$ | | $Pro_{76}$ | |
| $Asn_{70}$ | 3.81 | $Lys_{73}$ | 5.77 |
| $Lys_{72}$ | 3.38 | $Tyr_{74}$ | 6.15 |
| $Lys_{73}$ | 5.42 | | |

$cm^{-1}$ is observed [168]. On the other hand, in IR-MPD spectra of singly protonated bare proline ions in the gas phase, no IR active band is observed between 1400 $cm^{-1}$ and 1720 $cm^{-1}$ [169]. Fourier transform infrared spectra of gas-phase proline at 570 K also shows no vibrational modes of the molecule in this region [170]. No hydrogen bonded conformers are present at this temperature. In the IR spectra of $(Pro)_2H^+$, a falling slope down to 1500 $cm^{-1}$ from a peak at 1400 $cm^{-1}$ is observed [171]. This indicates possibly the presence of vibrational bands.

Nevertheless, a highly intriguing observation underlining the possible involvement of proline in the appearance of the 1483 $cm^{-1}$ band is found in other gas-phase studies: In IR-MPD spectra of the protonated [proline – methylamine ($CH_3NH_2$)] complex a peak at 1484 $cm^{-1}$ is observed [172]. The position of the band is consistent with calculated $CH_2$ scissor modes of proline as well as protonated methylamine. Two hydrogen bonds are present in the proline – methylamine complex. In proteins, ammonium ions are present in charged side chains of amino acids and it might be possible that those groups are also involved in the formation of hydrogen bonds to proline. Clearly, the formation of charged side chains depends on the charge state of the molecule.

Other studies suggest that the presence of $Na^+$ stabilizes zwitterionic proline [166] and in calculated infrared spectra a band appears in the 1480 $cm^{-1}$ region. As a consequence, the presence of K+ and ammonium groups close to proline could lead to the appearance of a band in this region, which is absent in the spectra of bare protonated proline. In summary, charged groups or metal ions interacting with proline enhance the infrared intensity of infrared active modes in proline. However, it is not clear if the intensity would be sufficient to account for the strength of the 1483 $cm^{-1}$ band observed in the infrared spectrum of cytochrome c.

In Table 3.4, amino acids close to proline in the three-dimensional structure of cytochrome c obtained via X-ray diffraction (see Figure 3.1) are shown. The distance between proline and the other amino acid is given in Å and is measured between the two $C_\alpha$ atoms of the amino acids. It might be possible that one of those residues coordinates with a proline resulting in the appearance of the 1483 $cm^{-1}$ band in the IR

spectra of higher charged cytochrome c ions.

The intensity of the CN stretching vibration of proline is found to be sensitive to the backbone structure. The higher the strain of this group, the higher the intensity of the vibrational mode. Proline is a very rigid molecule. Therefore, if the backbone is under high strain – for example due to charge repulsion of equally charged residues, which might be possible in higher charge states – the intensity of this band might be increased.

### 3.4.4 Differences in the IR spectra between cytochrome c and tendamistat

Figure 3.12 shows the infrared spectra of protonated and deuterated cytochrome c in charge state 14+ together with the spectrum of tendamistat in charge state 5+. As described, a red-shift of the amide-I band is observed for tendamistat with respect to this charge state of cytochrome c. In solution, the main band position of the amide-I band in a $\beta$-sheet motif of the backbone is found at lower frequencies compared to an $\alpha$-helical motif [5, 173]. Usually, a second side band characteristic for a $\beta$-sheet structure is found at 1690 cm$^{-1}$ in solution. Upon transition from solution to the gas phase, a shift to higher energies of those bands can occur as described in Section 3.4. Based on the solution structure of tendamistat, nearly half of the amino acids are involved in a $\beta$-sheet motif (mostly antiparallel). The position of the amide-I band in the infrared spectrum of gas-phase tendamistat suggests an increased strength in the hydrogen-bonding network with respect to cytochrome c in charge state 14+. The similarity of the infrared spectra of tendamistat and protonated cytochrome c in low charge states (7+ to 10+) with respect to the amide-I band position contradicts big structural differences between the molecules (at least for the structures of cytochrome c in low charge states). If the amide-I band position for a $\beta$-sheet motif with respect to an $\alpha$-helical motif is shifted to lower energies in the gas phase as observed in solution, tendamistat might be partly in a $\beta$-sheet conformation if this structural motif is conserved or at least partly preserved upon transition to the solvent-free environment. However, salt bridges stabilizing a $\beta$-sheet motif might be destroyed in the gas-phase protein at charge state 5+ due to Coulombic repulsion of the charges.

The position of the amide-II band in the infrared spectrum of tendamistat is red-shifted compared to its position for cytochrome c. This usually indicates a weaker hydrogen bonding strength involving the backbone N−H groups for tendamistat with respect to cytochrome c in the higher charge states. Nevertheless, the amide-II band is found to be relatively insensitive towards conformational changes of the molecule [148, 149].

Coulombic charge repulsion inside the molecule might influence the observed structure. It is unclear where exactly the protons are located in tendamistat. Out of the 74 amino acids, six are basic (three arginine, two histidine, and one lysine) and eight are acidic (five aspartic acid and three glutamic acid). As a consequence, if all basic residues are protonated and all acidic residues de-protonated, the 'native' charge state would be 2- (according to CDNT). Due to the fact that tendamistat at charge state 5+ is detected in the mass spectrometer, it is most likely that all basic residues are protonated as well as seven out of eight acidic residues. Lower charged cytochrome c molecules reflect the charging pattern of tendamistat in charge state 5+ more than the higher charged cytochromes, where all acidic residues are protonated and additional charges

are present in the molecule.

As described in Section 3.4, amino acid side chain IR absorption together with absorption of the backbone groups could give a vibrational signature in the IR spectra. About 10% of all amino acids in tendamistat are tyrosine. Vibrations resulting from Tyr–OH could also cause the double peak structure in the amide-II band. Comparing the IR spectra of the deuterated cytochrome c proteins and tendamistat, the band at 1520 cm$^{-1}$ is observed in both spectra. As for cytochrome c, the amino acid side chain of tyrosine could lead to the peak at this position. The high intensity for this band in tendamistat compared to that observed in the spectrum of cytochrome c could be due to a higher relative (and absolute) number of tyrosine in the molecule. This behavior is found in IR spectra of condensed proteins [151]. The real peak position of the amide-II band is probably located at slightly higher frequencies if the band at 1520 cm$^{-1}$ is correctly assigned to an amino acide side chain vibration. It seems therefore feasible that the amide-II band can be assigned to the position where the shoulder in the peak is situated and consequently it is found at 1530 cm$^{-1}$.

## 3.5    Conclusions

Infrared spectroscopy is performed on charged proteins in the gas phase in the so-called 'fingerprint' region between 1400 cm$^{-1}$ and 1800 cm$^{-1}$. Protonated and deuterated cytochrome c from equine heart, known for a high $\alpha$-helical proportional content in the secondary structure in solution, is investigated in the gas phase in various charge states (7+ to 19+). The other protein is tendamistat whose secondary structure in solution is mostly a $\beta$-sheet motif. It is investigated in charge state 5+. A blue-shift of the amide-I mode is observed with increasing charge state for protonated and deuterated cytochrome c. The position is found for protonated cytochrome c in charge state 7+ at 1663 cm$^{-1}$ and in charge state 15+ at 1671 cm$^{-1}$. In the infrared spectra of the deuterated species, the amide-I band position is found for the protein in charge state 8+ at 1656 cm$^{-1}$ and is shifted with increasing charge state up to 1676 cm$^{-1}$ as observed in charge state 19+. The shift observed in the spectra of protonated and deuterated cytochrome c indicates a decrease in the hydrogen bonding network strength of the peptidic backbone with increasing charge state of the protein. In the infrared spectrum of tendamistat, the amide-I band position is observed at 1664 cm$^{-1}$.

A new band at 1483 cm$^{-1}$ starts to appear in charge state 12+ of protonated cytochrome c and the intensity of the band increases with increasing charge state. The band is observed in previously published work on cytochrome c from bovine heart and the origin is not clear yet. However, the observed band starts to appear at charge state 13+ of the protein from bovine heart. In the primary structure of cytochrome c from equine heart, more adjacent oppositely charged amino acids are present than in the sequence of cytochrome c from bovine heart. Formation of zwitterionic motifs due to those adjacent oppositely charged amino acids, for example a basic residue next to an acidic one, might locally disturb the structure of the protein leading to a slightly different conformation in both proteins at the same charge state. An indicator for this conclusion could be the

appearance of the band observed at 1483 cm$^{-1}$ in different charge states of both proteins. Amino acid side chain vibrations of proline residues can possibly be assigned to this band. Due to the absence of the band in the spectra of the deuterated species, an involvement of a hydrogen atom in the oscillator of this vibrational band is plausible.

The infrared spectrum of gas-phase tendamistat is similar to those observed for cytochrome c in low charge states. Other (especially lower) charge states than the recorded ones should be spectroscopically investigated and compared to the here presented ones. Structural changes are expected to be observed.

So far, only infrared spectra of relatively highly charged cytochrome c proteins in the gas phase are investigated. It is possible that the structures of those charged proteins are more elongated due to Coulombic repulsion of the charges compared to the native structure, where solvent molecules are present and interact with the protein. It is of importance to investigate gas-phase cytochrome c spectroscopically also in low charge states due to a possible reduction of the effect of Coulombic repulsions. More compact structures might be found and amide-I band positions should shift to lower frequencies. Those spectra should be compared to the infrared spectra of other proteins (for example tendamistat) to validate the significance of infrared spectroscopy on big molecular systems and sensitivity to secondary structure elements.

As shown in this Chapter, it is possible to record infrared spectra of proteins in the gas phase. However, calibration points in the infrared spectra as for example amide-I band positions for severals secondary structural motifs need to be found to allow statements regarding the three-dimensional structure of these big biological molecular systems.

In order to assign observed band positions to secondary structure elements with infrared spectroscopy, it is necessary to perform measurements of well defined systems in the gas phase with well-known structures. In order to allow for comparison of findings obtained by theory with experimental results, small systems are required. In the next Chapters of this work, spectroscopic investigations on small model systems with different secondary structure elements are presented.

# Chapter 4

# Polyalanine (based) peptides as a model for $\alpha$-helical systems in the gas phase

## 4.1 Introduction

One of the most common short-length structural motifs in proteins is the $\alpha$-helix [174]. This structural motif is characterized by a repeating pattern of hydrogen bonds between the $i$th residue amide carbonyl oxygen and the $i+4$th residue amide NH group, forming hydrogen-bonded rings of 13 atoms (named $C_{13}$). In contrast, the not so common $3_{10}$ helical motif involves ten atoms ($C_{10}$), resulting from a hydrogen bonding network between the amide groups of residues $i$ and $i+3$. The helical motif is also found to be the most frequent secondary structure element present in transmembrane proteins [175]. This is of interest because inside the membrane, hydrophobic interactions with the transmembrane protein occur. The gas phase provides a solvent-free, hydrophobic environment. Hence, no interactions with polar groups from the environment are present and electrostatic interactions are the driving force for the conformation of the molecule. Therefore, it is to be expected that intrinsic properties and intramolecular interaction of such molecules can be best investigated in the gas phase where no intermolecular interactions has to be considered.

Experiments are performed using a chain of alanine amino acids (polyalanine) as a potential $\alpha$-helical model system to study the physical properties of this secondary structure element. Among the amino acids, alanine is a strong $\alpha$-helical former [174] and in solution at physiological pH values, long-chain polyalanine peptides form $\alpha$-helices [176, 177]. In general, polypeptide chains form an electric macro-dipole in solution, if they are folded in an $\alpha$-helical conformation. Polar peptide residues are oriented such that their electrical dipole moments point along the axial direction [178]. The overall dipole moment of the helix, known as the helical macro-dipole, is the sum of all electric dipoles of the residues. The dipole points from the N-terminus to the C-terminus of the peptide.

Multicanonical simulations using the SMMP program package [179] have been performed to study helix formation of polyalanine peptides in a solvent-free environment. For the peptide $Ala_{10}$ an $\alpha$-helical conformation is found at low temperatures [180]. It is

known that interactions of charged residues within an amino acid chain, depending on the position, can favor or inhibit a helical formation. To control the location of the charge, a lysine residue can be incorporated into an alanine chain. The proton affinity of the lysine residue is higher than that of the alanine residue. When the lysine is located at the N-terminus of the peptide, the positively charged residue leads to an unfavorable interaction with the helical macro-dipole, therefore possibly disrupting a helical motif and favoring a charge-solvated globular conformation. On the other hand, when the lysine residue is located at the C-terminus of the polyalanine chain, charge – helix macro-dipole interactions favor an $\alpha$-helical motif [180, 181].

Ion mobility experiments in combination with molecular dynamics simulations on different protonated polyalanine peptides (with and without lysine residues) have been carried out by Jarrold *et al.* [182]. Their results will be discussed below and compared to the results obtained by infrared spectroscopy. Infrared-ultraviolet (IR-UV) double resonance spectroscopy on a series of gas-phase polyalanine peptides containing lysine and phenylalanine[1] was performed by Rizzo *et. al.* at low temperatures [183]. The conclusion of those studies will also be discussed below. To test the capability of mid-IR spectroscopy on peptides and proteins in the gas phase to obtain information on secondary structure elements in the three-dimensional arrangement of the molecule, well-characterized systems are examined. Infrared spectra of $Ala_5H^+$, $Ala_{15}H^+$, $Ala_{20}H^+$, $Ac-Ala_{20}H^+$, $Ac-Ala_{20}K^+$, $Ac-Ala_{20}Na^+$, $Ac-LysH^+-Ala_{19}$, and $Ac-Ala_n-LysH^+$ with n = 5, 10, 15, and 19 were recorded.

## 4.2   Experimental

The peptides are synthesized using solid phase peptide synthesis (SPPS) [184]. The synthesizing process works from the C-terminal to the N-terminal end on solid support by applying the Fmoc/tert-butyl protecting group strategy [185]. Synthesis is performed using an automatic peptide synthesizer Syro XP (MultiSynTech) equipped with 10 ml polypropylene reactors. The protocol used is shown in Table 4.1 and the corresponding chemicals for the abbreviations used are given in the footnote[2].
The utilized amino acid derivatives (Fmoc-L-Lys(Boc)-OH and Fmoc-L-Ala-OH) and resins (Fmoc-L-Lys-O-Wang, Fmoc-L-Ala-Lys-O-Wang) are purchased from Fa. Gerhardt and Bachem, respectively. Peptide coupling reagents are obtained from Acros (DIPEA), Fluka (DBU) and Fa. Gerhardt (HOBt, TCTU, HCTU). N-terminal acetylation is performed by incubation the resin three times with 20 ml 5% acetic anhydride and 5% DIPEA in DMF.
After automatic synthesis, the resin is removed from the reactor and washed three times with 10 ml dichloromethane. The washed resin is shaken with 3 ml cleavage solution (90% triflouroacetic acid (TFA), 10% triisopropylsilane (TIS) and 1% $H_2O$) for 4 h. Subsequently, the cleavage solution is filtered and the resin washed with 100% TFA,

---

[1]The latter amino acid serves as a chromophore to perform spectroscopy.
[2]Fmoc → 9-fluorenylmethoxycarbonyl; PIP → piperidine; DMF → dimethylformamide; HOBt → N-hydroxybenzotriazole; TCTU → see CAS # 330641-16-2; HCTU → see CAS # 330645-87-9; DIPEA → N,N-diisopropylethylamine; NMP → N-methyl-2-pyrrolidon

**Table 4.1:** Protocol for synthesizing the polyalanine peptides. The abbreviations for the reagents used here are given in the footnote[2].

| Process | Reagent | Time |
| --- | --- | --- |
| Fmoc deprotection | 2 ml 40% PIP in DMF | $2 \times 10$ min |
| Washing | 2 ml DMF | $6 \times 1$ min |
| Coupling | 4 eq. Fmoc-amino acid-OH + 4 eq. HOBt in DMF (0.8 ml) 4 eq. TCTU or HCTU in DMF (1.2 ml) and 4 eq. DIPEA in NMP (0.4 ml) | 30 min |
| Washing | 2 ml DMF | $6 \times 1$ min |
| Coupling | 4 eq. Fmoc-amino acid-OH + 4 eq. HOBt in DMF (0.8 ml) 4 eq. TCTU or HCTU in DMF (1.2 ml) and 4 eq. DIPEA in NMP (0.4 ml) | 30 min |
| Washing | 2 ml DMF | $6 \times 1$ min |

keeping the total amount less than 4 ml. For precipitation, up to 100 ml ice-cold and dried diethylether is added to the combined mixture. The resulting suspension is stored overnight at -20° C to ensure complete precipitation. Afterwards, the supernatant of the cold suspension is removed using centrifugation. The precipitated raw peptide is washed with 15 ml ice-cold diethylether and again centrifuged. This washing procedure is repeated seven times and the peptides are kept in air until they are completely dry. After synthesis of the peptides, mass spectra are recorded with a commercial Agilent 6120 LC mass spectrometer to control the produced product samples. It is confirmed by mass spectrometry that the desired peptide samples are synthesized but those mass spectra are not shown in this thesis.

The peptides are hardly soluble and the strong acid TFA needed to be used. The best ion signal intensity in the FT ICR is obtained by dissolving about 1 mg peptide in 1 ml TFA and 100 $\mu$l $H_2O$. When the peptides are dissolved in formic acid or in acetonitrile with TFA, a decrease in ion signal intensity is observed. To produce polyalanine with a $Na^+$ ion attached, about 10 $\mu$l of a 100 mM NaCl aqueos solution is added to the TFA/$H_2O$ solution. To attach a $K^+$ ion to the peptide, 150 $\mu$l of a 100 mM KCl aqueos solution is added to the TFA/$H_2O$ solution. All solvents and salts are purchased from Sigma Aldrich and used without further purification.
Chains of polyalanine peptides with varying numbers of alanine residues are synthesized. In Figure 4.1, mass spectra recorded in the FT ICR in the mass range from 200 amu to 2000 amu are shown for $Ala_5H^+$ (black), $Ala_{15}H^+$ (red), and $Ala_{20}H^+$ (blue). All mass spectra are averaged after ten recordings. To obtain the mass spectrum of the smallest peptide as shown in Figure 4.1, the transient length is set to about 32 ms. For the longer peptides, the transient length is set to about 131 ms. Beside the transient length for recording the mass spectra, all experimental conditions to record
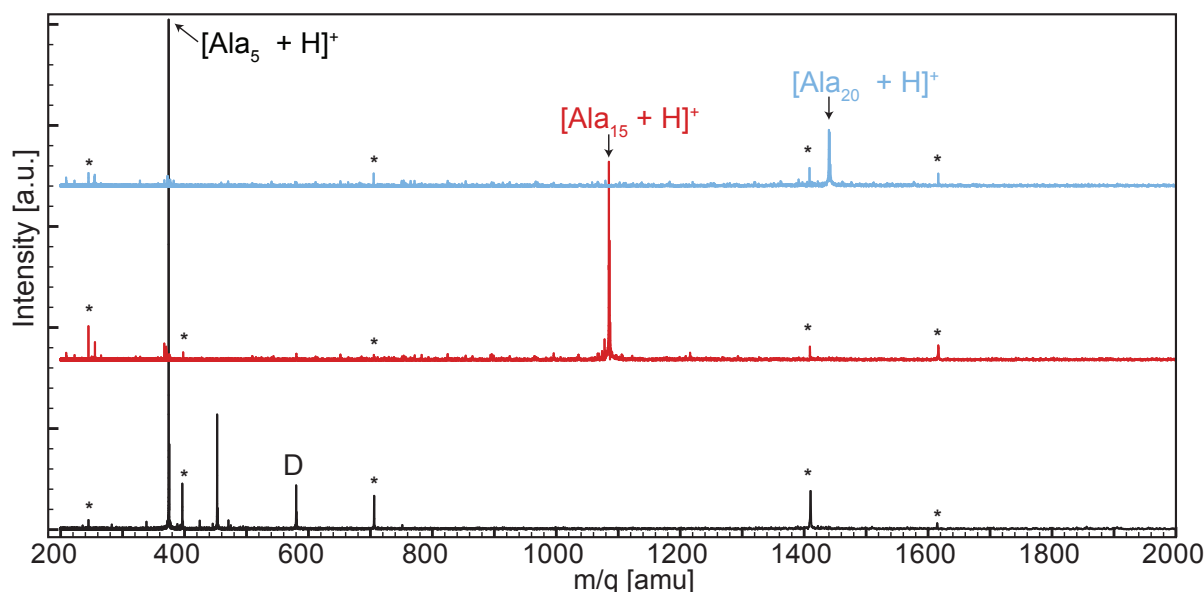
**Figure 4.1:** Mass spectrum of synthesized peptide samples of $Ala_5H^+$ (black, bottom), $Ala_{15}H^+$ (red, center), and $Ala_{20}H^+$ (blue, top). Noise peaks are marked by asterisks (see Section 1.4.4). Mass of $Ala_5H^+$: 374 amu; $Ala_{15}H^+$: 1085 amu; $Ala_{20}H^+$: 1440 amu. D marks a doubly charged unassigned ion as indicated by the isotopic distribution.

infrared spectra of those molecules are identical[3]. The most intense signal in the spectra correspond to the respective synthesized peptide. The ion signal intensity scale is the same in all three graphs and it is observed that the absolute signal intensity decreases with increasing peptide length. An asterisk marks noise peaks (see Section 1.4.4). In the mass spectrum of the shortest molecule, other unassigned mass peaks are present. Analyzing the isotopic distribution for the $Ala_5H^+$ sample, the peak observed at 453 amu corresponds to an unassigned monomer ion and the peak at 580 amu corresponds to an unassigned doubly charged ion (marked with D in the Figure).

Polyalanine peptides with a modification on the N-terminus are also synthesized. To prevent the proton being bound to the free amide group at the N-terminus, an exchange of the two protons of the functional group with an acetyl group ($COCH_3$) is performed via acetylation and the N-terminus is then capped[4]. Capping the N-terminus minimizes unfavorable interactions between the charge and the helix macro-dipole (see page 72)[5] and stabilization of a helical structure for the polyalanine peptides is expected. Mass spectra of $Ac-Ala_{20}H^+$ (black), $Ac-Ala_{20}K^+$ (violet), and $Ac-Ala_{20}Na^+$ (orange) are shown in Figure 4.2.
All mass spectra are recorded with a transient length of 131 ms and ten scans are averaged. On the left hand side of Figure 4.2, the mass range is set from 1000 amu to

---

[3]To record infrared spectra, the transient length is set to 32 ms for those peptides.
[4]Another possibility to prevent charging at the C-terminus is amidation, where the $O-H$ group is exchanged with a $NH_2$ group, however, this is not performed in the peptide samples in the current study.
[5]The situation here is slightly different than explained before: the capping prevents the binding of a proton at the N-terminus of the peptide. In the previously described situation, the proton is bound to a lysine residue located at either the N- or C-terminus of the peptide.
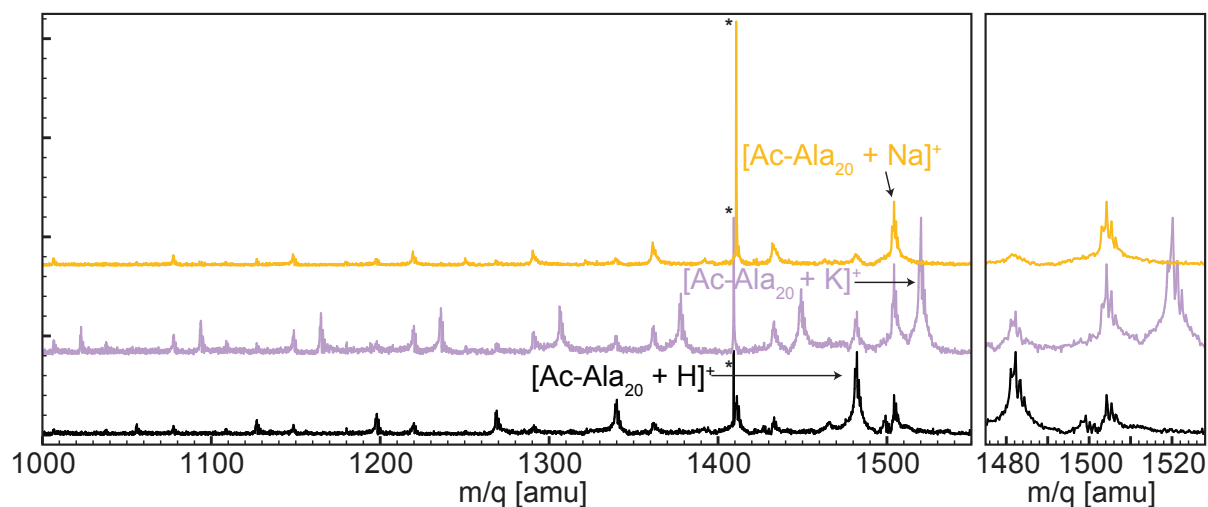
**Figure 4.2:** Mass spectrum of synthesized peptide samples of Ac−Ala$_{20}$X$^+$ with X corresponding to H (black, bottom), K (violet, center), or Na (orange, top) respectively. On the right, the mass range between 1475 amu and 1528 amu is shown and the isotopic distribution of Ac−Ala$_{20}$H$^+$ (1483 amu), Ac−Ala$_{20}$K$^+$ (1522 amu), and Ac−Ala$_{20}$Na$^+$ (1506 amu) is observed. Noise peaks are marked by an asterisk (see Section 1.4.4).

1550 amu. A distribution of mass peaks with a different amount of alanine residues is observed. The synthesized peptide sample is not purified and contain shorter peptide chains from the synthesizing process, which are observed in the mass spectrum. On the right side of Figure 4.2, the region from 1475 amu to 1528 amu is shown. The isotopic distribution of Ac−Ala$_{20}$X$^+$ with X corresponding to H, K, or Na is partially resolved. In the mass spectra of Ac−Ala$_{20}$ the protonated peptide mass peaks (bottom, black) and the potassiated peptide mass peaks (center, violet) show also contributions from attached sodium ions. Impurities of the glass container and of the syringe used to store and to spray the peptide solution are most likely the reason for this observation. By adding NaCl to the solution, the ion signal intensity of the protonated peak decreases and the intensity of the sodiated peak increases (see Figure 4.2 top part, orange).

A third group of peptides is investigated: In addition to the modification of the N-terminus by acetylation, one lysine (denoted as Lys, or, using the one letter code: 'K') residue at different positions in a polyalanine chain with varying chain lengths are also synthesized. The lysine is located either at the C-terminus of the polypeptide with a chain containing 5, 10, 15, or 19 alanine residues or at the capped N-terminus with 19 alanine amino acids following. The capped N-terminus prevents charging of this particular amide group. In those charged polyalanine peptides, it is reasonable to assume that the proton is located at the lysine residue. This is based on a higher gas-phase basicity of the residue (around 16 kcal/mol in energy) compared to the carboxamide C-terminus [186]. The charged lysine residue located at the C-terminus of the polyalanine peptide stabilizes a helical secondary structure by providing hydrogen bonding partners to the otherwise freely dangling C=O groups of the closest

alanine residues[6]. Mass spectra of $Ac-Ala_n-LysH^+$ with n = 5, 10, and 19 are shown in Figure 4.3. The mass spectrum of the $Ac-Ala_5-LysH^+$ (mass: 545 amu) sample is presented in black at the bottom of the Figure, the spectrum of the sample containing 10 alanine residues (mass: 900 amu) is shown at red in the center of the Figure and the mass spectrum of the 20 amino acid long peptide sample (mass: 1539 amu) is given in blue at the top of the Figure. In the latter spectrum, a size distribution of peptides containing less than 19 alanine residues is observed. The sample is used without further purification and in the synthesizing process, a mixture of shorter peptides is also produced.
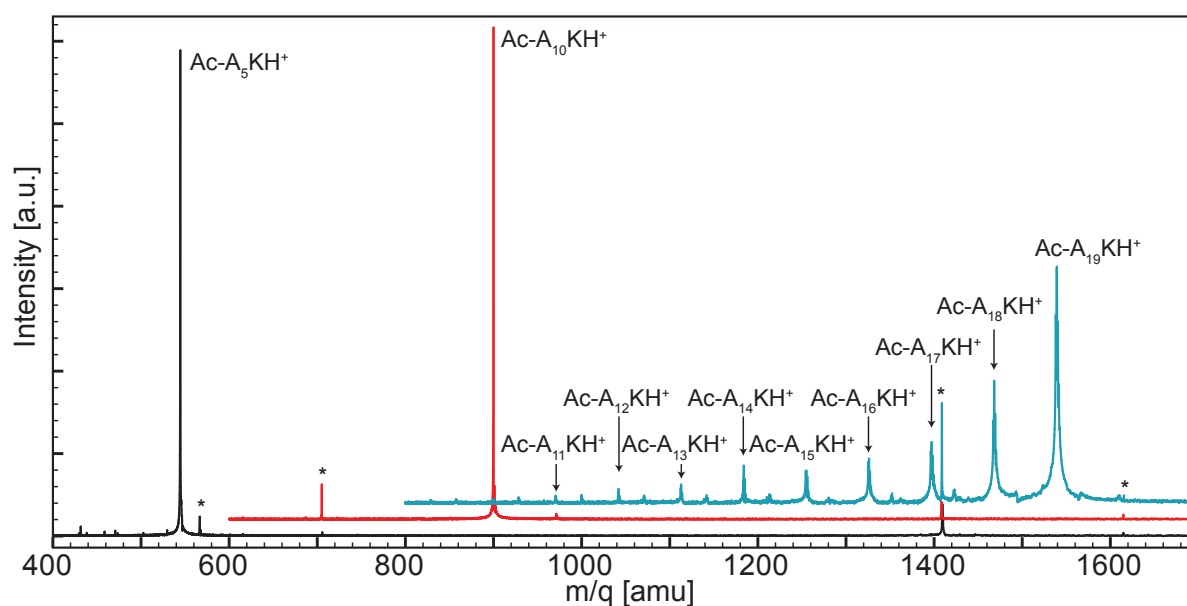


**Figure 4.3:** Mass spectrum of synthesized peptide samples of $Ac-Ala_n-LysH^+$ with n = 5 (black, bottom), 10 (red, center), and 19 (blue, top). In the mass spectrum of the longest peptide sample, a size distribution with a decreasing amount of alanine residues is observed. Noise peaks are marked by an asterisk (see Section 1.4.4).

The samples of the smaller peptides seem to have a higher degree of purity and the dominant peak corresponds to the synthesized peptide. Note that there is no mass spectrum shown for the synthesized peptide sample containing 15 alanine residues. The presented mass spectrum of the smallest polyalanine peptide is recorded using a 32 ms transient and averaged over 20 scans. To record the mass spectra of the longer peptides, a transient of 131 ms is used and ten scans are averaged for the $Ac-Ala_{10}-LysH^+$ sample, while 20 scans are averaged for the longest molecule sample.

In Figure 4.4, mass spectra of $Ac-Ala_{19}-LysH^+$ are shown with different transient lengths and theoretically at different resolutions (see Section 1.4.4). The transient length is 131 ms in the uppermost mass spectrum and is increased to 1048 ms in the lowest mass spectrum. Between the spectra, the transient length is doubled in each spectrum starting from the top. In the Figure on the right, the isotopic pattern of the peptide containing 19 alanine residues is shown. The observed isotopic distribution indicates a singly charged ion due to the spacing of the peaks by one amu. No peaks in

---

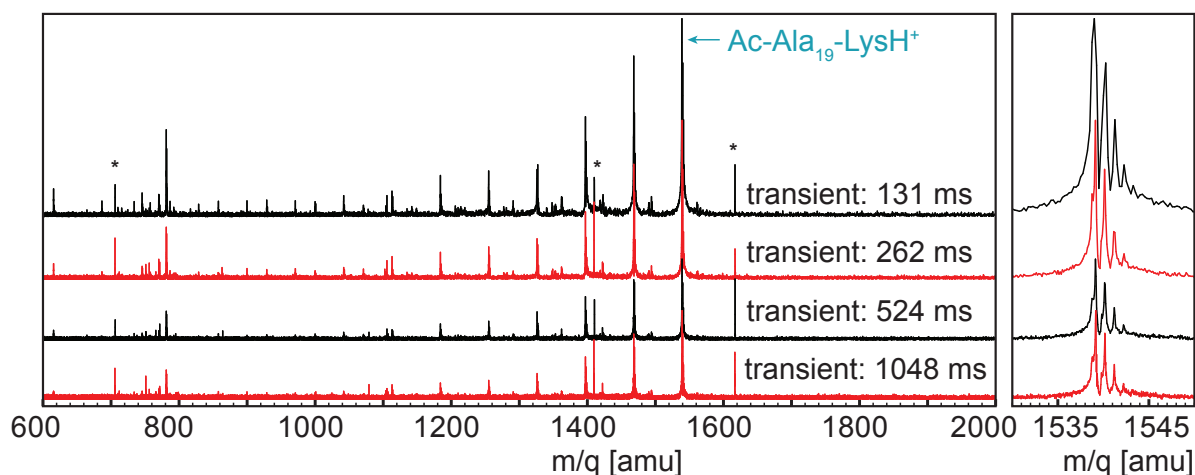[6]in addition to the favorable charge – helix macro-dipole interaction.

**Figure 4.4:** Mass spectra of synthesized peptide samples of Ac−Ala$_{19}$−LysH$^+$ for different transient lengths. From the top to the bottom, the transient length increases from 131 ms to 1048 ms. Noise peaks are marked by an asterisk (see Section 1.4.4).

between are observed, pointing to an absence of a dimer (two non-covalently bound Ac−Ala$_{19}$−LysH$^+$ peptides). It is observed that the resolution of the apparatus in this case is not limited by the transient length. The reason for low resolution is rather caused by the inherently high pressure in the ICR cell (see Chapter 1, Section 1.4.4).
In the experiments by Jarrold *et al.* dimers are also observed for the 20 amino acid long peptide [182]. Here, an analysis of the isotopic pattern shows that no doubly charged ions are present in the mass spectrum of Ac−Ala$_{19}$−LysH$^+$.

In Figure 4.5, mass spectra of Ac−LysH$^+$−Ala$_{19}$ taken under the same experimental conditions (transient lengths) are shown. In the close up view of the mass peak of



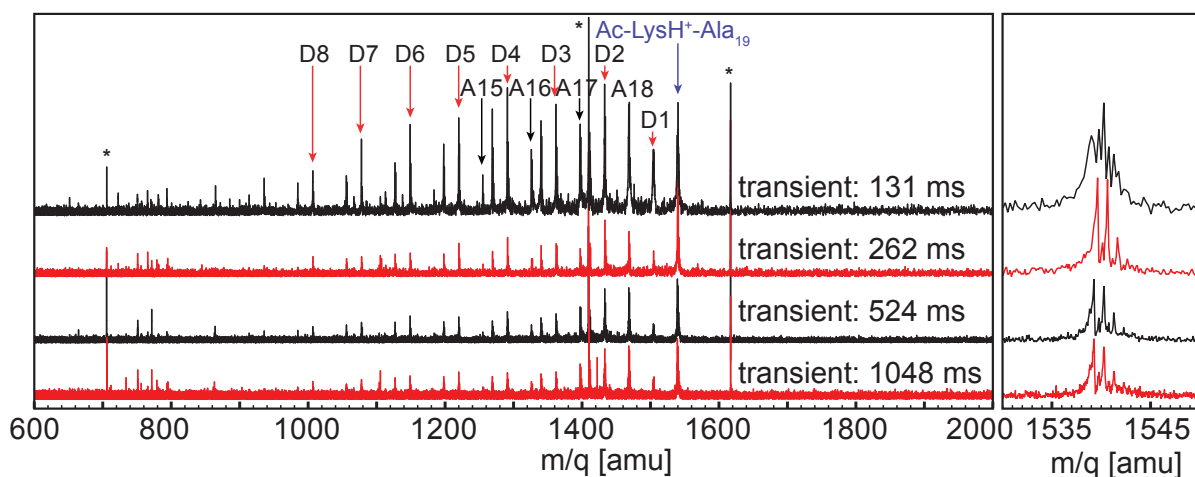**Figure 4.5:** Mass spectra of synthesized peptide samples of Ac−LysH$^+$−Ala$_{19}$ for different transient lengths. From the top to the bottom, the transient length increases from 131 ms to 1048 ms. Noise peaks are marked by an asterisk (see Section 1.4.4). The other symbols are explained in the text.

the polyalanine-based peptide with 19 alanine residues in the Figure on the right side, small additional peaks are observed in the isotopic distribution. These are resolved on transient length equal or longer than 262 ms. The spacing of the small peaks in the isotopic distribution between the more prominent ones is 0.5 amu, thus indicating a presence of a doubly charged ion on the same mass. This is in accordance with the observations reported by the Jarrold group [182] with one exception: In Jarrold's ion mobility based studies, no monomers of the peptide are present, which is not supported by the observations obtained here. The observed isotope pattern clearly results from one distribution spaced by one amu (singly charged) and a second distribution spaced by 0.5 amu (doubly charged). The intensity of the dimer in the mass spectrum relative to the monomeric ion is low. The mass peaks marked with A followed by a number n in the Figure 4.5 corresponds to the molecule $Ac-LysH^+-Ala_n$. The peaks marked with D and a number are due to dimeric ions in the mass spectra. D1 for example corresponds to a dimer consisting of $Ac-LysH^+-Ala_{19}$ and $Ac-LysH^+-Ala_{18}$, resulting in $[Ac-LysH-Ala_{19} + Ac-LysH-Ala_{18}]^{2+}$. The observed molecules with corresponding masses are shown in Table 4.2. The other peaks observed in the mass spectrum are not assigned.

**Table 4.2:** Assignment of mass peaks observed in the mass spectrum the synthesized sample of $Ac-LysH^+-Ala_{19}$ as observed in Figure 4.5.

| Molecule | Monomer m/q [amu] | Dimer m/q [amu] |
|---|---|---|
| $Ac-LysH^+-Ala_{19}$ | 1540 | |
| $[Ac-LysH-Ala_{19} + Ac-LysH-Ala_{18}]^{2+}$ (D1) | | 1504 |
| $Ac-LysH^+-Ala_{18}$ (A18) | 1469 | |
| $[Ac-LysH-Ala_{18} + Ac-LysH-Ala_{17}]^{2+}$ (D2) | | 1433 |
| $Ac-LysH^+-Ala_{17}$ (A17) | 1398 | |
| $[Ac-LysH-Ala_{17} + Ac-LysH-Ala_{16}]^{2+}$ (D3) | | 1362 |
| $Ac-LysH^+-Ala_{16}$ (A16) | 1327 | |
| $[Ac-LysH-Ala_{16} + Ac-LysH-Ala_{15}]^{2+}$ (D4) | | 1291 |
| $Ac-LysH^+-Ala_{15}$ (A15) | 1255 | |
| $[Ac-LysH-Ala_{15} + Ac-LysH-Ala_{14}]^{2+}$ (D5) | | 1220 |
| $Ac-LysH^+-Ala_{14}$ | 1184 | |
| $[Ac-LysH-Ala_{14} + Ac-LysH-Ala_{13}]^{2+}$ (D6) | | 1149 |
| $Ac-LysH^+-Ala_{13}$ | 1113 | |
| $[Ac-LysH-Ala_{13} + Ac-LysH-Ala_{12}]^{2+}$ (D7) | | 1078 |
| $Ac-LysH^+-Ala_{12}$ | 1042 | |
| $[Ac-LysH-Ala_{12} + Ac-LysH-Ala_{11}]^{2+}$ (D8) | | 1007 |
| $Ac-LysH^+-Ala_{11}$ | 971 | |

In Figure 4.6, two mass spectra of $Ac-LysH^+-Ala_{19}$ are shown in the m/q range of 1525 amu to 1555 amu. Both spectra are recorded using a 131 ms long transient and 50 scans are averaged. The spectrum presented on the left is recorded without infrared irradiation of FELIX in the ICR cell, whereas IR irradiation at about 6 $\mu$m
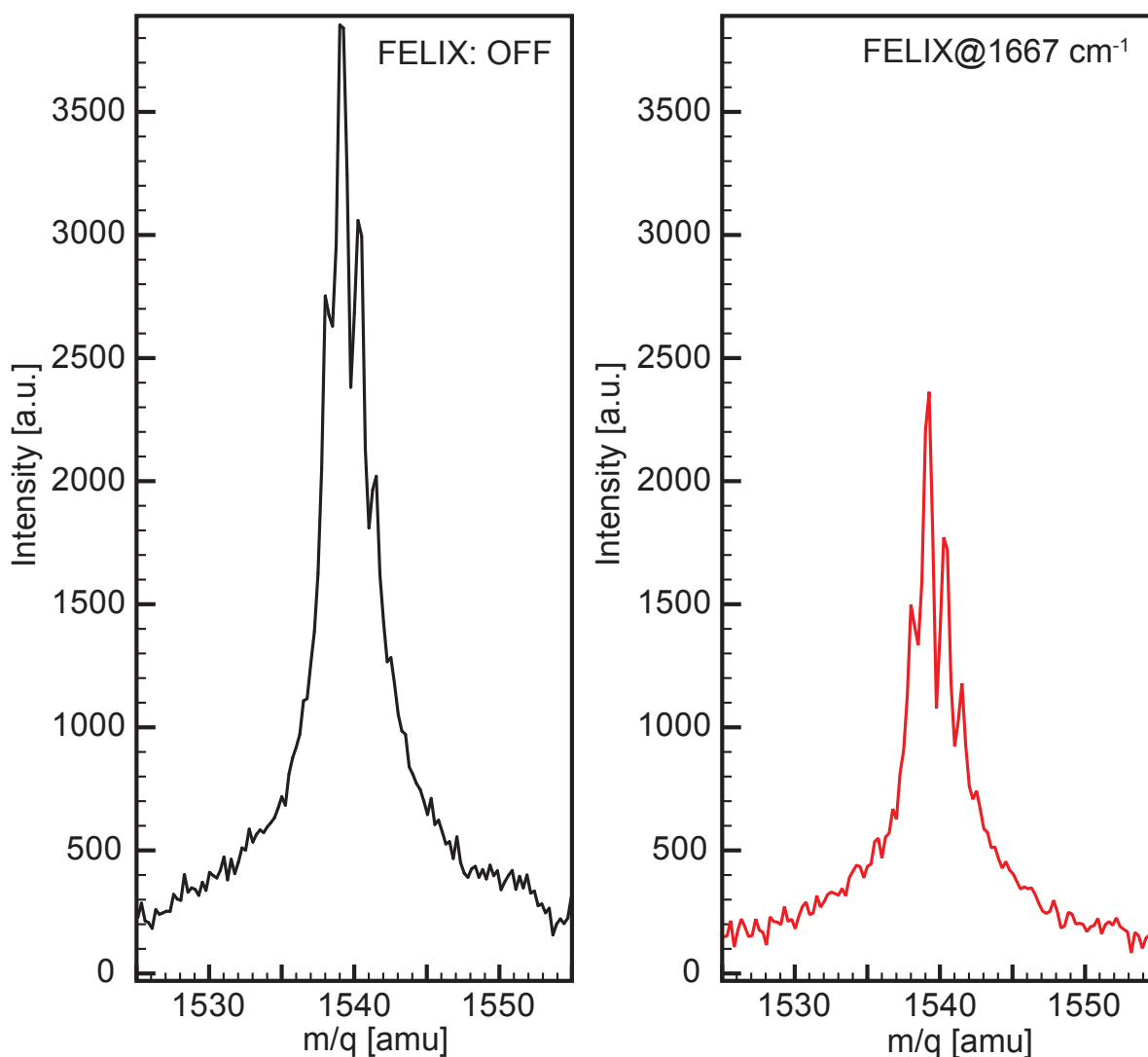
**Figure 4.6:** Mass spectra of synthesized peptide samples of $Ac-LysH^+-Ala_{19}$ in the region of 1525 amu to 1555 amu. On the left side, no infrared radiation is present in the ICR cell. On the right side, IR radiation at a wavelength of about 6 $\mu$m enters the ICR cell prior to recording of the mass spectrum. The ion signal intensity is depleted upon resonant absorption of the infrared light to about 65% of the initial value. The isotopic distribution does not change upon IR absorption. Unfortunately, the resolution is not high enough to completely exclude the presence of a doubly charged dimer in this mass range.

wavelength is present for four seconds prior to recording the spectrum shown on the right. The depletion in the ion signal intensity of about 30% indicates an absorption of infrared light and an infrared active transition of the molecule is observed. The resolution of the mass spectrometer does not allow for an accurate determination of the isotopic distribution, whether a dimer of the molecule is present or not. Therefore it cannot be excluded that the measured infrared spectrum contains contributions from the dimer. It is expected that the binding energy of the doubly charged non-covalently bound dimer is lower than the binding energy of the monomer. As a consequence it is

expected that upon resonant absorption of IR light, the dimeric ion would dissociate prior to the monomer. Possibly, only monomer ions appear here as fragments. The consequence would be an increase in the measured ion signal intensity of the monomer mass (=doubly charged dimer) if no other fragmentation occurs. Also, the underlying broader peak observed at this mass position would be expected to diminish or even vanish, and a higher relative intensity between the isotopic peaks spaced by one amu in the distribution should be observed. This is not the case and further, it is not clear whether the dimer dissociates only into a monomer. The monomer itself would be able to absorb at a molecular transition frequency and possibly dissociate. Therefore, it seems reasonable to deduce that the measured infrared spectrum corresponds to a superposition of the dimeric and monomeric ions.
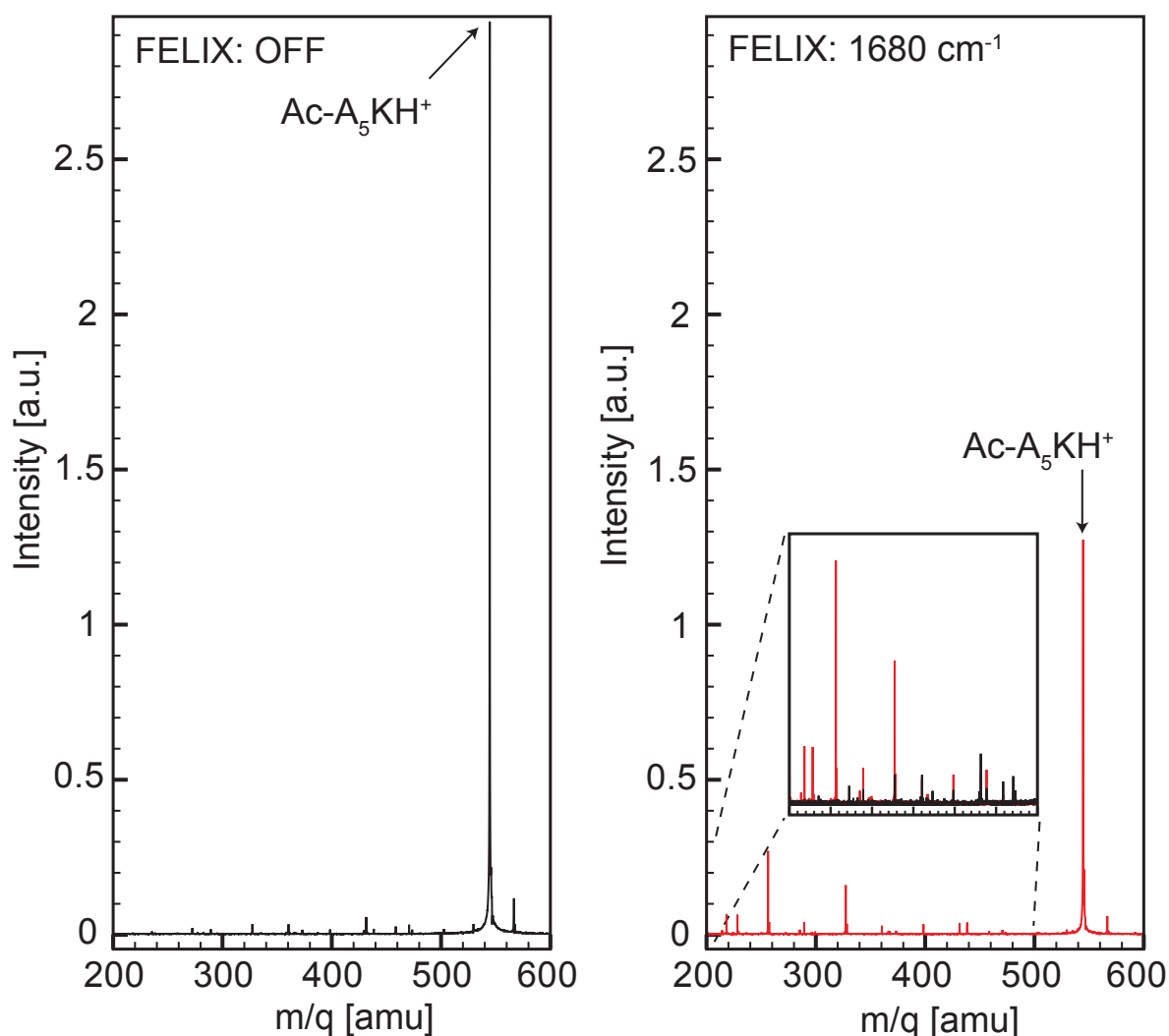
**Figure 4.7:** Mass spectra of synthesized peptide samples of $Ac-Ala_5-LysH^+$ in the region of 200 amu to 600 amu. On the left (black), no infrared light is radiated into the ICR cell. On the right (red), IR radiation with a wavelength of about 5.95 $\mu$m enters the ICR cell prior to recording of the mass spectrum. The inset shows both mass spectra in the range of 200 amu to 500 amu with an upscaled ordinate. A clear appearance of fragment ions upon resonant absorption of the molecule is observed.

In Figure 4.7, mass spectra of $Ac-Ala_5-LysH^+$ are shown. The Figure on the left (black) is recorded in the absence of infrared light. Shown on the right (red) is a spectrum recorded after irradiation of the molecules with light at a resonant transition. A clear depletion of the parent ion signal intensity up to about 45% upon resonant absorption is observed. The mass range between 200 amu and 500 amu is magnified and displayed in the small inset on the right. A clear appearance of weak mass peaks is observed. Those fragment mass peak intensities together with the parent ion signal intensity are taken to obtain the infrared spectrum of $Ac-Ala_5-LysH^+$. Table 4.3 gives an overview of the observed fragment ion peaks, which are used to obtain the infrared spectra of the peptides. The fragment ion molecules are not assigned. For the molecules $Ac-LysH^+-Ala_{19}$ and $Ac-Ala_{19}-LysH^+$, however, only a depletion of the parent ion signal intensity is observed.

**Table 4.3:** Measured fragment masses due to IR-MPD of polyalanine peptides on a resonant IR transition. The wavelength of irradiation from FELIX is set to about 6 $\mu$m, corresponding to 1667 cm$^{-1}$. The mass of the corresponding peptide is given in atomic mass units next to the peptide name in brackets. The presented fragments are used to obtain infrared spectra of the peptides. The fragment ion molecules are not assigned.

| Peptide | Fragment mass [amu] |
|---|---|
| $Ala_5H^+$ (374) | 115, 143, 161, 169, 214, 232, 240, 257 |
| $Ala_{15}H^+$ (1085) | 854, 924, 995 |
| $Ala_{20}H^+$ (1441) | 540, 711, 782, 825, 854, 895, 924, 996, 1138, 1209, 1280, 1351 |
| $Ac-Ala_{20}H^+$ (1483) | 540, 611, 682, 754, 825, 896, 966, 1037, 1108, 1180, 1321, 1393 |
| $Ac-Ala_{20}Na^+$ (1506) | only depletion of parent ion signal |
| $Ac-Ala_{20}K^+$ (1522) | appearance of $K^+$, mass 39 |
| $Ac-Ala_5-LysH^+$ (545) | 44, 84, 129, 130, 143, 147, 157, 185, 200, 218, 228, 256 |
| $Ac-Ala_{10}-LysH^+$ (900) | 129, 256, 327, 398, 399, 469, 540, 541, 611, 613, 682, 684 |
| $Ac-Ala_{15}-LysH^+$ (1255) | 147, 169, 185, 214, 256, 285, 311, 327, 398, 469, 540, 611, 682, 754, 825, 896, 967, 1149 |
| $Ac-Ala_{19}-LysH^+$ (1540) | only depletion of parent ion signal |
| $Ac-LysH^+-Ala_{19}$ (1540) | only depletion of parent ion signal |

## 4.3 Results

Infrared spectra of N-terminal modified and unmodified polyalanine peptides (with and without lysine residues) are shown and compared in this Section. Both differences and common features between the obtained spectra are highlighted.

### 4.3.1   Protonated polyalanine peptides

In Figure 4.8, infrared spectra of $Ala_5H^+$, $Ala_{15}H^+$, and $Ala_{20}H^+$ are shown. The number of the alanine residues increases from the spectrum presented at the bottom of the Figure to the top one. Spectra are generated by averaging at least two individual scans and for the regions marked in orange however, only one scan is recorded. In the spectrum of $Ala_{15}H^+$ at 1105 cm$^{-1}$ the drop of the signal is due to a short temporal breakdown of the infrared light.
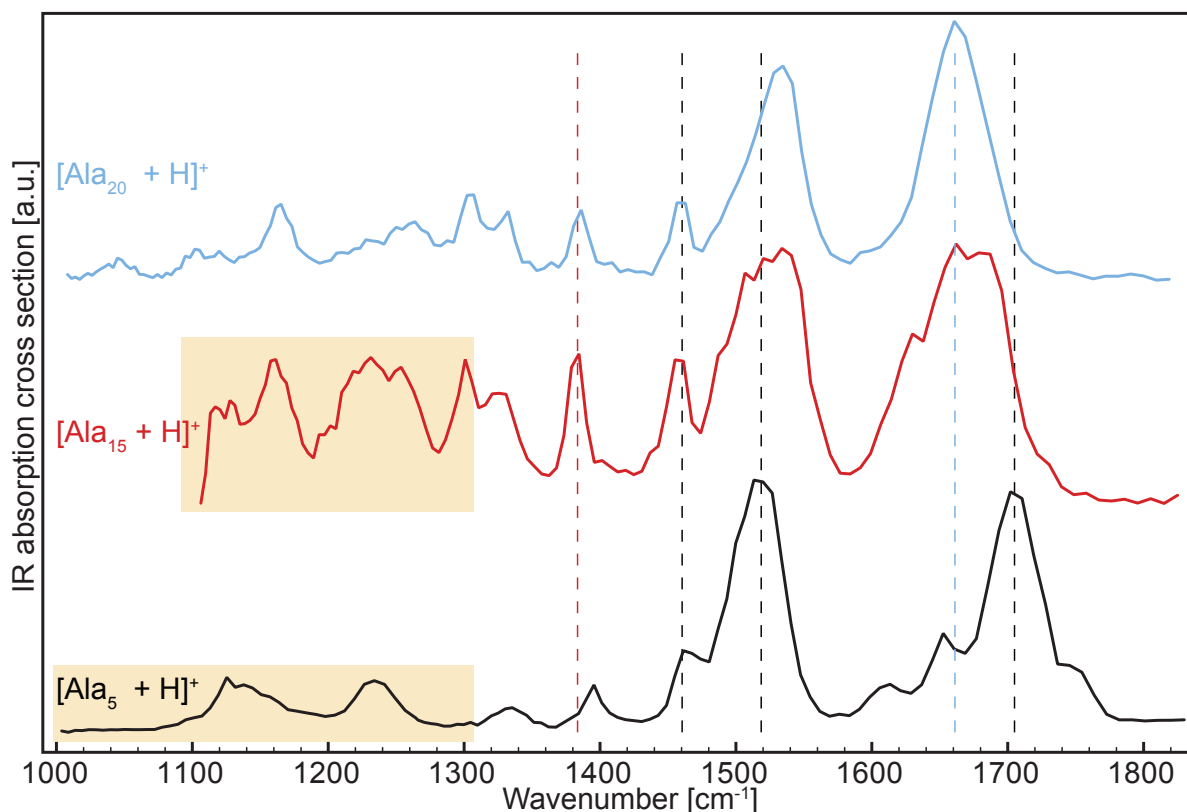


**Figure 4.8:** Infrared spectra of singly protonated polyalanine peptides. Spectra for $Ala_5H^+$ (black, bottom), $Ala_{15}H^+$ (red, center), and $Ala_{20}H^+$ (blue, top) are shown together. All spectra are normalized to the highest peak. The spectra are shifted up with increasing peptide length for better readability. The vertical dotted lines indicate assigned peak positions and these lines should guide the eye and underline differences and common features between the spectra. See text for further explanations.

The spectral feature observed in the infrared spectra of both longer peptides at around 1670 cm$^{-1}$ and in the spectrum of $Ala_5H^+$ at around 1700 cm$^{-1}$ can be assigned to the C=O stretching vibration (amide-I band) of the carboxyl groups of the peptidic backbone. In the spectrum of $Ala_5H^+$, four different underlying peaks in the amide-I region are resolved, located at 1613 cm$^{-1}$, 1652 cm$^{-1}$, 1704 cm$^{-1}$, and 1746 cm$^{-1}$. The amide-I band in the spectrum of the 15 alanine long peptide is relatively broad (FWHM 93 cm$^{-1}$), a small peak appears at the red side of the band at 1630 cm$^{-1}$ and a small dip is found in the amide-I band at 1671 cm$^{-1}$, separating two peaks at 1662 cm$^{-1}$ and 1682 cm$^{-1}$. Also, a small shoulder on the blue side of the amide-I band at 1726 cm$^{-1}$ is

observed. In the infrared spectrum of $Ala_{20}H^+$, the amide-I band is located at 1662 $cm^{-1}$ with a FWHM of 55 $cm^{-1}$ – narrower compared to the 15 alanine long peptide. No other underlying structural features are resolved.

The band located at 1516 $cm^{-1}$ in the spectrum of the shortest peptide as well as the band found at 1534 $cm^{-1}$ in the spectra of the longer peptides can be assigned to the $N-H$ bending vibration (amide-II band) of the amide groups of the peptidic backbone. The FWHM of the amide-II band in the infrared spectrum of $Ala_5H^+$ and of $Ala_{20}H^+$ is 45 $cm^{-1}$, whereas the FWHM of the band in the spectrum of $Ala_{15}H^+$ is about 30% larger (66 $cm^{-1}$). The position of the amide-II band is blue-shifted in the spectrum of the biggest peptide with respect to its position in the spectrum of the smallest peptide of about 20 $cm^{-1}$. In the infrared spectrum of $Ala_{15}H^+$, an underlying peak structure in the amide-II band is observed, meaning that the absolute position cannot be assigned with high accuracy. However, the band position seems blue-shifted with respect to that observed for $Ala_5H^+$.

The peak found at 1459 $cm^{-1}$ is present in the spectra of all three molecules and appears independent from the length of the alanine chain. A peak at the same position is found in IR-MPD studies of protonated dialanine [187] and is assigned to a $NH_3^+$ deformation vibration. Due to the universal character of this peak position it is reasonable to conclude that this assumption is valid and that this vibration might be insensitive to the intrinsic structure.

In the infrared spectra of the longer peptides, a peak is observed at 1384 $cm^{-1}$, which is blue-shifted by 11 $cm^{-1}$ up to 1395 $cm^{-1}$ in the spectrum of $Ala_5H^+$. In a deuterated aqueous solution, a symmetric in-plane bending vibration of the $CH_3$ group ($\delta_s(CH_3)$) of an alanine residue is found at 1378 $cm^{-1}$ [150]. In a solvent-free environment, this transition can be shifted and the observed peak positions can possibly assigned to that vibrational mode.

The other observed peaks in the spectra below 1365 $cm^{-1}$ could potentially be assigned to bending or wagging vibrations of CH and $CH_2$ groups, which are found in solution at [1315 - 1350] $cm^{-1}$ and [1170 - 1382] $cm^{-1}$ [150, 188]. Nevertheless, all the modes seem to be blue-shifted compared to the positions found in solution. A detailed assignment is not undertaken in this region due to the lack of other experimental data.

As mentioned before, a vibrational band is found in the infrared spectrum of $Ala_5H^+$ at 1746 $cm^{-1}$, which is not observed at this position in the spectra of the longer peptides. This band might be assigned to the vibrational mode of the terminal carboxyl, the C=O stretching vibration of the C-terminal carbonyl. For example, in studies on gas-phase di- and trialanine peptides with attached metal ion, it is observed that this mode shifts from 1650 $cm^{-1}$ for $Ala_2Ca^{2+}$ to 1750 $cm^{-1}$ for $Ala_2Cs^+$ with decreasing metal-ion coordination strength [189]. The peak is systematically higher in frequency for $Ala_3$ – metal ion complexes than the $Ala_2$ complexes, which is said to be due to the greater extent of chelation of the metal ion (by two oxygens of $Ala_3$) and therefore weakening the metal-ion interaction with the C-terminal carbonyl oxygen of the longer peptide compared to $Ala_2$. Obviously, no metal ions are present in the studies here. However, a similar interpretation for the appearance of the vibration in the spectrum of $Ala_5H^+$ at 1746 $cm^{-1}$ seams feasible: The C-terminal carbonyl might be not or at least only weakly hydrogen bounded in the shortest peptide and stronger bounded for the alanine chains whose vibrational frequencies of this particular oscillator are red-shifted and not

resolved in the peak structure observed at around 1660 cm$^{-1}$.
In the infrared spectra of the protonated polyalanine peptides, the band positions of the amide-I and amide-II band differ for the shortest molecule compared to the longer ones. A full discussion of these results is provided in Section 4.4.

The second group of peptides investigated, which contain only alanine residues are Ac−Ala$_{20}$ with either a proton or an alkali metal ion attached (most probably at the C-terminus). In Figure 4.9, infrared spectra of Ala$_{20}$H$^+$ (blue, bottom), Ac−Ala$_{20}$H$^+$ (black, center), Ac−Ala$_{20}$K$^+$ (violet, center)[7], and Ac−Ala$_{20}$Na$^+$ (orange, top) are shown. All spectra are normalized to the highest peak and shifted upwards for better comparability. The infrared spectrum of the sodiated molecule is measured in depletion. The spectra of the other molecules are given as appearance spectra and the measured fragment ion masses are listed in Table 4.3. Note that in the spectrum of Ac−Ala$_{20}$K$^+$, only the appearance of K$^+$ ions is recorded. For all molecules, at least two scans are recorded and averaged.
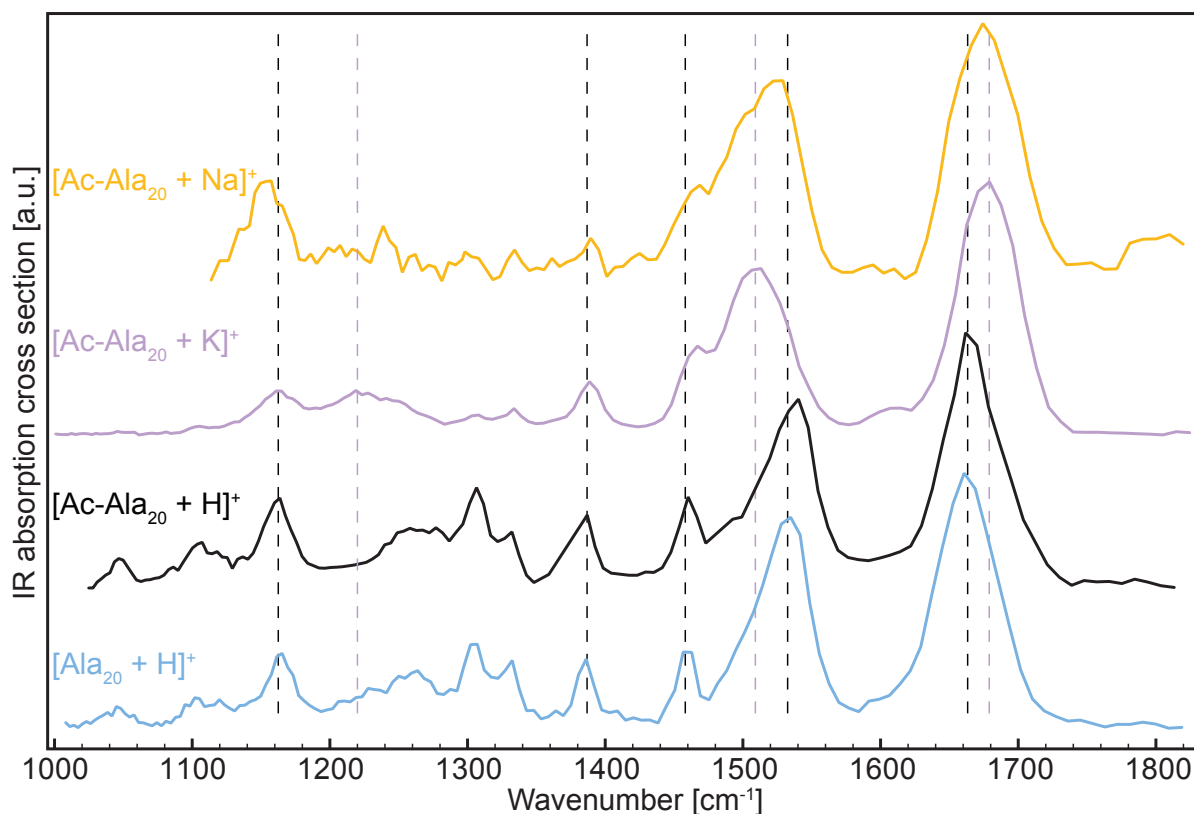


**Figure 4.9:** Infrared spectra of polyalanine peptides containing 20 alanine residues with and without an acetyl group at the N-terminus. Spectra of Ala$_{20}$H$^+$ (blue, bottom), Ac−Ala$_{20}$H$^+$ (black, center), Ac−Ala$_{20}$K$^+$ (violet, center), and Ac−Ala$_{20}$Na$^+$ (orange, top) are shown. All spectra are normalized to the highest peak and the spectra are shifted upwards for better readability. Dotted lines should guide the eye and underline differences and commonalities between the spectra. Further explanations are given in the text.

---

[7]Note: K$^+$ being a potassium ion

The measured infrared spectra of $Ala_{20}H^+$ and $Ac-Ala_{20}H^+$ appear to be identical. For the molecules with an attached alkali ion, the amide-I band in the infrared spectra is blue-shifted with respect to the protonated counterparts. The position of the amide-I band in the spectrum of $Ac-Ala_{20}K^+$ is found at 1680 $cm^{-1}$ and the FWHM of the band is 52 $cm^{-1}$, comparable to the width of the amide-I band for the protonated alanine peptide. In the spectrum of the sodiated polyalanine peptide, the amide-I band is located at 1675 $cm^{-1}$ and the width of the band is 62 $cm^{-1}$, which is slightly broadened. The latter spectrum, however, is calculated only taken the depletion of the parent ion signal intensity upon resonant absorption and the signal-to-noise ratio is usually lower[8] compared to appearance spectra. In the infrared spectra of the sodiated or potassiated peptide it appears to arise a small peak at 1605 $cm^{-1}$.

A red shift of the amide-II band position is observed in the infrared spectra of the alkali ion attached peptides compared to the protonated ones. The band observed for the potassiated molecule is located at 1510 $cm^{-1}$ with a FWHM of about 61 $cm^{-1}$ (measured from the dip at 1476 $cm^{-1}$ to 1537 $cm^{-1}$). The band position in the spectrum corresponding to the sodiated molecule is observed at 1522 $cm^{-1}$, with a width of 71 $cm^{-1}$. The amide-II band is broader in the spectra of the peptides with attached alkali ion compared to the protonated ones (FWHM about 46 $cm^{-1}$).
The peak observed in the infrared spectra of the protonated species at 1459 $cm^{-1}$ is blue-shifted by about 8 $cm^{-1}$ to 1467 $cm^{-1}$ compared to the position in the spectra of potassiated/sodiated molecules. The gap between the amide-II band and this peak is shallow in the latter group of molecules, indeed the 1467 $cm^{-1}$ vibration appears more as a shoulder of the amide-II band.
In all shown infrared spectra a peak is present at 1384 $cm^{-1}$ with a roughly constant relative intensity.
In the infrared spectrum of $Ac-Ala_{20}K^+$, a broad band at 1218 $cm^{-1}$ is observed, which is absent in the other spectra (the signal-to-noise ratio in the spectrum of $Ac-Ala_{20}Na^+$ is low and accurately, no statement of the presence or absence of this band can be given). The appearance of this feature might be connected with an unassigned interaction of the alkali metal ion with binding sites of the peptide.
A relatively strong absorption of IR light is observed at 1162 $cm^{-1}$ in all spectra. In solution, wagging vibrations of $CH_2$ groups ($\gamma_w(CH_2)$) are found in the range of [1170 - 1382] $cm^{-1}$ [150]. This vibrational mode couples with adjacent $CH_2$ groups and is sensitive towards the hydrocarbon chain conformation [150]. Some other vibrations are also found in solution of pure alanine, see [188]. Due to the manifold of different vibrations in this wavelength region as well as the presence of many chemical groups with IR active modes, it is difficult to accurately assign the observed peaks.

### 4.3.2 Lysine containing polyalanine peptides

Polyalanine based peptides containing one lysine residue, either at the N-terminus or the C-terminus (see Figure 4.10) are investigated in the gas phase by infrared spectroscopy.

---

[8]due to ion signal fluctuations during the scan

**Figure 4.10:** This image of a peptide with 15 alanine residues and one lysine residue at the C-terminus is based on a calculated structure by M. Rossi.
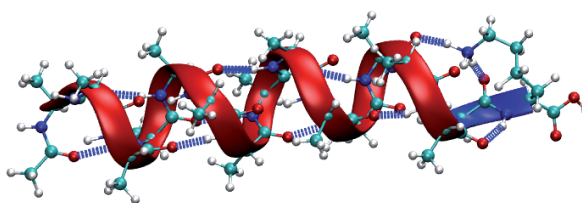
Figure 4.11 shows infrared spectra of protonated singly charged peptides with the lysine located at the C-terminus. The N-terminus is capped with an acetyl group to prevent protonation at this position. The infrared spectra of $Ac-Ala_5-LysH^+$ (black, bottom), $Ac-Ala_{10}-LysH^+$ (red, center), $Ac-Ala_{15}-LysH^+$ (green, center), and $Ac-Ala_{19}-LysH^+$ (light blue, top) are presented. The infrared spectrum of the longest peptide is determined by depletion of the parent ion signal intensity and the spectra of the other molecules are calculated by the appearance of fragment ions. The masses of the fragment ions used to calculate the IR absorption cross section are shown in Table 4.3. All spectra are normalized to the highest peak intensity and the spectra are shifted upwards for better comparability. All spectra are recorded at least twice and the scans are averaged. At a first glance, all the spectra appear rather similar.

In all these spectra, the amide-I band at appears around 1680 cm$^{-1}$ and the amide-II band appears at around 1524 cm$^{-1}$. Figure 4.12 shows the spectra between 1600 cm$^{-1}$ and 1750 cm$^{-1}$ in the region of the amide-I band. With increasing chain length of the peptide, the band position is shifted to lower wavenumbers in the infrared spectra. The peak position for the peptide with 5 alanine residues is located at 1691 cm$^{-1}$ (FWHM: 58 cm$^{-1}$), with 10 alanine residues at 1683 cm$^{-1}$ (FWHM: 54 cm$^{-1}$), with 15 alanine residues at 1677 cm$^{-1}$ (FWHM: 45 cm$^{-1}$), and with 19 alanine residues at 1667 cm$^{-1}$ (FWHM: 61 cm$^{-1}$).

In the infrared spectrum of $Ac-Ala_5-LysH^+$ a small peak is observed at 1790 cm$^{-1}$. This peak might be assigned to the C=O stretching vibration of the C-terminal carbonyl group, similar to the peak at 1746 cm$^{-1}$ observed in the spectrum of $Ala_5H^+$ (see page 83). A similar trend is found for the amide-II band position in the different peptides. With increasing chain length of the molecule, this band shifts to higher wavenumbers although the shift is not as strong as that observed for the amide-I band position. The amide-II band position is found for the peptide with 5 alanine residues at 1518 cm$^{-1}$ (FWHM: 68 cm$^{-1}$), with 10 alanine residues at 1523 cm$^{-1}$ (FWHM: 48 cm$^{-1}$), with 15 alanine residues at 1527 cm$^{-1}$ (FWHM: 42 cm$^{-1}$), and with 19 alanine residues at 1529 cm$^{-1}$ (FWHM: 61 cm$^{-1}$).

In infrared spectra of the different molecules in the lower wavenumber regime, there are tiny differences in peak positions observed but not further discussed here. For an advanced understanding of the molecular structure, theoretical model calculations need to be performed and calculated infrared spectra need to be compared with the experimental spectra recorded here.

Figure 4.13 shows infrared spectra of both $Ac-Ala_{20}H^+$ (black, bottom) and $Ac-Ala_{19}-LysH^+$ (light blue, top). Both peptides contain 20 amino acids and have the same N-terminal
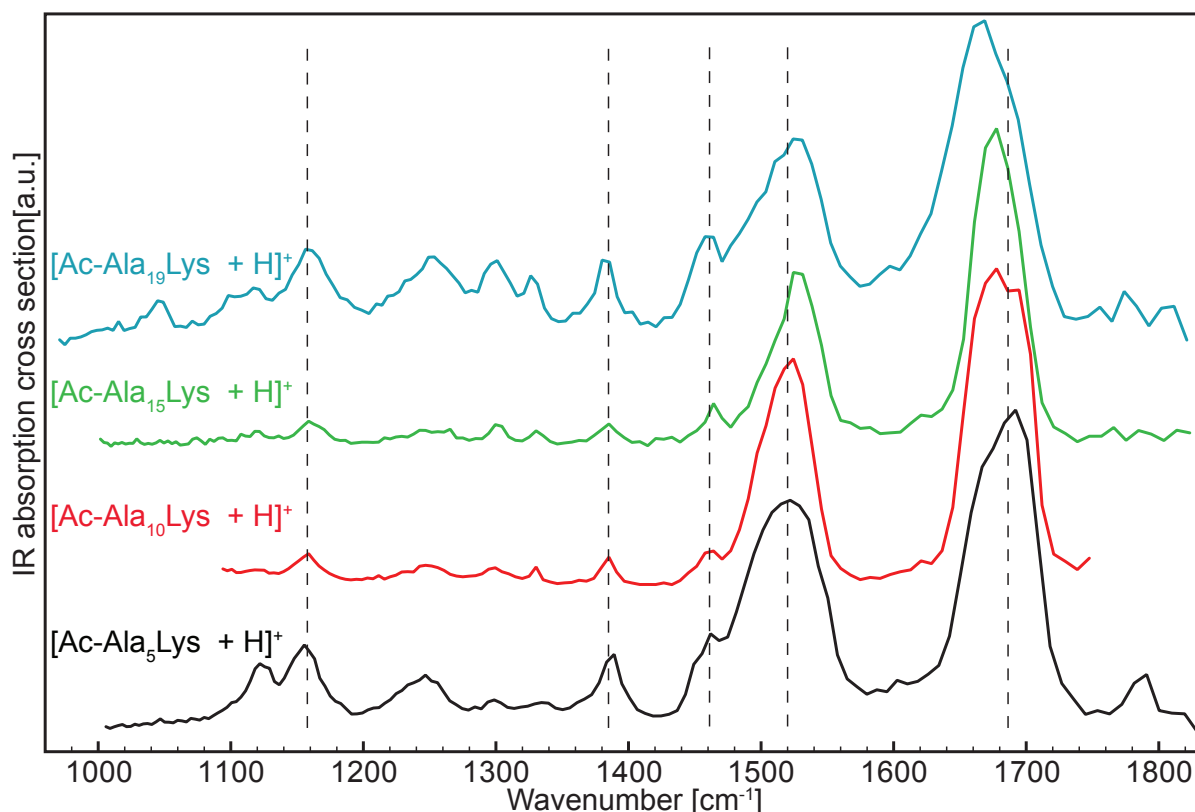
**Figure 4.11:** Infrared spectra of polyalanine based peptides with lysine at the C-terminus. Ac−Ala$_5$−LysH$^+$ (black, bottom), Ac−Ala$_{10}$−LysH$^+$ (red, center), Ac−Ala$_{15}$−LysH$^+$ (green, center), and Ac−Ala$_{19}$−LysH$^+$ (light blue, top) are investigated. All spectra are normalized with respect to the highest peak intensity and the spectra are shifted upwards for better readability. Dotted lines are guidelines for the eye and emphasize differences and common features between the spectra. Further explanations are given in the text.

cap modification – the only difference is the exchanged amino acid at the C-terminus (Ala → Lys). It is important particularly for Ac−Ala$_{19}$−LysH$^+$ that the protonation site is located close to the C-terminus because of a stabilization of an $\alpha$-helical motif. Dotted lines with an "x" mark peaks at identical positions in both spectra and red dotted lines without a x mark differences the infrared spectra of the molecules. The amide-I band is found to be narrower in the spectrum of the pure alanine chain peptide (difference in FWHM of about 12 cm$^{-1}$) but its position seems to be unshifted. For the other peptide the amide-II band position is red-shifted by about 15 cm$^{-1}$ and appears broader (the difference of the FWHM is about 15 cm$^{-1}$). A possible explanation for the slightly broadened amide-I and amide-II band in the lysine containing peptide might be explained by the presence of an infrared active mode of lysine. A symmetric ($\delta_s(NH_3^+)$) and antisymmetric ($\delta_{as}(NH_3^+)$) deformation vibration of the NH$_3^+$ group is found in solution at [1526 - 1527] cm$^{-1}$ and [1626 - 1629] cm$^{-1}$ respectively [150]. The peak in the spectrum of Ac−Ala$_{20}$H$^+$ at 1307 cm$^{-1}$ is red-shifted by about 8 cm$^{-1}$ in the spectrum of the other molecule and exhibits a reduced relative intensity. For the other peaks, the relative intensities are the same in both spectra.

**Figure 4.12:** Infrared spectra of Ac−Ala$_5$−LysH$^+$ (black, bottom), Ac−Ala$_{10}$−LysH$^+$ (red, center), Ac−Ala$_{15}$−LysH$^+$ (green, center), and Ac−Ala$_{19}$−LysH$^+$ (light blue, top) in the amide-I band region. All spectra are normalized with respect to the highest peak intensity and the spectra are shifted upwards for better readability. The color of the dotted lines represents the assignment of the corresponding molecule to the peak position.
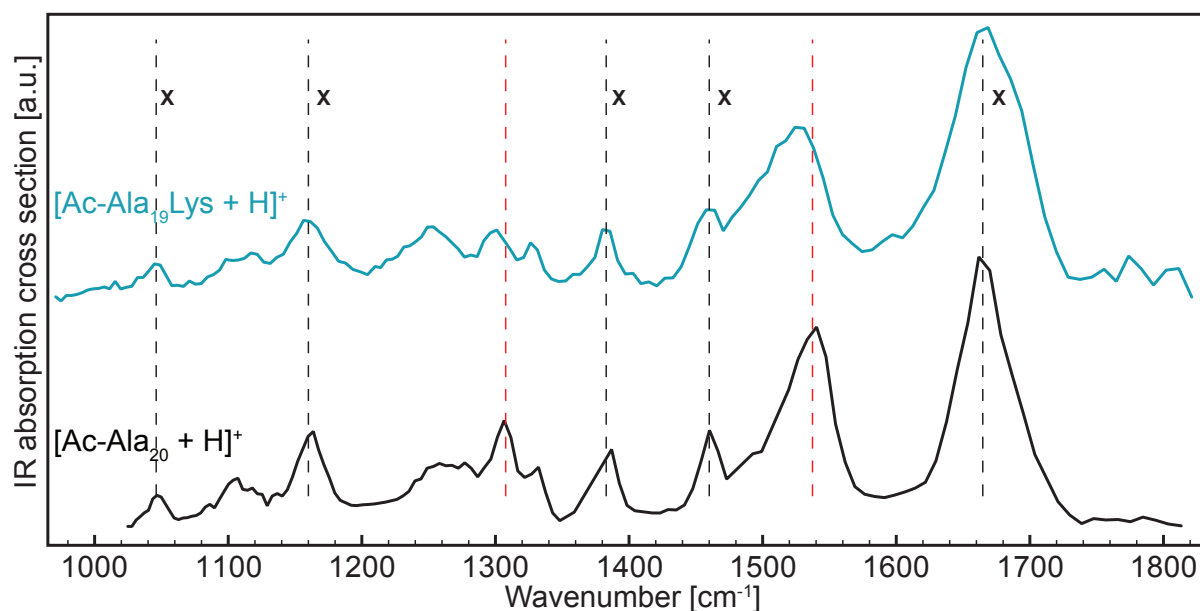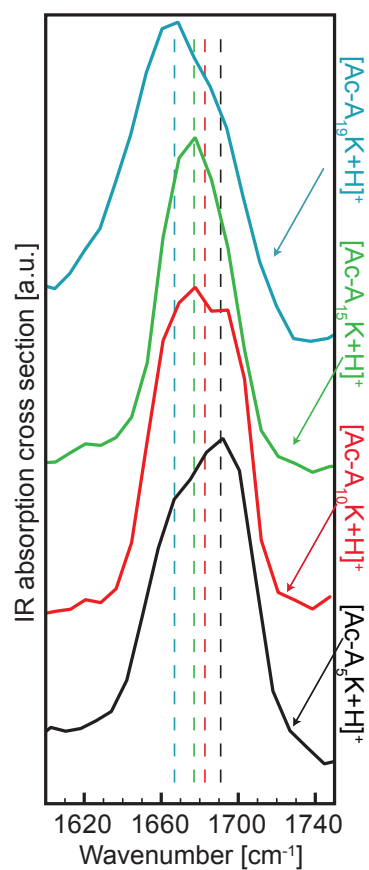
**Figure 4.13:** Infrared spectra of Ac−Ala$_{20}$H$^+$ (black, bottom) and Ac−Ala$_{19}$−LysH$^+$ (light blue, top). Both spectra are normalized to the highest peak intensity and the spectrum of Ac−Ala$_{19}$−LysH$^+$ is shifted upwards for better readability.

To study the influence of the lysine position in the peptide chain, an infrared spectrum

is measured for a molecule where the lysine residue is replaced from the C-terminus to the N-terminus giving $Ac-LysH^+-Ala_{19}$. In the literature it is stated that for a peptide chain with this sequence, a globular structure is found. The IR absorption cross section is calculated by the depletion of the parent ion signal intensity. In Figure 4.14, infrared spectra of $Ac-LysH^+-Ala_{19}$ (dark blue, bottom) and $Ac-Ala_{19}-LysH^+$ (light blue, top) are shown together.



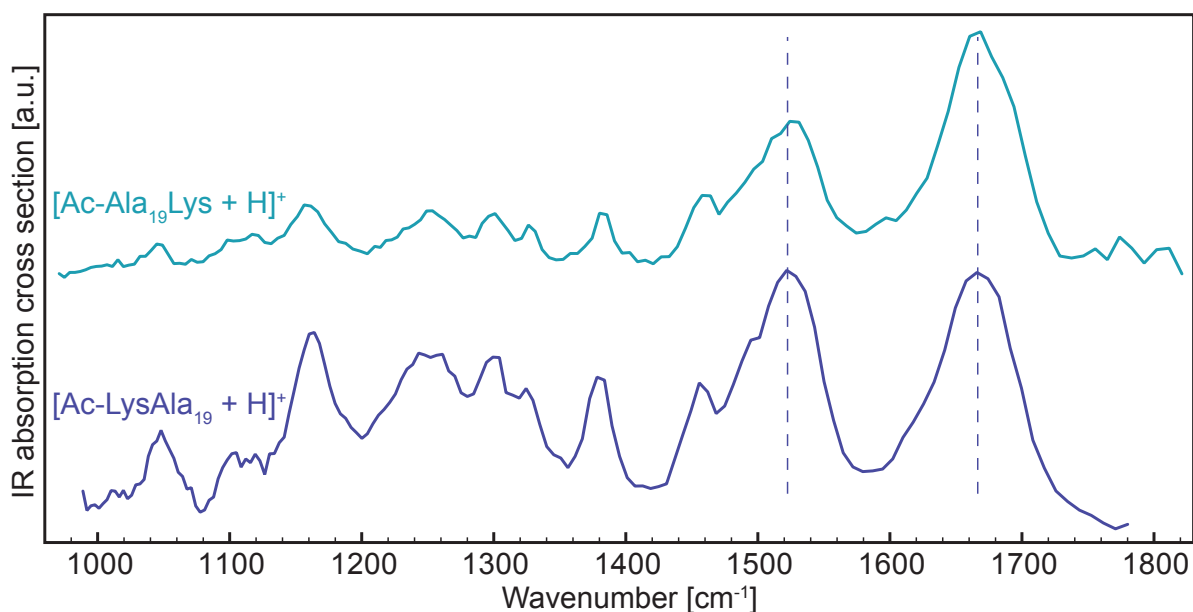**Figure 4.14:** Infrared spectra of $Ac-LysH^+-Ala_{19}$ (dark blue, bottom) and $Ac-Ala_{19}-LysH^+$ (light blue, top). Both spectra are normalized to the highest peak intensity and the spectrum of $Ac-Ala_{19}-LysH^+$ is shifted upwards for better readability.

Both infrared spectra appear identical apart from small exceptions: Firstly, the FWHM of the amide-I band of $Ac-LysH^+-Ala_{19}$ is about 71 cm$^{-1}$ with the FWHM of the same band observed in the spectrum of $Ac-Ala_{19}-LysH^+$ being about 10 cm$^{-1}$ narrower (61 cm$^{-1}$). The narrowing occurs at the red side of the band and in both spectra the peak is found to disappear at the same wavenumber on the blue side. One possible reason for the wider band in the spectrum of the molecule where the lysine residue is located at the C-terminus might be a difference in the H-bonding network involving the lysine residue. The antisymmetric ($\delta_{as}(NH_3^+)$) deformation vibration of the $NH_3^+$ group might be shifted in $Ac-Ala_{19}-LysH^+$ due to lysine serving as a hydrogen bonding partner to the next alanine residues (thus, stabilizing a helical macro-dipole of the molecule). It is possible that the hydrogens in the case of the N-terminus located lysine are involved differently in the hydrogen bonding network with the alanine residues because a charge located at the N-terminus would lead to an unfavorable charge – helix macro-dipole interactions. Secondly, the observed relative intensities of the peaks in the spectra of the two peptides are different. The amide-II band position is slightly red-shifted with respect to the position found for $Ac-Ala_{19}-LysH^+$ and is observed at 1521 cm$^{-1}$.

## 4.4    Discussion

Among the infrared spectra of the protonated polyalanine peptides containing 5, 15, and 20 alanine residues, differences are observed (see Figure 4.8). A blue-shift of the amide-I band position together with a red-shift of the amide-II band position in the spectrum of the shortest polyalanine peptide with respect to the longer ones indicate a weaker hydrogen bonding network for the shortest alanine chain. As found in spectroscopic studies of small uncapped polyalanine peptides with two or three residues, several different conformations are present in the gas phase. A superposition of the different structures is measured and it is concluded from calculations that the present structures are mostly globular (with a minor contribution of linear structures) [190, 191]. In these studies, it is indicated that the protonation occur as well on the backbone carbonyl oxygen for these short peptides and a mixture of N- and O-protonated conformations, which likely interconvert is present. The most favored O-protonated species have a nearly linear backbone, which is stabilized due to hydrogen bonds and charge delocalization. Infrared studies for a protonated dialanine molecule in the amide-I and amide-II region together with computational studies, nevertheless conclude that a structure with the protonation site at the N-terminus is favored [187].

In the region of the amide-I mode, a double peak structure is found at 1700 cm$^{-1}$ (strong) and 1760 cm$^{-1}$ (stronger), and is considerably blue-shifted compared to the position found for the longer peptides (more than five residues) in the study here. The peak observed at 1700 cm$^{-1}$ is consistent with the assigned amide-I mode of Ala$_5$H$^+$ (1704 cm$^{-1}$). A weak shoulder at 1746 cm$^{-1}$ is also present in the spectrum of Ala$_5$H$^+$. This position is clearly red-shifted compared to the vibrational mode observed at 1760 cm$^{-1}$ for the dialanine peptide. In the spectrum of the dialanine peptide, the amide-II band is found at 1525 cm$^{-1}$ and is blue-shifted by nine wavenumbers compared to Ala$_5$H$^+$. The infrared spectrum of the calculated lowest energy conformation of protonated dialanine, which has a planar structure fits best to the experimental spectrum. By comparing the relative band positions of the amide-I and amide-II band as observed in the spectra of Ala$_5$H$^+$ and the protonated dialanine peptide, a stronger hydrogen bonding network strength is found for the longer molecule. This suggests a more compact structure of Ala$_5$H$^+$ than that of the smaller peptide.

In a spectroscopic study in the N−H stretching region on polyalanine peptides, it is observed that the protonation site in Ala$_5$H$^+$ is located at the N-terminus, and the observed conformations are charge solvated [192]. This is in agreement with the results obtained for the amide-I and amide-II band positions found here.

Computational studies on protonated polyalanine peptides with 8, 14, and 17 alanine residues show that protonation of the N-terminus – especially for the shortest peptide – does not lead to $\alpha$-helical structures [193]. The helical character of the peptide is more favored, if protonation occurs close to the C-terminus. In Figure 4.9, infrared spectra of N-terminal protected polyalanine peptides and an unprotected polyalanine peptide are shown together and no differences are observed. This leads to the conclusion that the charge might not be located at the N-terminus (either in the case of the unprotected peptide).

It is also known that the solvent plays an important role for whether an alanine based

peptide possesses a helical structure or not [194, 195] and that protection of the N- and C-terminus promotes helix formation [196]. Helical structure promoting solvents are, for example, water with trifluoroethanol rather than pure water.

So far, the only structural study on protonated polyalanine peptides in the gas phase is performed in the group of Jarrold in 1999, using ion mobility experiments (see Chapter 3, Section 3.4.2 for a short introduction to this method). The measured collision cross sections compared with results obtained by molecular dynamics (MD) simulations indicates self-solvated globular conformations of the peptides. No $\alpha$-helical conformations of $Ala_nH^+$ peptides for n = 5 to 20 are observed [197]. However, when instead of the proton a singly charged metal ion ($Li^+$, $Na^+$, $K^+$, ..) binds to the molecule, $\alpha$-helical conformations for peptides with n $\geq$ 12 are observed [198, 199]. This is due to the location of the metal ion near the C-terminus of the peptide, which stabilizes a helical structure element with favorable interactions of the charge to the helix macro-dipole. The similarity of the infrared spectra of N-terminal protected and unprotected peptides indicates, as mentioned, that the binding site of the proton is not necessarily at the N-terminus. Helical conformations of protonated peptides in the gas phase have been reported by other groups: It is observed with mass analyzed ion kinetic energy (MIKE) spectrometry, performed on multiply protonated melittin (a 26-residue long peptide), that the helical conformation of the molecule is retained in the gas phase [200].

From the infrared spectra observed here, the hydrogen bonding network involving the backbone amide- and carboxyl groups appears to be stronger for the longer peptides (15 or 20 alanine residues) compared to the shorter one (5 alanine residues). Also, in the presence of a singly charged alkali metal ion instead of a proton, the amide-I band position is slightly blue-shifted, indicating a weaker hydrogen bonding network strength for that molecule. The blue-shift of the band position in the spectra of the alkali metal ion attached polyalanine peptides with respect to the position observed for $Ala_{20}H^+$ is not as strong as for $Ala_5H^+$. Therefore it is not clear whether the obtained results are in agreement with the results of the ion mobility studies as there appear to be some differences.

One important note to make is that the amide-I and amide-II band are relatively broad in the infrared spectrum of $Ala_{15}H^+$ with respect to $Ala_5H^+$ and $Ala_{20}H^+$ respectively. This might be an indication for a larger number of different conformers of $Ala_{15}H^+$ with respect to the others at 300 K in the gas phase. In order to assign those structures, DFT calculations on the investigated peptides need to be performed and infrared spectra need to be calculated and compared to the experimentally measured spectra. Since the measured infrared spectrum for each peptide is a sum of the individual infrared spectra of all present conformers in the ICR cell at 300 K, one must include different conformations in the calculations and compute the superposition of their individual infrared spectra.

Addressing the series of investigated polyalanine peptides containing one lysine residue, either located at the N- or to the C- terminus of the molecule, it can be observed that the lysine residue next at the C-terminus favors a helical structure in two ways. Firstly, if the lysine residue is protonated, the charge located at this position of the

peptide interacts 'positively' with the helix macro-dipole. Secondly, the lysine side chain provides hydrogen bonding partners for the last carbonyls of the alanine residues and it is believed that those bonding partners stabilize helical conformations within proteins [201, 202]. On the other hand, the location of the charged lysine at the N-terminus leads to unfavorable interactions of the charge with the helix macro-dipole and the carbonyls of the alanine residues next to the C-terminus lack hydrogen bonding partners. Also multicanonical simulations in which helix-coil transitions in the gas phase are investigated emphasize the importance of the location of the lysine residue [180]. Ion mobility studies (IMS) are performed on peptides with a varying number of alanine residues and the lysine located either at the C-terminus or N-terminus [182, 203, 204], the results of which, when compared to MD calculations, indicate the presence of a helical motif when the charged lysine residue is located at the C-terminus of the peptide for sufficiently long alanine chains. Ion mobility studies on those peptides with adsorbed water molecules indicate that the transition from a globular conformation to a helical conformation occurs when at least $8 - 10$ alanine residues are present in the peptide [205]. In the unsolvated peptides, it is found for $Ala_nLysH^+$ that peptides with $n < 7$ adapt a globular conformation, while those with 7 or more alanine residues exhibit a helical conformation [182]. Experiments are also performed at various temperatures – as high as 725 K – and it is observed that the $\alpha$-helical structural motif is conserved at even these high temperatures in a 15 alanine long peptide with a lysine residue at the C-terminus [206]. For $Ac-LysH^+-Ala_n$ globular conformations are found for $n < 13$ whereas only dimers are observed for $n > 13$. When the peptide $Ac-LysH^+-Ala_n$ is aggregated into dimers, the conformation is predominantly helical [182]. Figure 4.15 is adapted from [204] and shows drift time distributions for protonated $Ac-Ala_{15}-LysH^+$ and $Ac-LysH^+-Ala_{15}$. The arrow points in the direction of the helix dipole from the C-terminus to the N-terminus in the molecule $Ac-Ala_{15}-LysH^+$ (a) and conformations obtained by MD calculations of the dimer $[Ac-LysH^+-Ala_{15}]_2^{2+}$ (b) together with a conformation of the globular monomer of $Ac-LysH^+-Ala_{15}$ (c) are shown.

Some of the published results by the group of Jarrold on polyalanine based peptides containing lysine are different from the results obtained in the FT ICR apparatus: In the ion mobility studies, mass spectra of $Ac-LysH^+-Ala_n$ with $n > 13$ show **only** dimers [182] whereas in the FT ICR apparatus monomers of the corresponding molecules are observed, indicated by the isotopic distribution pattern in the mass spectra. It is not possible to separate $M^+$ from $M_2^{2+}$ ions in the FT ICR setup because the corresponding peaks appear at the same m/q position in the mass spectrometer.
Nevertheless, it is likely that the IR spectrum recorded for $Ac-LysH^+-Ala_{19}$ corresponds to a superposition of monomer and dimer ($[Ac-LysH-Ala_{19}]_2^{2+}$) ions present in the ICR cell. For the ion mobility experiments, it is found that the dimers of the peptide $Ac-LysH^+-Ala_{19}$ appear to be helical. The lysine from one molecule interacts with the C-terminus from the other molecule and *vice versa*, thus forming a 'head-to-toe' arrangement of antiparallel helices. This observation could explain why no differences are observed in the infrared spectra of the peptides where the lysine is adjacent to either of the two termini. The primary cause of the signal depletion could result from the weakly bound dimer and so, a helical conformation is most probably probed by infrared spectroscopy.
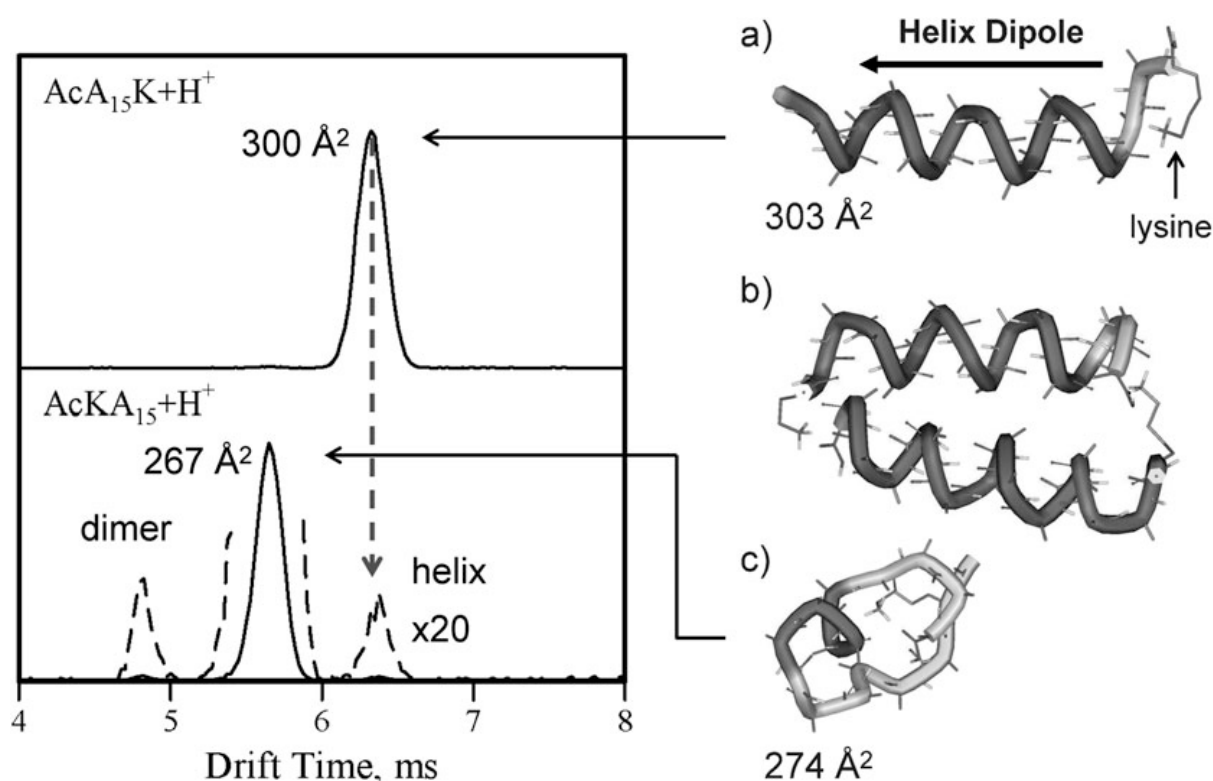
**Figure 4.15:** Drift time distribution measured for protonated $Ac-Ala_{15}-LysH^+$ and $Ac-LysH^+-Ala_{15}$. The structures on the right are snap shots taken from MD simulations. (a) shows an $Ac-Ala_{15}-LysH^+$ helix; (b) shows an $Ac-LysH^+-Ala_{15}$ helical dimer, and (c) shows an $Ac-LysH^+-Ala_{15}$ globule. This picture is presented with permission of M. Jarrold and taken from [204] (Figure 2 therein).

In the group of Rizzo, conformation-specific IR-UV double resonance spectroscopy in a cold ion trap (around 10 K) is performed on alanine based (5 or 10 alanine residues), lysine containing peptides with one additional phenylalanine residue, serving as a chromophore for spectroscopy [207, 208]. The spectra are recorded in the O−H and amide N−H stretching region between 3100 $cm^{-1}$ and 3600 $cm^{-1}$. Nitrogen-15 isotopic substitution is used for the smaller peptides to identify different conformers. DFT calculations are performed and calculated spectra are assigned to the experimentally obtained spectra [183]. Comparison of infrared spectra between the molecules $Ac-Phe-Ala_5-LysH^+$ and $Ac-Phe-Ala_{10}-LysH^+$ lead to the conclusion that infrared transitions between 3320 $cm^{-1}$ and 3350 are due to internal N−H groups involved in an $\alpha$-helical motif as found in the longer peptide [207]. However, those transitions are also found in the infrared spectrum of the peptide $Ac-LysH^+-Phe-Ala_{10}$, which should adapt a more globular conformation due to the lysine located at the N-terminus of the peptide [183]. It seems that the band positions for the N−H groups in that region are not specifically diagnostic for an $\alpha$-helical peptide. The vibrations found there might be due to a short helical turn in $Ac-LysH^+-Phe-Ala_{10}$ or represent medium-strength hydrogen bonded N−H groups whose frequencies are located in the same region as the vibrations of molecular groups involved in an $\alpha$-helical motif [183]. It has

been concluded, the band position of carboxylic acid OH group is a more decisive indicator for an $\alpha$-helical motif present in the peptide. Its position is red-shifted by 300 cm$^{-1}$ in the globular peptide (lysine residue at the N-terminus) relative to one in the helical peptide (lysine residue at the C-terminus). This observation, together with the observations obtained by infrared spectroscopy in the amide-I and amide-II region, lead to the conclusion that theoretical calculations with a higher level of theory need to be performed to obtain better structures and better infrared spectra. It seems intruding to investigate a broader range of the infrared spectrum and include formerly excluded regions in the energy range. There might be different indicators necessary to assign structures and to distinguish between structural motifs, for example the position of the stretching vibration of the OH group as to take only the amide 'standard' vibrations, for example the C=O stretching vibration into account.

### 4.4.1   Results obtained from theory compared to measured IR spectra

Computational studies on polyalanine based peptides with a lysine residue are carried out in the group of Matthias Scheffler by Mariana Rossi and Volker Blum in the Fritz-Haber-Institut. Calculated infrared spectra for different conformers of the shortest protonated peptide $Ac-Ala_5-LysH^+$ are compared to the measured infrared spectrum [209]. A very detailed description of the theoretical methods used and the obtained results is given in the doctoral thesis of Mariana Rossi [28]. In this Section, only a brief summary on the results compared to the measured spectra is given. DFT using the FHI-aims program package [210] in the PBE [211] generalized gradient approximation corrected for van der Waals (vdW) interactions [212] is applied to find low energy structures to a highly accurate degree. Helical assignment is confirmed for $Ac-Ala_n-LysH^+$ with n = 10 and 15, while the smallest peptide investigated with n = 5 is found not to be a simple helix. Infrared spectra are calculated – including anharmonicities in the vibrations – from ab initio molecular dynamics (AIMD) trajectories. To quantify the consistency between theory and experiment, the Pendry reliability factor $R$ is used [213]. The lowest energy conformation found for the molecule with 5 alanine residues is not an $\alpha$-helix, but rather a globular conformation named g-1 (see Figure 4.16).
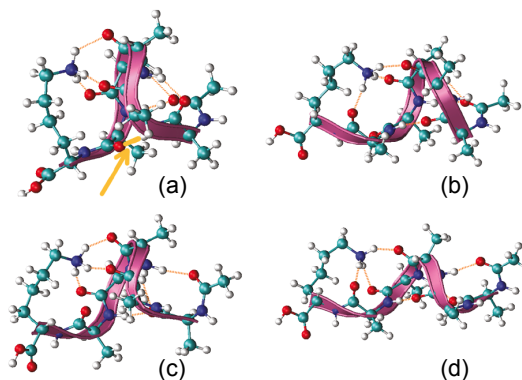


**Figure 4.16:** Different calculated conformers of $Ac-Ala_5-LysH^+$. (a) g-1, (b) $\alpha$-1, (c) $\alpha$-2, (d) $3_{10}$-1. This Figure is adapted from [209] (Figure 2 therein) and further details are explained there.

In the energy hierarchy, $\alpha$-helical conformations ($\alpha$-1 and $\alpha$-2) and a $3_{10}$-helical conformation ($3_{10}$-1) are close to the lowest energy structure (about 90 to 110 meV higher as

found with DFT-PBE + vdW). However, the calculated infrared spectrum that fits the experimental data for the smallest investigated peptide best corresponds to a mixture of different conformers, namely 25% g-1, 60% $\alpha$-1, 15% $\alpha$-2 and no contribution from $3_{10}$-1. Figure 4.17 presents the calculated infrared spectra for the different conformers together with the actual experimental results.
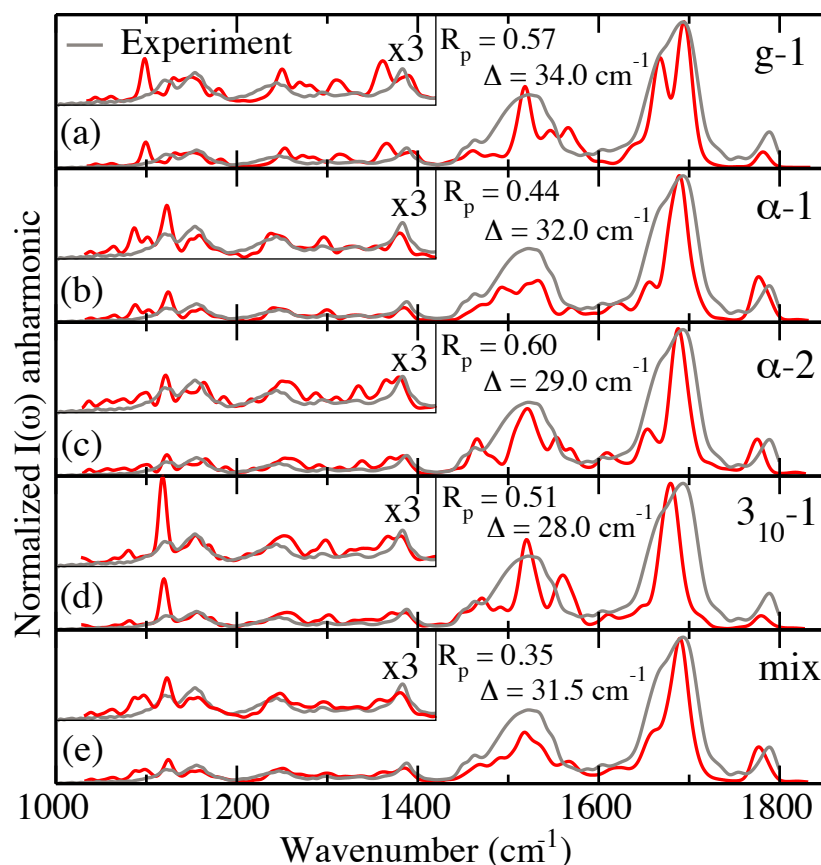


**Figure 4.17:** Comparison between experimental (gray lines) and theoretical (red lines) vibrational spectra, all normalized to 1 for the highest peak. (a,b) Ac−Ala$_{15}$−LysH$^+$: calculated spectra based on the harmonic approximation, for a $3_{10}$-helical (a) and an $\alpha$-helical (b) local minimum of the potential energy surface. (c) Ac−Ala$_{15}$−LysH$^+$: calculated spectrum from AIMD (including anharmonic effects), starting from an $\alpha$-helix and $\alpha$-helical in character throughout the simulation. (d) Same as panel c, for Ac−Ala$_{10}$−LysH$^+$. Pendry $R$-factors and rigid shifts $\Delta$ between measured and calculated spectra are included in each graph (calculated spectra are shifted by $\Delta$ for visual comparison). This Figure is adapted from [209] (Figure 1 therein) and further details are explained there.

In Figure 4.18, the calculated infrared spectra for the peptides containing 10 and 15 alanine residues together with the experimental spectra are shown. The calculated infrared spectra for the longer peptides – based on only one $\alpha$-helical conformer – agree very well with the experimental spectra. Particularly, the small bands between 1000 cm$^{-1}$ and 1400 cm$^{-1}$ are almost reproduced quantitatively.

**Figure 4.18:** Ac−Ala$_5$−LysH$^+$: (a-d) Theoretical anharmonic vibrational spectra from AIMD trajectories (red lines) for the four chosen conformers (see text), compared with experiment (gray line); (e) optimum calculated spectrum when assuming a coexistence of more than one conformer in experiment. Pendry $R$-factors and rigid shifts $\Delta$ between measured and 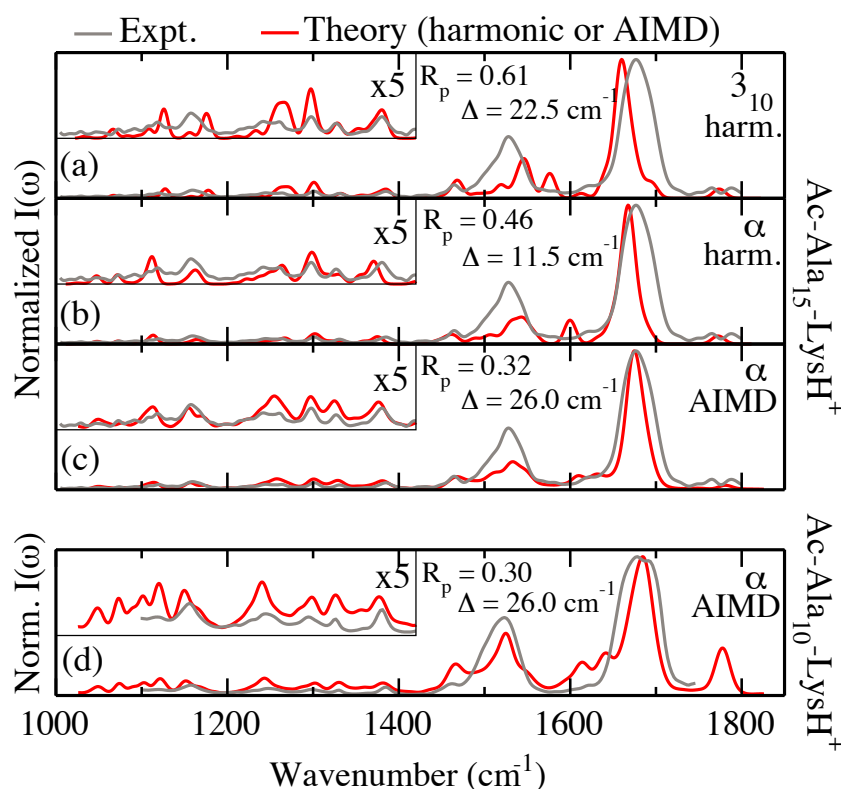calculated spectra are included in each graph (calculated spectra are shifted by $\Delta$ for visual comparison). All spectra have been normalized to 1 for the highest peak. This Figure is adapted from [209] (Figure 4 therein) and further details are explained there.

## 4.5  Conclusions

N-terminal protected and unprotected peptides containing alanine chains (protonated or an attached alkali metal ion) with a varying number of alanine residues are spectroscopically investigated. The amide-I band position for those peptides is found at 1662 cm$^{-1}$ and the amide-II band position at 1534 cm$^{-1}$. The amide-I and amide-II band position for Ala$_5$H$^+$ is located at 1704 cm$^{-1}$ and 1516 cm$^{-1}$, consistent with a more globular conformation. For the 20 alanine residue long peptides, where an alkali metal ion is bound, the amide-I band position is located at 1680 cm$^{-1}$ and the amide-II band position at 1510 cm$^{-1}$. This indicates a stronger hydrogen bonding network and hence a more globular structure for these molecules. The apparently identical infrared spectra for Ac−Ala$_{20}$H$^+$ and Ala$_{20}$H$^+$ possibly indicate that the protonation site is similar for both peptides and not necessarily directly located at the N-terminus (this has a destabilizing effect to an $\alpha$-helical motif). It would be interesting to record infrared spectra of shorter alanine peptides (two or three residues) in the FT ICR apparatus and compare the obtained results with published data (DFT calculations and infrared

spectra).

The band positions of the amide-I and amide-II band in the infrared spectrum of $Ac-Ala_5-LysH^+$ indicate an increased hydrogen bonding strength that is consistent with other measurements where a globular conformation is suggested. The amide-I band position is observed at 1691 cm$^{-1}$ and the amide-II band position is observed at 1518 cm$^{-1}$. With increasing number of alanine residues, the peptide adapts a more $\alpha$-helical structure and the amide-I band for that motif is assigned to be located at around 1667 cm$^{-1}$ with the amide-II band position being at around 1529 cm$^{-1}$ (see also the results for $Ala_{20}H^+$ and $Ac-Ala_{20}H^+$).

Comparing the infrared spectra for $Ac-Ala_{19}-LysH^+$ and $Ac-LysH^+-Ala_{19}$ (the latter might actually also contain the dimer $[Ac-LysH^+-Ala_{19}]_2^{2+}$), no differences are observed. Though, it is possible that the dimer is observed in the measurements instead of the monomer, which would mean that a helical conformation could be present for both molecular systems. Thus, the infrared spectrum is expected to correspond to an $\alpha$-helical motif of the molecules, as observed by others. The role of the dimerization was underestimated in the experiment presented here and to study this in particular more infrared spectra of those systems need to be recorded under better controlled conditions. Especially of interest is an infrared spectrum of the monomer of the molecule with lysine at the N-terminus. Another possibility is to use a non-capped peptide, without the acetyl at the N-terminus, and record infrared spectra of the doubly charged peptide (it seems implausible that the proton is located at the N-terminus and not at the lysine residue for the singly charged ion). The charge at the N-terminus should destabilize the helical motif. A third method might be to synthesize a peptide, where both the N-terminus and the C-terminus are protected. A protection at the C-terminus might hinder the 'head-to-toe' arrangement of the antiparallel helix. Furthermore, using crown ether molecules to cap the protonated lysine residue, thereby sterically hindering an interaction of the charge with the backbone, could also lead to interesting structures of the peptide. Of course, in such an approach it would be essential to record infrared spectra with the crown ether capped lysine at the C-terminus. As a next step, experiments to obtain infrared spectra are planned for shorter polyalanine based peptides with a lysine residue (molecules of interest are for example $Ac-Ala_{10}-LysH^+$ and $Ac-LysH^+-Ala_{10}$) because a transition between helical dimer aggregates of those peptides occurs when more than 13 alanine residues are present. Comparing those results with the results shown here could lead to assignment of the amide-I and amide-II band positions and vibrations in the lower frequency range to globular conformations of the molecule.

To record infrared spectra in the fingerprint region of alanine based peptides, most attention is given to the region of the amide-I and amide-II band. The lower frequency region between 1000 cm$^{-1}$ and 1400 cm$^{-1}$ was measured but it was assumed that the present vibrational bands do not shed much light on the conformation of the investigated molecule and theoretical methods do not describe those vibrations well. However, as shown in Section 4.4.1, theoretical approaches work very well nowadays in this wavelength region. Differences between different conformers are represented in the

calculated spectra and indicative also in this range. Therefore, when new mid-infrared spectroscopic experiments are performed on biomolecules in the gas-phase, it is critical to not only measure the amide-I and amide-II region intensively but also to include the higher wavelength/lower frequency range.

# Chapter 5

# Amide-I and -II vibrations of the cyclic $\beta$-sheet model peptide gramicidin S in the gas phase[*]

In the condensed phase, the peptide gramicidin S is often considered as a model system for a $\beta$-sheet structure. Here, we investigate gramicidin S free of any influences of the environment by measuring the mid-IR spectra of doubly protonated (deuterated) gramicidin S in the gas phase. In the amide-I (i.e., C=O stretch) region, the spectra show a broad split peak between 1580 cm$^{-1}$ and 1720 cm$^{-1}$. In order to deduce structural information, the conformational space has been searched using molecular dynamics methods and several structural candidates have been further investigated at the density functional level. The calculations show the importance of the interactions of the charged side chains with the backbone, which is responsible for the lower frequency part of the amide-I peak. When this interaction is inhibited via complexation with two 18-crown-6 molecules, the amide-I peak narrows and shows two maxima at 1653 cm$^{-1}$ and 1680 cm$^{-1}$. A comparison to calculations shows that for this complexed ion, four C=O groups are in an antiparallel $\beta$-sheet arrangement. Surprisingly, an analysis of the calculated spectra shows that these $\beta$-sheet C=O groups give rise to the vibrations near 1680 cm$^{-1}$. This is in sharp contrast to expectations based on values for the condensed phase, where resonances of $\beta$-sheet sections are thought to occur near 1630 cm$^{-1}$. The difference between those values might be caused by interactions with the environment, as the condensed phase value is mostly deduced for $\beta$-sheet sections that are embedded in larger proteins, that interact strongly with solvent or that are part of partially aggregated species.

## 5.1   Introduction

The function of biological molecules is determined by their three-dimensional structure and shape. For peptides and proteins, these three-dimensional structures are the consequence of the primary sequences of amino acids, the resulting intrinsic intramolecular

interactions as well as the interactions of the molecules with their environments. The main interactions include bond stretching, angle bending and twisting terms, as well as intra- and intermolecular repulsive, dispersive and electrostatic interactions. The subtle balance between all those intra- and intermolecular interactions gives rise to extremely complex potential energy surfaces in the folding of biological molecules, such as proteins. Traditionally, studies on biological molecules are performed in the condensed phase, where information directly relevant to the structure of the molecule under physiological conditions is obtained. However, in order to get an in-depth understanding of the individual contributions to the potential energy landscape, the structures of peptides and proteins can be investigated in the gas phase and therefore in the absence of solvent effects.

Initial experiments on biological molecules in the gas phase were performed on neutral species in molecular beams [46]. The molecules can be brought into the gas phase by either thermal evaporation or laser desorption. The former technique is limited to small, thermally stable species while laser desorptions also allows for the investigation of larger species such as neutral gramicidin peptides [47]. Over the last years, several techniques have been developed to allow for studies of large, charged biomolecules in the gas phase. However, obtaining structural (or dynamic) information on such molecules is difficult. For condensed phase samples, numerous spectroscopic and scattering techniques exist. For gas-phase species choices are more limited. Two techniques that deliver direct structural information are gas-phase ion mobility measurements [158, 214, 215] and infrared spectroscopy [43, 146, 216]. These two techniques are rather complementary, as ion mobility probes the overall shape and is not very sensitive to the local structure whereas IR spectroscopy probes the bonds between the atoms and is not as sensitive to the overall structure.

For small gas-phase biomolecules, mid-IR spectroscopy has in many instances shown to yield detailed structural information [49, 55, 217–221]. This is especially true when the molecules are internally cold and when individual oscillators give rise to individual peaks. A comparison with theoretical predictions can then give detailed structural information. For larger systems, the situation is more complicated as resonances from individual oscillators overlap and give rise to a broad envelope [146]. When the samples are warm, several mechanisms can cause peak broadening and the possible presence of several conformers can result in overlapping spectra of different species. Further, theory is becoming increasingly difficult when the size of the system increases. For condensed phase proteins, IR spectroscopy is a standard technique to determine secondary structures and even small changes in peak shapes and position can be attributed to structural changes [42]. In part, this is possible because an extended database of condensed-phase IR spectra has been recorded for proteins with known secondary structures. This has allowed a structure-spectrum correlation, where diagnostic band positions, notably the amide-I band (C=O stretching vibration), are used to make a structural identification. In the condensed phase, $\alpha$-helices as well as disordered structures have amide-I band positions near 1654 cm$^{-1}$, while $\beta$-sheet type structures have band positions near 1633 cm$^{-1}$ [42]. However, the environment can have a strong influence on those values. For example, in a study on a $\beta$-hairpin peptide, for which it was shown by nuclear magnetic resonance (NMR) that its structure is folded, $\beta$-sheet-like, the amide-I modes are found centered around 1640 cm$^{-1}$ and 1670 cm$^{-1}$, at identical positions to those of a reference

peptide having a random coil structure [222]. Only at high concentrations, when the peptides start to aggregate, an additional peak near 1616 cm$^{-1}$ appears. IR studies on the cyclic peptide gramicidin S highlight the important role of the environment as well [223]. In polar protic solvents, the main amide-I band is found near 1629 cm$^{-1}$, however it is observed to shift when the protic character of the solvent is reduced and for dimethyl sulfoxide (DMSO), an amide-I position of 1652 cm$^{-1}$ is observed while the secondary structure of the peptide is assumed not to change [223]. The spectral signatures of $\beta$-sheet peptides can thus be highly dependent on the environment.

For peptides and proteins in the gas phase, no database which correlates structure to spectra is available. It is thus important to investigate in the gas phase medium to large size model systems with well-defined structures to establish the corresponding mid-infrared signatures and compare them to the spectral signatures of large proteins [146], small neutral model systems for $\beta$-sheet structures [49, 217–219, 221], helices [204], turns [220], *coiled coil* systems [224] and to data from the condensed phase. While for small neutral species, model systems for $\beta$-sheet structures were investigated [49, 217–219, 221, 225], the search for the presence of such structures in charged species in the gas phase remained elusive so far [204, 225]. However, understanding the spectral signatures of charged, $\beta$-sheet rich systems in the gas phase is of tremendous importance since it might pave the way for a variety of future experiments where IR spectrometry is combined with mass spectrometric techniques. In contrast to most condensed-phase techniques, MS-based approaches are capable of analyzing specific species within an ensemble of enormous inherent heterogeneity. The IR-MS combination, therefore, exhibits an outstanding potential to obtain information about the secondary structure of the toxic, but still poorly characterized oligomers that precede fibril formation of proteins involved in a variety of amyloid diseases [226].

Due to internal constraints, the cyclic peptide gramicidin S (GramS) could serve as a model for a $\beta$-sheet structure even for charged species. Gramicidin S was discovered in the 1940's in the former Soviet Union as a substance with antibiotic character that is produced by *Bacillus brevis* [227]. Subsequent characterization revealed its amino acid sequence [228]. Its structure is proposed to consist of an antiparallel $\beta$-sheet, stabilized by four intramolecular hydrogen bonds between two opposing leucine and valine residues and two $\beta$-turns consisting of two D-phenylalanine and proline residues [229]. In addition, gramicidin S contains two ornithine groups that have basic side chains (see Figure 5.1).

Gramicidin S and derivatives have been extensively studied in the condensed phase. As gramicidin S is difficult to crystallize, comparatively few X-ray crystallography studies exist. X-ray structural characterization has been performed on the hydrated gramicidin S – urea complex [230] as well as on a derivative in which the Orn and Phe peptidic nitrogen atoms are methylated and in which the side chain of Orn contains a protecting group [231]. In these studies, the proposed antiparallel $\beta$-sheet structure is found with additional interactions of the Orn side chains with the D-Phe C=O groups. In case of the urea complex, only one side chain is found to be involved in this interaction [230], while in the other study both side chains are found to interact [231]. There are two possibilities how this interaction can take place. In the hydrated gramicidin S – urea complex [230], the side chain is found to form an H-bond to i $\longrightarrow$ (i+2) Phe C=O (the nearest Phe clockwise in Figure 5.1) and in the study on the derivative, this interaction is found to
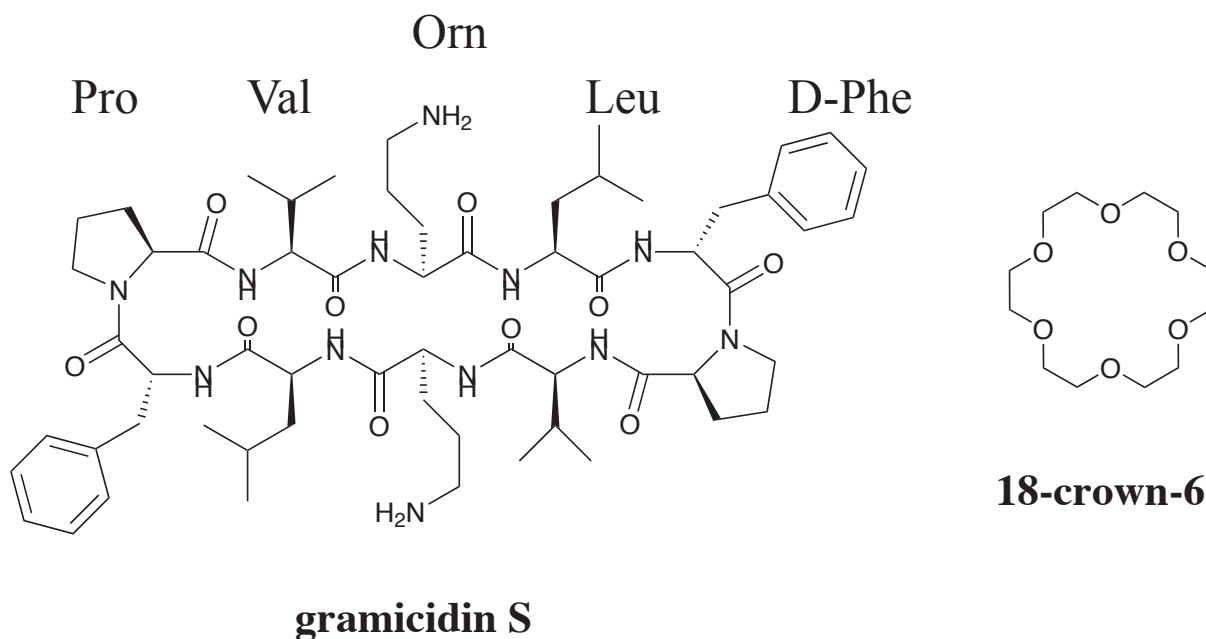
Orn

Pro          Val          NH$_2$          Leu          D-Phe

**gramicidin S**

**18-crown-6**

**Figure 5.1:** The cyclic decapeptide gramicidin S (cyclo(Pro-Val-Orn-Leu-D-Phe)$_2$) and the crown-ether 18-crown-6.

occur to the i $\longrightarrow$ i-3 Phe residue (counterclockwise in Figure 5.1). Many NMR studies on gramicidin S and its derivatives have been carried out. The backbone is generally found to adopt the antiparallel pleated $\beta$-sheet conformation [232–235]. The Orn side chains are found to adopt the i $\longrightarrow$ i+2 configuration [232, 234], however in one study, a combination of both orientations [233] is found. Gramicidin S is also investigated in the gas phase as a charged molecule using mass spectrometric methods [236–238], and from ion mobility studies, it is concluded that a $\beta$-sheet structure is present [238]. Spectroscopic studies on neutral gas-phase gramicidin S in the X−H stretching region give a N−H stretching signature that is compatible with a $\beta$-sheet structure as well [47]. Here, we present a mid-IR study of doubly protonated gramicidin S in the gas phase.

## 5.2   Experimental

The experiments were performed using the Fourier-transform ion cyclotron resonance (FT ICR) mass spectrometer at the free electron laser facility FELIX in Nieuwegein, the Netherlands [57], see Section 1.4.4. Gramicidin S was synthesized in the group of A. S. Ulrich [239]. It was dissolved at a 50 $\mu$M concentration in a H$_2$O/methanol/AcOH solution of 60/39/1% . In some experiments, 18-crown-6 ether was added at a concentration of 200 $\mu$M. The solution was sprayed using a Z-Spray electrospray source (Micromass, Manchester, UK) and the ions were transferred to the ICR trapping cell. Deuteration was performed by dissolving gramicidin S in deuterated solvents. The observed mass increases by 14 amu, indicating that all exchangeable hydrogen atoms were replaced by deuterium. After isolation of the mass-to-charge ratio of interest, the ions were irradiated for a few seconds by the IR output of the free electron laser FELIX. When the IR frequency of the light is resonant with an IR active vibration mode of

the ions, the sequential absorption of many photons and subsequent fragmentation (infrared multiple photon dissociation (IR-MPD), see Section 1.4.1) occurs. IR spectra can be recorded by monitoring the fragmentation yield as a function of IR wavelength. Although the so obtained spectra are not identical to linear IR absorption spectra, it was shown that they can be very close to them [58], see Section 1.4.2.

As IR light source, the free electron laser FELIX [65] was used, see Section 1.4.3. It is continuously tunable over the 5–250 $\mu$m range. The light output comes in macropulses of about 5 $\mu$s length at a repetition frequency of up to 10 Hz (5 Hz is used in the present experiment). The macropulses contain micropulses which can be adjusted in length between 300 fs and several ps. The bandwidth is transform limited and can range from 0.5% FWHM of the central wavelength to several percent. The micropulse repetition rate can be selected to be either 25 MHz or 1 GHz, resulting in a micropulse spacing of 40 ns or 1 ns, respectively. In the 1 GHz mode, the output energy can be up to 100 mJ / macropulse.

## 5.3 Results

**Measured IR spectra**

When the IR radiation of FELIX is resonant with an IR active vibrational mode of the molecule, IR-MPD can be observed. For the here investigated doubly protonated gramicidin S, fragmentation is observed to be distributed over a multitude of different fragmentation channels and the corresponding total ion intensity in each of those channels individually can be quite low. As a consequence, instead of monitoring the appearance of fragment ions, the depletion of parent ions is monitored.

In the top part of Figure 5.2, the gas-phase IR spectra of $[GramS + 2\,H]^{2+}$ as well as of perdeuterated $[d-GramS + 2\,D]^{2+}$ are shown. The spectrum of $[GramS + 2\,H]^{2+}$ shows a broad split peak between 1580 cm$^{-1}$ and 1720 cm$^{-1}$, another intense peak around 1500 cm$^{-1}$ and some less intense structure between 1300 and 1410 cm$^{-1}$. The split peak is in the region where one expects the amide-I (C=O stretch) vibration to occur while the peak at 1500 cm$^{-1}$ is where the amide-II (N−H bending) vibration is expected to be present. The IR spectrum of $[d-GramS + 2\,D]^{2+}$ between 1580 cm$^{-1}$ and 1720 cm$^{-1}$ shows a structure consisting of at least three underlying peaks. The peak observed for $[GramS + 2\,H]^{2+}$ near 1500 cm$^{-1}$ appears to be changed in shape and shifted to ~1410 cm$^{-1}$.

Assuming pure amide-I and -II vibrations to be present in the molecule, one would expect essentially no change in shape for the amide-I band upon deuteration. A small shift of this band of 1-5 cm$^{-1}$ can be due to coupling of the C=O stretch motion with the N−H (N−D) in plane bending motion. The amide-II band on the other hand consists to a large extent of N-H bending motion and in solution, it can be found near 1550 cm$^{-1}$. Upon deuteration, the mode shifts to 1460-1490 cm$^{-1}$ and is then termed amide-II'. It also changes its character and consists to a large extent of C−N stretching motion in deuterated molecules.

When comparing the two spectra in the top part of Figure 5.2, the expected isotopic shift in the amide-II band can be clearly observed. At first glance surprising, however, might be the change in peak shape in the amide-I region, as it implies that this structure
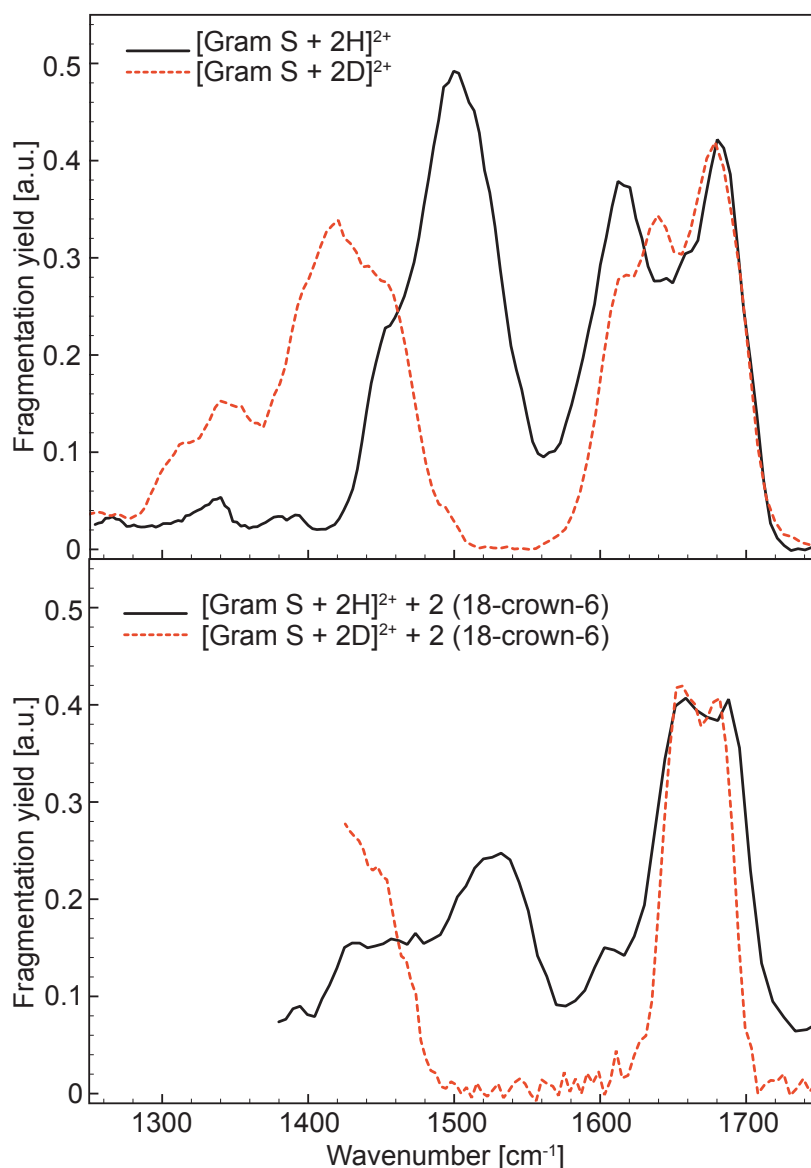
**Figure 5.2:** Top: Experimental IR spectra of $[\text{GramS} + 2\,\text{H}]^{2+}$ and perdeuterated $[\text{d}-\text{GramS} + 2\,\text{D}]^{2+}$. Bottom: Experimental IR spectrum of $[\text{GramS} + 2\,\text{H}]^{2+}$ complexed with two 18-crown-6 molecules and perdeuterated $[\text{d}-\text{GramS} + 2\,\text{D}]^{2+}$, complexed with two 18-crown-6 molecules. Amide-I (C=O stretch) vibrations are expected to be found above $1600$ cm$^{-1}$.

is not solely due to C=O stretching motion. On the other hand, looking at the structure of the molecule and considering the fact that we are looking at the doubly protonated cation, it is also clear that a simple picture of pure amide-I and -II vibration does not hold. The two protons (deuterons) will attach to the most basic sites in the molecule, the two ornithine side chains (see Figure 5.1), and the $-(\text{CH}_2)^3-\text{NH}_3^+$ side chains will then coordinate to backbone C=O groups. The resulting strong hydrogen bonds can then significantly shift the frequency of those C=O oscillators. In addition, the $-\text{NH}_3^+$ groups by themselves are IR active and give for lysine in solution a $\delta_{as}$ mode around $1627$ cm$^{-1}$ and a $\delta_s$ mode around $1526$ cm$^{-1}$ [42]. Upon deuteration, those modes then shift to 1201

and 1170 cm$^{-1}$, respectively [42]. This is in line with the observation that the red side of the structure in the amide-I region in the spectrum of [GramS + 2 H]$^{2+}$, where the $\delta_{as}$ mode of $-$NH$_3^+$ is expected, undergoes the largest change upon deuteration.

For a further investigation of the effect of the charged side chain and, in particular, to study the peptidic backbone with reduced backbone $-$side chain interaction, experiments were performed in which [GramS + 2 H]$^{2+}$ is complexed with two 18-crown-6 molecules. These crown-ether molecules have a very high affinity to protonated amines [240] and it is expected that in such a complex, the side chain $-$NH$_3^+$ groups are coordinated to 18-crown-6 molecules and that the interaction of the $-$NH$_3^+$ groups to the backbone is inhibited due to steric constraints.

The bottom part of Figure 5.2 shows the IR spectra of [GramS + 2 H]$^{2+}$ and of deuterated [d$-$GramS + 2 D]$^{2+}$, both ions being complexed with two 18-crown-6 molecules. When comparing those spectra with the spectra of their uncomplexed counterparts, it can be observed that the spectra of complexed species show a significantly altered spectral structure in both, the amide-I and -II region. Slightly narrower lines can be expected, as fewer absorbed photons are required to fragment this complex. Most striking in the amide-I region is that the peak near 1600 cm$^{-1}$ has a much lower intensity (and is even completely absent for the deuterated complex). In that region, contributions from C=O oscillators involved in strong hydrogen bonds as well as from $-$NH$_3^+$ $\delta_{as}$ modes would be expected. The crown-ether probably shields the $-$NH$_3^+$ groups from interacting with and forming strong hydrogen bonds with backbone C=O groups. Thus we can assign a large fraction of the intensity near 1600 cm$^{-1}$ in the spectra of the uncomplexed ions to C=O groups involved in such an interaction. Additional deuteration also shifts the $-$NH$_3^+$ $\delta_{as}$ mode outside the experimental range. This is in line with the reduction in intensity near 1600 cm$^{-1}$ in the spectrum of [GramS + 2 H]$^{2+}$ and the total absence of intensity in that range in the spectrum of deuterated [d$-$GramS + 2 D]$^{2+}$ complexed with 18-crown-6. In the amide-II region around 1500 cm$^{-1}$, the situation is more complex. There, deuteration changes the nature of the modes from predominantly N$-$H bending to N$-$C stretching upon deuteration. Concomitant with that is a change in relative intensity, compared to the amide-I mode. In addition, the $-$NH$_3^+$ $\delta_s$ mode is shifted to lower wavenumber. Also, the complexation with 18-crown-6 might affect both the positions and intensities of the $-$NH$_3^+$ $\delta$ modes. All those factors make the interpretation of the amide-II region difficult.

## 5.4   Calculations

In order to learn more about the structure of gas-phase [GramS + 2 H]$^{2+}$, different conformeric structures and their IR spectra were calculated. There are two fundamental problems that need to be addressed. First, the conformational space is very large and it is difficult to select possible candidate structures for further investigations. Second, even when reasonable candidate structures are found, their structures need to be optimized, and the energies as well as the IR spectra need to be calculated. For the calculations of IR spectra of small peptides containing only a few amino acids, density functional theory (DFT) methods have been shown to give at least reasonable geometries and IR spectra [220]. For species of the size of [GramS + 2 H]$^{2+}$, DFT methods are, however,

computationally expensive.

We choose a molecular dynamics (MD) based search scheme described below to find candidate structures which are then further investigated at the B3LYP level with the def-SVP basis set using Turbomole [241] with the ri-approximation. The MD calculations are performed with the Amber 9 [242] program using the ff96 force field [243]. The parameters for the non-standard amino acid ornithine are adapted from parameters of lysine by removing one $CH_2$ group and distributing its net charge to the neighboring atoms. In order to efficiently sample the conformational space, replica exchange molecular dynamics (REMD) [244] has been performed on 16 parallel runs with temperatures ranging in a geometric series from 270 K to 800 K. Time steps are set to one femtosecond, replica exchange attempts occur every picosecond and the total length of each simulation is 100 nanoseconds.

At each simulation, 1000 structure snapshots from the trajectory at 312 K are picked and further analyzed. First, the structures are energy minimized at the force field level. Next, they are grouped in families of structural similarity. This is done by using the root mean square atom displacement (RMSD) values between two structures to judge the similarity between them. In the calculations of the RMSD values, all atoms of the peptidic backbone plus the N−H hydrogen as well as the $-NH_3^+$ nitrogen on the ornithine side chains are used. The RMSD values are then calculated between all pairs of structures. Those values are plotted in a two-dimensional color coded matrix and this matrix is sorted by grouping species with low RMSD values such that blocks with similar structures appear. From each block, several low energy structures are taken, their energies minimized at the B3LYP level and the IR spectra calculated using numerical differentiation.

From the calculations on $[GramS + 2H]^{2+}$, visual inspection shows that for most structures, the dominant interaction occurs between the $-NH_3^+$ group of the Orn side chain and the C=O groups of the Phe, Orn, Val and Pro amino acids. In general, there are two possibilities for such interactions. The Orn side chains can either coordinate to the backbone C=O groups that are found clockwise or counterclockwise (see Figure 5.1), *i.e.* the side chain on top orients to the right and the one below to the left (clockwise) or *vice versa*. For an interchange between those two orientations, both side chains have to act concertedly and at least four, presumably strong, hydrogen bonds have to be broken. Consequently, the barrier for such a transition is expected to be high. For that reason, two separate simulations were started, one with the Orn side chains oriented clockwise and one with counterclockwise orientation.

From the resulting 2000 structures, 21 were selected for further investigations using B3LYP. Ten structures had a counter-clockwise and eleven a clockwise Orn side chain orientation. Interestingly, the structure with the lowest B3LYP energy is also the one which has the lowest ff96 energy. The relative ff96 and B3LYP energies can be compared, by setting the lowest energy structure as the zero energy in both sets. For the ten counter-clockwise structures, the RMS difference in relative energy between ff96 and B3LYP is 11 kcal/mol, for the eleven clockwise structures 7 kcal/mol. An inspection of those 21 structures resulted in five low energy families and the lowest energy structure of each of those families together with the calculated IR spectra is shown in Figure 5.3. Note that the structures in Figure 5.3 are turned upside down, compared to the structure in Figure 5.1. The main structural differences result from differences in hydrogen

bonding of the Orn side chains with the backbone. Spectra for hydrogenated and deuterated species are shown on the left-hand and right-hand side, respectively. The frequency positions in all calculated spectra as well as in the further discussion are scaled by 0.9614 [245].

The top two structures are characterized by a clockwise (see Figure 5.1) orientation of the Orn side chains. In case of the top structure (A), the Orn side chains coordinate to the C=O groups from their own residue as well as Phe C=O groups two residues away. The structure is very symmetric and distances between the H-atoms on the $-NH_3^+$ groups and the O-atoms of the corresponding C=O groups are 1.68 Å and 1.73 Å to Phe and Orn, respectively. Four $\beta$-sheet type interactions are present and H-bonding distances are 2.06 Å and 1.95 Å for the outer two and inner two H-bonds, respectively. In addition, the Phe aromatic rings are closely coordinated to the charged side chains. At both the ff96 and the B3LYP level, this structure is found to be lowest in energy. The structure shown in the second row (B) is also based on a clockwise orientation of the Orn side chains, this time however interacting with the C=O groups of the Phe and Pro residues (distances of 1.63 Å and 1.84 Å, respectively). Four $\beta$-sheet type interactions are present as well with 2.31 Å and 2.01 Å as distances for the outer and inner ones, respectively. At the B3LYP level, this structure is 7.2 kcal/mol higher in energy than structure A.

The three structures shown in the lower part of Figure 5.3 (C-E) all have counter-clockwise orientations of the Orn side chains. In structure C), the Orn side chains are coordinated to C=O groups of the Phe, Orn (opposite side) and Val (next to the Orn group) residues. The distances between the H-atoms on the $-NH_3^+$ groups and the O-atoms of the corresponding C=O groups range from 1.62 Å to 1.87 Å. The only other $-H\cdot\cdot O=$ distances closer than 2.5 Å are found between the two Leu C=O and corresponding Val N−H groups. Such interactions are expected for a $\beta$-sheet type structure. The distances between the corresponding atoms are with 2.18 Å and 2.22 Å relatively large, however. In relative energy, this structure is 1.4 kcal/mol higher in energy than A). Structure D) on the second row from below is similar to structure C). There as well, one Orn side chain coordinates to Phe, Orn and Val residues. The other Orn side chain coordinates to a Phe and an Orn residue as well as to the aromatic ring of the Phe residue. The H-bond lengths are similar to the ones found in C); H-bonds in which the Orn side chains are involved range from 1.60 Å to 1.87 Å and for the two $\beta$-sheet type interactions, H-bond lengths of 2.22 Å and 2.29 Å can be found. In addition, a weak $C_7$ type interaction with a distance of 2.37 Å can be noted between a Pro C=O and a Orn N−H. The relative energy of D) is +6.2 kcal/mol.

The structure shown at the bottom (E) is very symmetric and characterized by inter-actions of the two Orn side chains with the Phe and Pro residues. The distance to the Phe C=O is 1.67 Å and to the Pro C=O 1.83 Å. In this structure, four $\beta$-sheet type interactions are present. The distances for the outer H-bonds is 2.04 Å and for the inner H-bonds 1.84 Å. The Orn C=O groups have their closest H-bonding interaction with the Orn N−H in a $C_5$ type interaction. The distance is, however, with 2.96 Å large and the corresponding interaction, thus, weak. The relative energy of E) is with +15.7 kcal/mol the highest of the five structures considered.

When comparing the calculated IR spectra to the experimental ones, many of the general features shown in the experiment are reproduced. A perfect agreement, however, is

not observed for any of the calculated structures. In order to characterize the observed spectrum, it is instructive to analyze the normal modes in the calculated spectrum. For the following, we would like to discuss the normal modes in the amide-I region of the deuterated structures A) and C). For deuterated A) in the range between 1600 cm$^{-1}$ and 1720 cm$^{-1}$, the calculated spectrum in Figure 5.3 shows four groups of peaks. At the lowest frequency two almost degenerate vibrations at 1620 cm$^{-1}$ resulting from Phe C=O stretch motion occur. Next higher at 1643 cm$^{-1}$ and 1645 cm$^{-1}$, C=O stretching vibrations of the two Orn C=O groups are present. The four vibrations of the $\beta$-sheet Leu and Val C=O groups are found between 1681 cm$^{-1}$ and 1688 cm$^{-1}$. At the highest frequency at 1705 cm$^{-1}$ and 1706 cm$^{-1}$, vibrations of the Pro C=O groups are found. Qualitatively, those shifts can be understood by having the C=O groups with the strongest H-bonds (bond to the charged Orn side chain) most to the red, the medium strongly H-bonded $\beta$-sheet groups more to the blue and the essentially non H-bond Pro vibrations most to the blue.

For the deuterated structure shown in Figure 5.3 C), ten vibrational modes, all stemming from C=O stretching vibrations, are found in the region between 1600 cm$^{-1}$ and 1700 cm$^{-1}$. Most to the red, at 1624 cm$^{-1}$ and 1628 cm$^{-1}$, vibrations of the two Phe C=O groups occur. The Val C=O vibrations are found at 1655 cm$^{-1}$ and 1657 cm$^{-1}$, followed by the Orn C=O vibrations at 1660 cm$^{-1}$ and 1663 cm$^{-1}$. Between 1687 cm$^{-1}$ and 1695 cm$^{-1}$, the Leu and Pro C=O vibrations are found which appear rather mixed. As in the case of the vibrational modes of structure A), the red-shift qualitatively follows the strength of H-bonding. Most to the red, modes of groups are found that (strongly) interact with the charged Orn side chain, while the vibrations of the more weakly H-bonded Leu and Pro C=O groups are found on the blue side of the peak.

Structures and spectra for [GramS + 2 H]$^{2+}$ complexed with two 18-crown-6 molecules were calculated using the same methodology that was employed for uncomplexed [GramS + 2 H]$^{2+}$. At the ff96 level, four structural families are found and a total of nine structures were optimized and their IR spectra calculated at the B3LYP level. In all structures that were found, the 18-crown-6 molecules are tightly coordinated to the Orn side chains. In three of the four families, in addition some coordination of C=O groups towards the Orn $-$NH$_3^+$ groups is observed. This, however, comes at the expense of intramolecular H-bonding and at the B3LYP level, the lowest energy structure from those families is 12.8 kcal/mol higher in energy than the overall lowest energy structure. Figure 5.4 shows this overall lowest energy structure, together with the calculated IR spectra for protonated and deuterated species. The structure is characterized by Orn side chains coordinating 18-crown-6 molecules and pointing away from the peptide backbone. The Leu and Val C=O groups are found to be involved in $\beta$-sheet type H-bonding with distances for the outer Leu C=O H-bond of 2.00 Å and for the inner Val C=O H-bond of 2.37 Å. The latter C=O groups, however, have with 2.20 Å shorter distances to the neighboring Leu N-H groups to form $C_7$ type H-bonds. The Pro and Orn C=O groups form $C_7$ type H-bonds with the adjacent Orn and Phe N$-$H groups at distances of 2.09 Å and 2.00 Å, respectively. Only the Phe C=O groups have no close H-bonding partner.

Qualitatively, the calculated spectra are in good agreement with the ones obtained experimentally. For the non-deuterated molecule, both the peaks in the amide-I and -II region are predicted narrower than observed. For the deuterated species, the
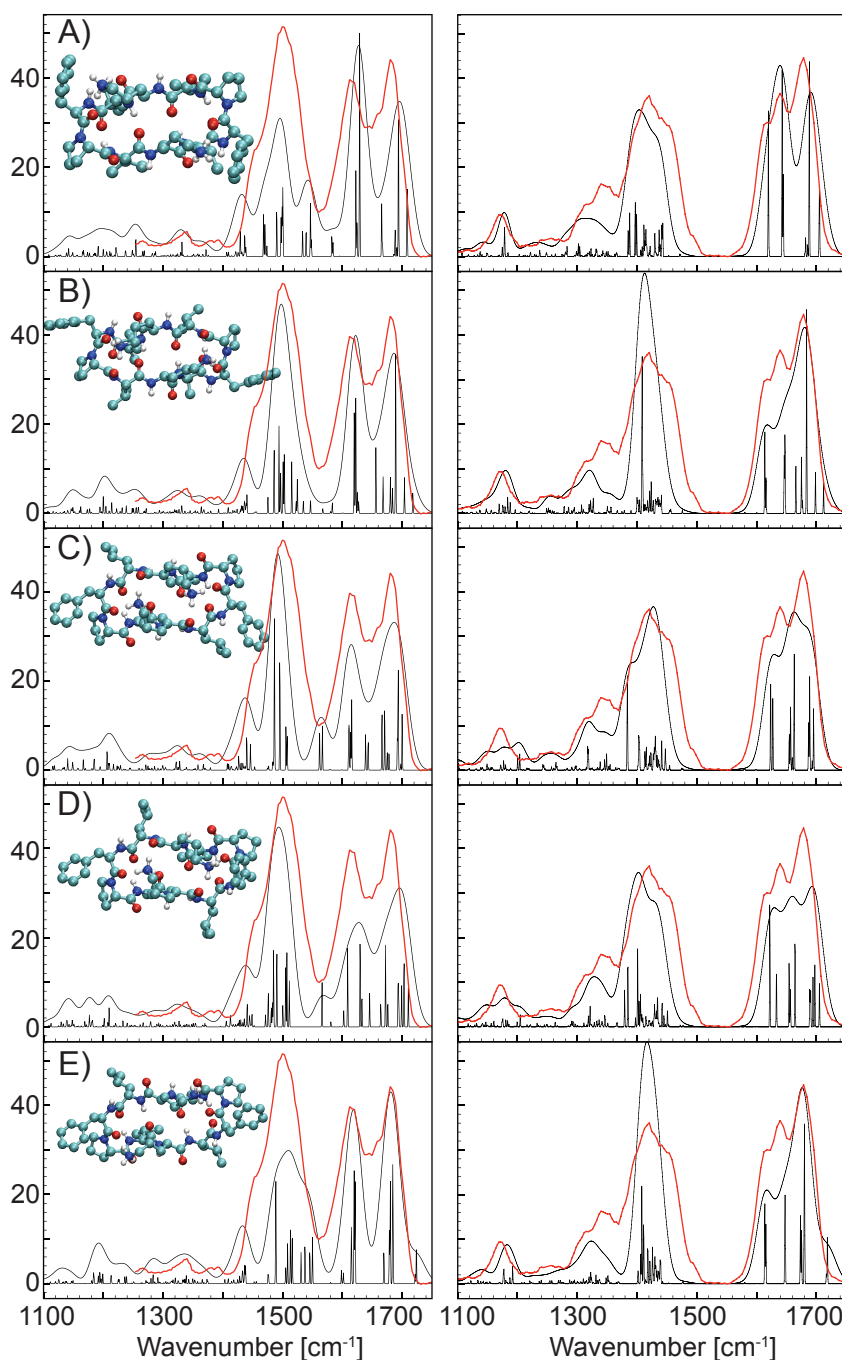
**Figure 5.3:** Experimental IR spectra of $[\text{GramS} + 2\,\text{H}]^{2+}$ and perdeuterated $[\text{d}-\text{GramS} + 2\,\text{D}]^{2+}$ (red solid line, left and right part respectively) compared to theoretical predictions for five calculated structures. The sharp spectra result from a convolution of the theoretical stick spectrum with a $1\,\text{cm}^{-1}$ full width gaussian and the broader spectra in black from a convolution with a 2% full width gaussian. All calculated line positions are scaled by 0.9614.

width of the amide-I peak is well reproduced. Unfortunately, for this molecule, the measurements did not cover the amide-II region. For the amide-I peak, the splitting which is experimentally observed for the protonated as well as the deuterated species

is not reproduced by the calculations.

Visualizing the vibrations in the amide-I region for the deuterated species shows that most of the Orn C=O vibrations are present most to the red, at 1643 cm$^{-1}$ and 1644 cm$^{-1}$. They are, however, of low IR intensity. Pro C=O vibrations are predicted to be much stronger and are found at 1649 cm$^{-1}$ and 1650 cm$^{-1}$. Significantly shifted to the blue, at 1675 cm$^{-1}$ and 1676 cm$^{-1}$, Leu "$\beta$-sheet" vibrations occur. They are to some extent mixed with the vibrations of the Phe C=O groups which are found at 1677 cm$^{-1}$ and 1678 cm$^{-1}$. Most to the blue are the out-of-phase and in-phase combinations of the Val C=O groups at 1682 cm$^{-1}$ and 1694 cm$^{-1}$. Such a splitting might be indicative of different environments. Then, however, localized modes would be expected. In the present case, the two C=O groups are in very similar environments and as they are quite close to each other, they couple and give rise the the mentioned in- and out-of-phase vibrations. In conclusion for [GramS + 2 H]$^{2+}$ complexed with two 18-crown-6 molecules, it can be noted that, contrary to expectations, the red part of the amide-I peak results from Orn and Pro C=O vibrations while the Leu and Val C=O vibrations are rather found near the blue at around 1680 cm$^{-1}$
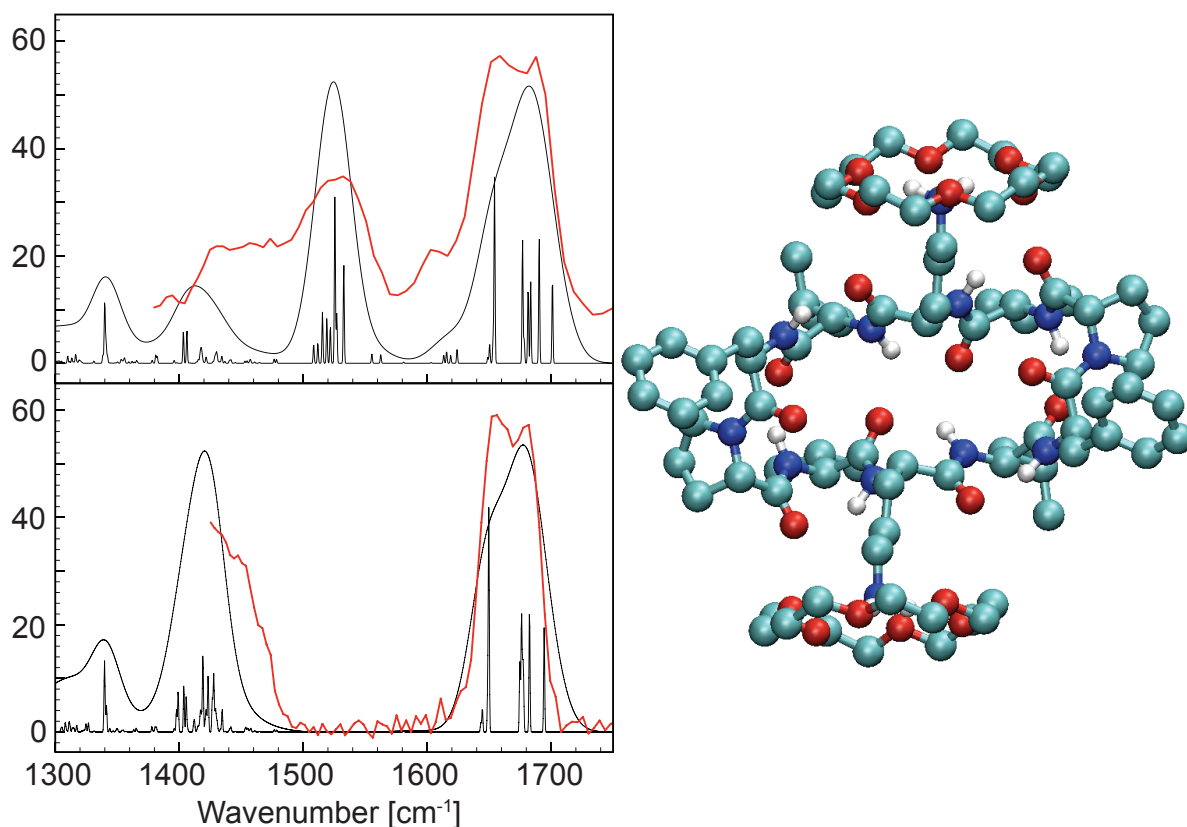


**Figure 5.4:** Experimental IR spectra of [GramS + 2 H]$^{2+}$ and perdeuterated [d−GramS + 2 D]$^{2+}$ complexed with two 18-crown-6 molecules (red solid line, top and bottom, respectively) compared to a theoretical prediction. The sharp spectra result from a convolution of the theoretical stick spectrum with a 1 cm$^{-1}$ full width gaussian and the broader spectra in black from a convolution with a 2% full width gaussian. All calculated line positions are scaled by 0.9614.

## 5.5   Discussion

The experimental results clearly show the importance of the Orn $-NH_3^+$ side chain interaction on the IR spectra, and, thus, also on the structure of gas-phase gramicidin S. The deuteration experiments show that the red shoulder in the experimental amide-I peak is caused by $-NH_3^+$ umbrella motion and further, that this mode couples to other C=O stretching modes, as the amide-I peak shape changes upon deuteration of the molecule. Further evidence for the $-NH_3^+$ – peptide backbone interactions comes from the 18-crown-6 experiments, which demonstrate that the red side of the amide-I peak is due to oscillators that interact with the $-NH_3^+$ group. The calculations shed further light on the nature of these side chain – backbone interactions. In all calculated structures, a $-NH_3^+-Phe(C=O)$ interaction is found to be strong, irrespective whether the orientation of the side chain occurs clockwise or counterclockwise. In addition, an interaction with the Orn C=O is present in the low energy structures A), C) and D) presented in Figure 5.3. An interaction with Val C=O is predicted for C) and D) and with Pro C=O for B) and E). Due to their strong H-bonded nature, these strongly H-bonded Phe, Val, Pro or Orn C=O groups give rise to the resonances which are found most to the red in the amide-I peak. Leu C=O groups on the other hand do not interact with the Orn $-NH_3^+$ side chains. Instead, they are found to bind to the opposing Val N−H groups in a $\beta$-sheet type arrangement. Their vibrations, however, are predicted in calculations for structures A) and C) in Figure 5.3 and for the crown-ether complexed structure in Figure 5.4 to occur on the blue side of the amide-I peak between 1675 cm$^{-1}$ and 1695 cm$^{-1}$. When the Val C=O is coordinated to an $-NH_3^+$, its resonance is predicted to be found around 1656 cm$^{-1}$ (see Figure 5.3 C)). When this is not the case and the Val C=O groups interacts with an opposing Leu N−H groups in a $\beta$-sheet arrangement (see Figure 5.3 A) and Figure 5.4), its resonances are found to be close to that of Leu C=O "$\beta$-sheet" in gramicidin S between 1682 cm$^{-1}$ and 1694 cm$^{-1}$.

When comparing the experimental spectra with the calculated ones, all calculations give reasonable agreements between theory and experiment. In the experiment, the amide-I band appears to consist of two components for the hydrogenated gramicidin S and of three for the deuterated species. In all calculations a splitting of the amide-I band is predicted as well. There are, however, also differences in the calculated and experimental relative intensities of the amide-I components, and it is difficult to decide which of the structures in Figure 5.3 gives the best match to experiment. Here, spectra are calculated at 0 K. The experiment on the other hand does not measure the linear IR absorption spectrum of a fully annealed sample at 0 K, but is rather measured as an IR-MPD spectrum. For many species, it was shown that IR-MPD spectra can be quite similar to the corresponding linear absorption spectra [58]. The experimental IR-MPD spectrum is linearly corrected for the variation in FEL power over the tuning range. This is however only a first order approximation and errors in relative intensities may result.

In addition to uncertainties in relative intensities, peak positions can also deviate from their positions in a linear, 0 K spectrum. This can be caused by mode anharmonicities, in combination with the IR-MPD mechanism [58] and the fact that the excitation starts from a room temperature sample. However, one can estimate the resulting shifts to be on the order of a few wavenumbers [58], thus relatively small compared to the widths

of the observed resonances.

Another complication stems from the possibility of the presence of more than one conformer. The presence of different conformers of gas-phase $[\text{GramS} + 2\,\text{H}]^{2+}$ is observed in recent low temperature experiments [48]. Ion mobility experiments at 300 K on the other hand did not indicate the presence of different conformers with vastly different overall shapes [238]. We have calculated cross-sections for the structures presented in Figure 5.3. The results show that structure A) has the lowest cross-section (highest mobility) and, for example, structure E) has an about 10% higher cross-section. Such a difference could stem from differences in the orientations of the rather flexible phenyl side chains and, in a 300 K experiment, thermal motion might average out those deviations to some extend. Indeed, performing cross-section calculations on structures where the phenyl groups are deleted reduces the differences in relative cross-sections to about 5%. With the addition of thermal motion of the other side chains, the resulting differences in mobilities might be insufficient to be observable in most ion mobility experiments. In the calculations, many local minima were populated. It is also observed, that in the simulations, transition from a clockwise to counterclockwise (or *vice versa*) side chain orientation occurred only infrequently with a high barrier between those two nearly isoenergetic conformational spaces. It is thus likely that in the experiment, a distribution of different conformers is present and the resulting IR-MPD spectrum would, consequently, be a superposition of several spectra. A possible indication for this is that our observed amide-II band is significantly broader than calculated.

Doubly protonated gramicidin S in the gas phase might thus not be a good model system for an isolated $\beta$-sheet peptide, as the $\beta$-sheet can be disturbed by side chain – backbone interactions. However, the spectrum of $[\text{GramS} + 2\,\text{H}]^{2+}$ complexed with two 18-crown-6 molecules might be expected to resemble more closely that of an unperturbed antiparallel pleated $\beta$-sheet conformation. There, the amide-I peak appears to consist of two components, one near 1653 cm$^{-1}$ and one near 1680 cm$^{-1}$. Gramicidin S in solution shows a band at 1629 cm$^{-1}$ when dissolved in $D_2O$, which shifts to 1652 cm$^{-1}$ when DMSO is used as a solvent [223]. $D_2O$ is a polar and protic solvent which can solvate the charged side chains and can be involved in H-bonding to the six C=O groups and four N−H groups which are not part of the $\beta$-sheet section of the molecule. DMSO is a polar but aprotic solvent, which can solvate the charged side chains, but which is, however, less effective in H-bonding. The situation in DMSO could thus be quite comparable to the situation when the charged side chains are complexed with 18-crown-6. This is in line with our observation that the peak at the red side in the amide-I band of gramicidin S occurs at essentially the same frequency. However, a comparison between experiment and theory indicates that this peak is not caused by the $\beta$-sheet C=O groups but rather by the Orn and Pro C=O groups and that the Val and Leu C=O groups contribute more to the blue part near 1680 cm$^{-1}$.

For proteins in the condensed phase, amide-I vibrations of $\beta$-sheet molecules are observed near 1633 cm$^{-1}$ [42]. Small gas-phase $\beta$-sheet model systems show amide-I resonances that are all above 1650 cm$^{-1}$ [221] and it has been pointed out, that a classification of the structure just based on the position of amide-I resonances is not possible for those systems [221]. However, it is also conceivable that the condensed phase value of 1633 cm$^{-1}$ is applicable only for $\beta$-sheet sections that interact strongly with their surroundings, such as $\beta$-sheet sections that are embedded in larger protein

structures or $\beta$-sheet sections that are part of an amyloid network. For isolated $\beta$-sheet sections, the 1633 cm$^{-1}$ value might not apply. In line with that is the observation that for $\beta$-sheet sections that interact only with polar solvents, a shifted value of 1640 cm$^{-1}$ can be observed, however at high concentrations, when aggregation takes place, a new peak at 1616 cm$^{-1}$ is observed [222]. The FTIR studies on gramicidin S dissolved in various solvents [223], the data on small $\beta$-sheet model systems [221] as well as the data presented here indicate that the intrinsic frequency of $\beta$-sheet peptides might be substantially to the blue, more in the range 1650 cm$^{-1}$ –1680 cm$^{-1}$.

For helical or random coil systems, the dependence on the environment might be less pronounced. In an isolated $\beta$-hairpin, only half of the C=O and N$-$H groups have a partner to interact with and the other half has to interact with either the environment or has to interact intramolecularly via weak C$_5$ hydrogen bonding. In helices, on the other hand, only the terminal C=O or N$-$H groups do not have a partner. In agreement with that is the observation that for cytochrome c in the gas phase, which has most likely a predominantly random coil or helical structure, an amide-I band between 1660 cm$^{-1}$ and 1670 cm$^{-1}$ [146] is observed - close to the value in the condensed phase [154], see Chapter 3 for further discussions.

## 5.6   Conclusions

The mid-infrared spectrum of the doubly protonated (deuterated) gas-phase peptide gramicidin S has been measured at 300 K using IR-MPD in an FT ICR cell. The spectra show a broad split peak between 1580 cm$^{-1}$ and 1720 cm$^{-1}$. A comparison to theory shows the importance of the interactions of the charged side chains with the backbone, which is responsible for the lower frequency part of the amide-I peak. When this interaction is inhibited via complexation with two 18-crown-6 molecules, the amide-I peak narrows and shows two maxima at 1653 cm$^{-1}$ and 1680 cm$^{-1}$. A comparison to calculations shows that for this complexed ion, four C=O groups are in an antiparallel $\beta$-sheet arrangement. The data thus represents the first evidence for the stability of a charged $\beta$-sheet peptide in the gas phase. An analysis of the calculated spectra shows that those $\beta$-sheet C=O groups give rise to the vibrations near 1680 cm$^{-1}$, in contrast to expectations based on values for the condensed phase, where resonances of $\beta$-sheet sections are thought to occur near 1630 cm$^{-1}$. The difference between those values may be caused by interactions with the environment, as the condensed phase value is mostly deduced for $\beta$-sheet sections that are embedded in larger proteins, that interact strongly with solvent or that are part of partially aggregated species.

# Chapter 6

# Gasphase IR spectra of *coiled coil* protein complexes[*]

Electrospray ionization (ESI) is the softest ionization method that is currently available and it is widely accepted, that ESI generated ions of proteins and protein assemblies at certain conditions retain characteristic aspects of their solution-state conformation. ESI mass spectrometry (MS) therefore evolved as a useful tool to obtain information on composition, stoichiometry, and dynamics of non-covalently associated protein complexes. While tertiary structure information of proteins can be obtained from ion mobility spectrometry (IMS), only a few techniques yield direct information on the secondary structure of gas-phase peptides and proteins. We present here the mid-IR spectroscopic secondary structural analysis of three *de novo* designed $\alpha$-helical *coiled coil* model peptides and their non-covalently associated complexes in the gas-phase. The conformational stability of such *coiled coil* peptides in solution is primarily driven by aggregation. Isolated monomers usually remain unfolded. Two of the investigated peptides were designed to assemble into stable $\alpha$-helical complexes in acidic solution, while the third one remains monomeric and unfolded at these conditions. Monomer ions of all three peptides show comparable photodissociation IR spectra and therefore suggest an unfolded conformation in the gas phase. In contrast, considerable C=O stretch (amide-I) and N−H bend (amide-II) band shifts have been observed for the dimers which is consistent with an elevated H-bond content. These findings provide evidence that at least a fraction of the condensed phase $\alpha$-helical structure is retained in the gas phase *coiled coil* complexes.

## 6.1 Introduction

In recent years, electrospray ionization (ESI) mass spectrometry (MS) evolved as a valuable tool to enable the analysis and characterization of the stoichiometry, dynamics, and shape of non covalently associated protein complexes [27, 246–250]. Instrument modifications that involve changes in the differential pumping region [251] as well as the development of nano-ESI sources [252] have been pioneering achievements,

which nowadays make mass spectrometry a key-technique in structural genomics and proteomics [27]. Very recent investigations showed that even membrane protein complexes, which are usually extremely difficult to analyze, can be transferred intact into a mass spectrometer [253].

A crucial requirement for the analysis of an intact protein complex in the gas phase is the (at least partial) conservation of the condensed-phase conformation. Several investigations on large protein assemblies provided evidence that important structural features such as a compact globular shape, secondary structure elements, hydrogen bonding patterns, and characteristic quaternary contacts can be maintained at certain instrument and solution conditions [25]. Soft landing approaches for example showed that viruses retain their shape and infectivity after ES ionization, transfer, and deposition on a surface placed at the end of the mass spectrometer [254]. In addition, ion mobility spectrometry (IMS) [214], especially when coupled to MS, can provide valuable insights on the gas-phase conformation of proteins and protein complexes [215, 255]. Investigations on the undecameric complex of the tryptophan-RNA binding attenuation protein (TRAP) for example revealed, that the characteristic ring-like topology present in solution can be retained in the absence of solvent [256]. Furthermore, IMS was successfully applied to characterize the quaternary organization of small, soluble A$\beta$42 oligomers, which are thought to be the toxic component in Alzheimer's disease [257]. On the other end of the molecular scale, there have been various combined IMS and molecular dynamics (MD) studies on small peptides and peptide complexes, which adopt unique and defined *in vacuo* conformations [204, 258]. Most of these peptides are alanine- and/or glycine-rich which makes them easy to calculate and understand, but also hardly soluble in aqueous solutions. As a consequence, often rather harsh source conditions and a high acetic acid or trifluoroacetic acid (TFA) content are required to electrospray these peptides. However, various well defined conformations such as isolated $\alpha$-helices [203, 259], helix-turn-helix motifs [258], and non-covalently associated helices [260] with partially impressive thermal stability have been observed for these systems. According to the properties of the gas-phase environment, mainly charge-charge, charge-dipole, and dipole-dipole interactions have been found to contribute to their conformational stability.

Despite the enormous differences between small *in vacuo* peptides and large solution-like protein assemblies, there is a common problem – a tremendous lack of knowledge on the secondary structure level. In the condensed phase, direct measurements of the secondary structure can be easily performed using circular dichroism [6] or infrared (IR) spectroscopy [261]. Until a few years ago, there was no similar method for gas-phase structural analyses, but recently IR spectroscopy became available to gain secondary structure information of biomolecular ions in the gas-phase. A limited amount of species has been investigated since then, including amino acids [262], dipeptides [187, 263], tripeptides [190], longer polypeptides [105, 183, 207, 264, 265], and the full length protein cytochrome c [146]. However, to date relatively little has been done on the gas-phase IR spectroscopic analysis of protein complexes. More importantly, nothing is known on the secondary structural differences between protein complexes and their corresponding subunits in the gas phase. This kind of information is of outstanding importance, because a quaternary organization is often crucial for folding (and also misfolding) in solution.

Within this chapter gas-phase IR spectroscopic analysis of one of the smallest and most stable protein assemblies known from nature – the $\alpha$-helical *coiled coil* is presented. The structural characterization in solution was performed by CD spectroscopy, while infrared multiphoton dissociation (IR-MPD) spectroscopy was employed to obtain secondary structure data of dimeric *coiled coil* complexes and their corresponding monomers. Perceptible differences between both species suggest that the cooperative, oligomerization-driven folding known for *coiled coils* in solution to some extend also exists in the solvent-free environment. In addition, the here presented data provide evidence that at least a fraction of the helical conformation is retained in the gas phase.

## 6.2  Experimental

### 6.2.1  Peptide synthesis and purification

Model peptides "VW01", "VW02", and "VW03" were synthesized in the group of Beate Koksch by standard Fmoc chemistry on Fmoc-Leu-OWang resin (0,68/0,71 mmol/g) using a 431 A peptide synthesizer (Applied Biosystems, Darmstadt, Germany). The peptides were cleaved from the resin by reaction with 3 mL of a solution containing 5% (v/v) triisopropylsilane, 0.1% (v/v) water, and 94.9% (v/v) TFA. Purification was carried out by preparative reversed phase high-performance liquid chromatography (HPLC) on a Vydac C4 column. The molecular weight of the products was determined by MALDI-TOF mass spectrometry using a Voyager MALDI-TOF Mass spectrometer (Perseptive/Applied Biosystems, Darmstadt, Germany) and its purity was determined by analytical HPLC on a Merck LaChrom system (Merck, Darmstadt, Germany) equipped with a Vydac C4 (10 $\mu$m) column (Grace Vydac, Columbia, MD, USA). The gradient used was similar to those of the preparative HPLC.
"VW02Abz" was synthesized by a solid phase assembly using a Multi-Syntech Syro XP peptide synthesizer (MultisynTech, Witten, Germany) by Fmoc strategy on Fmoc-Leu-OWang resin (0.65 mmol/g). In contrast to "VW02", "VW02Abz" was N-terminally capped with anthranilic acid (Abz)[1]. The peptide was cleaved from the resin by reaction with 4 mL of a solution containing 10% (w/v) triisopropylsilane, 1% (w/v) water, and 89% (w/v) TFA. The crude product was purified by reversed-phase HPLC on a Knauer smartline manager 5000 system (Knauer, Berlin, Germany) equipped with a C8 (10 $\mu$m) LUNATM Phenomenex column (Phenomenex, Torrance, CA, USA). The peptide was eluted with a linear gradient of acetonitrile/water/0.1% trifluoroacetic acid and identified by MS using an Agilent 6210 ESI-ToF LC/MS spectrometer (Agilent Technologies, Santa Clara, CA, USA) with direct infusion via a Harvard Apparatus 11plus syringe pump (Harvard Apparatus, Holliston, MA, USA). Purity was determined by analytical HPLC on a Merck LaChrom system (Merck, Darmstadt, Germany) equipped with a C8 (10 $\mu$m) LUNATM Phenomenex column (Phenomenex, Torrance, CA, USA). The gradient

---

[1]2-Aminobenzoic acid (Abz, common name anthranilic acid) is often used as UV-label to enable photometric concentration determination in solution. Unpublished results from the group of Beate Koksch indicate that this modification has no noticeable influence on the condensed phase folding properties of a 26-residue *coiled coil* peptide.

used was similar to those of the preparative HPLC.

## 6.2.2   Circular dichroism spectroscopy

Circular dichroism (CD) measurements of peptides in buffer were carried out on a J-715 CD spectrometer equipped with a temperature controller (Jasco, Easton, MD, USA) using a quartz cell of mm path length. Spectra were recorded at 25° C from 200-240 nm at 0.5 to 0.2 nm resolution with a scan rate of 20 nm/min and a sensitivity of 100 mdeg. Three scans were acquired and averaged for each sample. Raw data were manipulated by smoothing and subtraction of buffer spectra. The concentration was 0.8 mg/mL, which corresponds to approximately 250 $\mu$M. CD values are given in mdeg and were not converted into molar ellipticities, since concentration in solution was not determined independently and only roughly estimated by the weighted amount of peptide. The following buffers were used: pH 4.0, pH 5.0: 10 mM acetate buffer; pH 7.4, pH 8.0: 10 mM Tris-HCl buffer.

## 6.2.3   Infrared photodissociation experiments

The experiments were performed at the free electron laser facility FELIX [65] (Nieuwegein, the Netherlands) by using the Fourier-transform ion cyclotron (FT ICR) mass spectrometer [57], see Section 1.4.3. Ions were generated via electrospray ionization. In order to avoid higher-order aggregation, which might affect the ESI process, exclusively freshly prepared samples were used for the presented experiments. Therefore, 0.15 mg of peptide was dissolved freshly in 1 mL $NH_4O$For buffer (10 mM, pH 4.5, Fisher Scientific, Schwerte, Germany). Subsequently, the pH was adjusted to 4.5 using diluted formic acid solution and 40% MeOH (v/v, Acros Organics, Geel, Belgium) was added to obtain a stable and reproducible spray. The solution was sprayed via a syringe pump Harvard Apparatus 11plus (Harvard Apparatus, Holliston, MA, USA) and a standard Z-spray source (Waters Corporation, Milford, MA, USA ) connected to the mass spectrometer. The ESI generated ions were transported and accumulated in a hexapole ion trap and subsequently transferred into a home-built FT ICR mass spectrometer that is optically accessible via a KRS-5 window at the back end. After trapping and mass-selective isolation of the charged molecules of interest inside the ICR cell, the ions were irradiated by IR photons of the free electron laser FELIX. When the IR light is resonant with an IR active vibrational mode in the molecule, this results in the absorbance of many photons, which causes dissociation of the ion (IR-MPD), see Section 1.4.1. Monitoring the fragmentation yield or the depletion of the parent ion signal as a function of IR wavelength leads to the IR spectra, see Section 1.4.2. Recent studies showed that the so-obtained spectra are not entirely identical to linear IR absorption spectra, but can be very close to it [58]. The output of FELIX is continuously tunable over a range of 40 to 2000 cm$^{-1}$ . In the presented experiment only the range from 1300 to 1850 cm$^{-1}$ was scanned. The light consists of macropulses of about 5 $\mu$s length at a repetition rate of 10 Hz, which contain 0.3 –5 ps long micropulses with a micropulse spacing of 1 ns. In the present experiment, macropulse energies were in the range 35 mJ with a bandwidth of approximately 1%.

## 6.3 Results and discussion

### 6.3.1 Peptide design

Within this project, three different $\alpha$-helical *coiled coil* model peptides have been investigated. *Coiled coils*, also often referred to as Leucine zipper proteins, are very common in nature and approximately 3-5% of all amino acids in naturally occurring peptides and proteins are involved in their formation [266]. Therefore, *coiled coils* have been studied extensively within the last 20 years, which led to a profound understanding of the interactions that govern thermodynamic stability and specificity of folding; making it one of the best understood protein folding motifs [267, 268]. Today, artificial coiled coils with properties that are optimized for a particular purpose or application can be designed *de novo* according to well established principles [269–271]. $\alpha$-helical *coiled coils* typically consist of two to five right-handed $\alpha$-helices which are wound around one another forming a left-handed supercoil. Figure 6.1 a and 6.1 b exemplarily show the molecular modeling structure of an antiparallel 41-residue coiled coil. The primary structure of each helix is characterized by a periodicity of seven residues, the so-called 4-3 heptad repeat which is commonly denoted $(abcdefg)_n$. For design purposes, these seven residues are often presented in a so-called helical wheel diagram, which, in a simplified manner, depicts the alignment of residues along the helical axis (figure 6.1 c). Positions a and d are typically occupied by apolar residues (Leu, Ile, Val, Met) that form a special interaction surface at the interface of the helices by hydrophobic core packing ("knobs-into-holes") [267–269]. In contrast, positions e and g are frequently occupied by charged amino acids (most commonly Glu, Arg, and Lys) that form inter-helical Coulomb interactions [268]. Polar residues are often found in the remaining heptad repeat positions b, c, and f, which are solvent exposed at the opposite side of the helical cylinder.

In solution, the hydrophobic core provides the major contribution to the thermodynamic stability of the $\alpha$-helical *coiled coil*. In contrast, the inter-helical ionic pairing positions e and g mainly dictate the specificity of folding (parallel versus antiparallel) and promote the preference for homo- or heterotypic $\alpha$-helical *coiled coil* formation [272–274]. An additional interaction domain, which is not related to the *coiled coil* oligomerization, is formed by intramolecular Coulomb interactions between positions c/g and b/e, respectively. These interactions indirectly influence the stability of $\alpha$-helical *coiled coil* complexes by stabilizing or destabilizing the single helices [275, 276]. The combination of all three previously described interactions makes *coiled coil* folding a highly cooperative process, *i.e.* the helical structure cannot exist without oligomerization and *vice versa*. As a result, isolated monomers as they for example occur at very low concentrations remain unfolded in solution.

Idealized *coiled coils* are often extraordinarily resistant to urea- and temperature-driven denaturation in solution. However, conditions in the gas phase are explicitly different, which may affect the previously described *coiled coil* recognition motifs tremendously. Due to the absence of bulk water hydrophobic interactions are thought to be reduced to their van der Waals fraction, while Coulomb interactions and hydrogen bonds are assumed to increase substantially in the gas phase [25, 277]. Even though hydrophobic core packing is the major determinant for *coiled coil* folding in solution, ESI generated
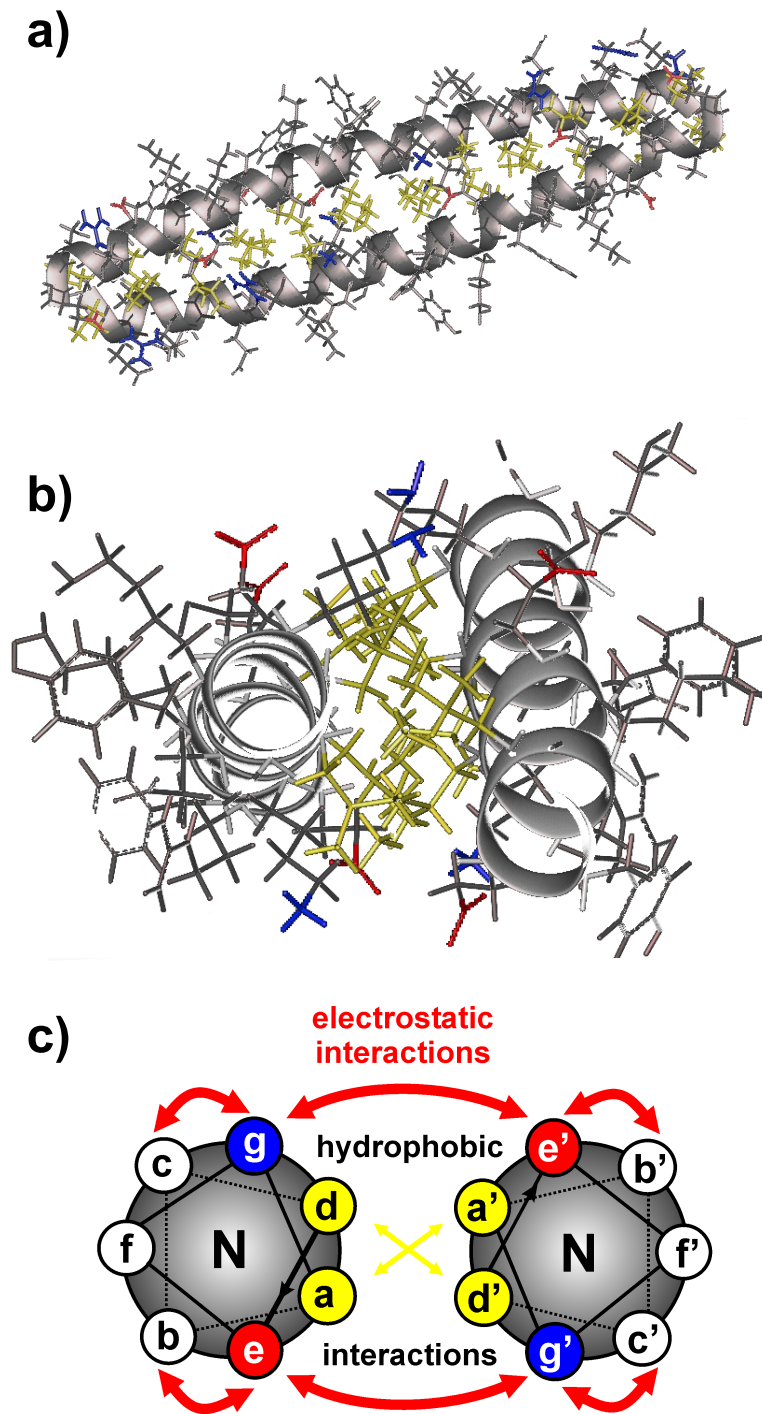
**Figure 6.1:** Molecular modeling structure representing the solution state conformation of a dimeric *coiled coil* peptide. a) View perpendicular to the axis of the helix and b) view along the helical axis. c) Helical wheel diagram of a parallel *coiled coil* showing the three recognition motifs. Yellow residues represent hydrophobic interactions, while coulomb interactions are occurring between negatively (red) and positively charged residues (blue).

ions of oligomers can be transferred, stored and detected in the solvent free environment

of a mass spectrometer [274, 278, 279]. Surprisingly, this is even possible at conditions and instrument setups which are not particularly suitable for the analysis of intact protein complexes. From these findings, the question arises how oligomers of *coiled coil* peptides are held together in the gas phase and, more importantly, whether this is accompanied by a conservation of the helical conformation. If so, perceptible differences in the gas phase IR spectra of monomers and oligomers should occur. To shed light on these questions, three different 26-residue *coiled coil* model peptides have been designed, characterized and investigated by gas-phase mid-IR photodissociation spectroscopy. Peptides "VW01" and "VW02" (see Figure 6.2) possess highly comparable sequences and were both designed to provide for maximum *coiled coil* stability. Positions a and d within the heptad repeat are exclusively occupied by leucine residues, which ensure an efficiently packed hydrophobic core. Furthermore, positions e and g are equipped with charged glutamic acid and lysine residues, respectively, forming attractive Coulomb interactions in case of a parallel helix alignment. The adjacent b and c positions are occupied by oppositely charged lysine or glutamate residues which further stabilize the α-helical structure by intramolecular ion-pairing. Due to their optimized sequence, both peptides are expected to adopt a stable α-helical *coiled coil* conformation at a wide range of pH values.
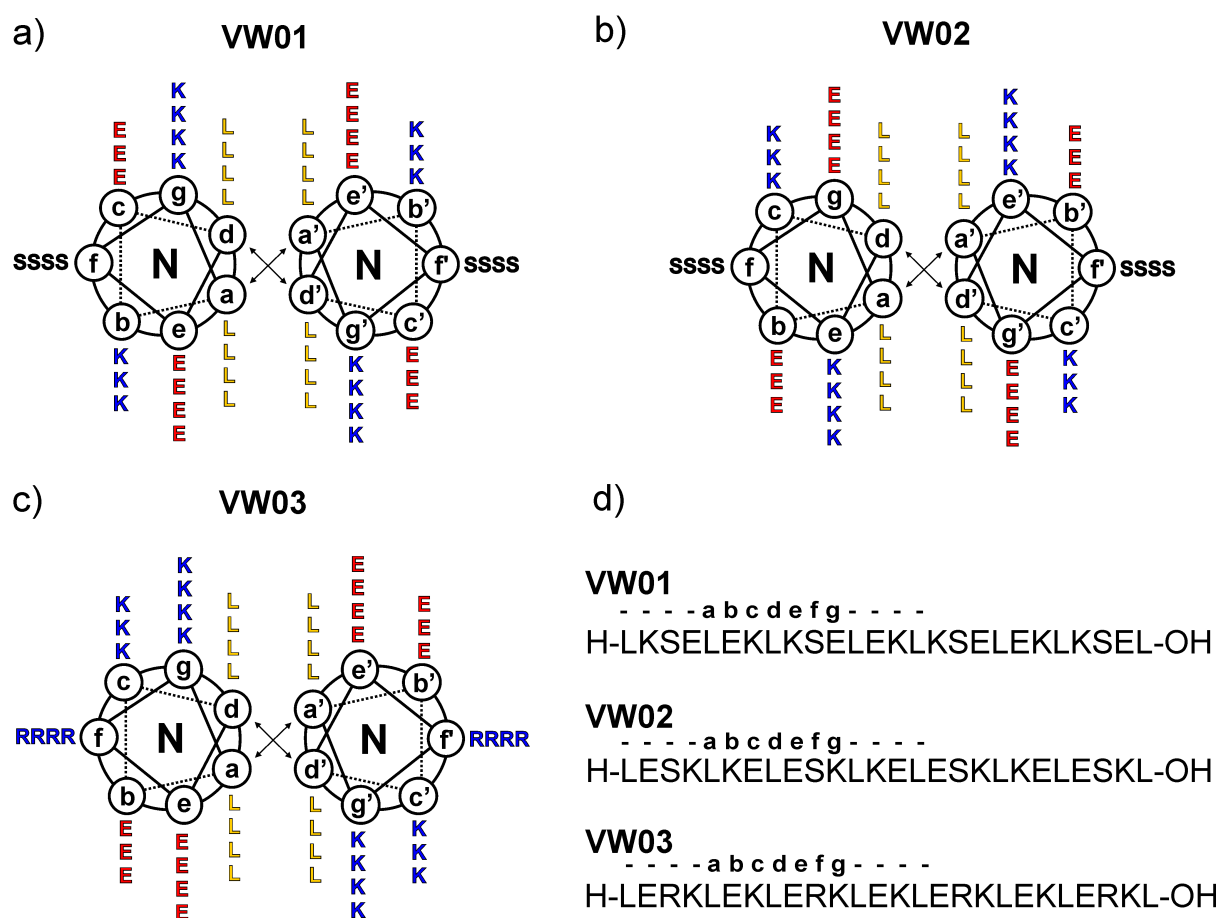


**Figure 6.2:** a-c) Helical wheel Diagram and d) sequence of peptides "VW01", "VW02", and "VW03". Hydrophobic residues are depicted in yellow, acidic residues in red, and basic residues in blue.

ESI ionization is facilitated at acidic conditions and peptide "VW03" was designed to serve as a negative control that remains unfolded at these conditions. As in "VW01" and "VW02", the hydrophobic core positions a and d are exclusively equipped with Leu residues and positions e and g exhibit a similar pattern of oppositely charged Glu and Lys residues, respectively. The intrinsic tendency to form a *coiled coil* structure should therefore be comparable to that of "VW01" and "VW02". However, in contrast to "VW01" and "VW02", positions c and f were additionally modified with similarly charged Lys and Arg residues. Intramolecular coulomb repulsions between these residues are known to destabilize the helical structure at acidic conditions and, consequently, inhibit *coiled coil* oligomerization. Such "pH switches" are a well established tool in solution studies and have been applied to generate pH and ionic-strength dependent amyloid forming model systems [280, 281].

### 6.3.2   Folding in solution

Prior to gas-phase experiments, peptides "VW01", "VW02", and "VW03" have been characterized by CD spectroscopy regarding their pH dependent folding behavior in solution. The spectra recorded at a peptide concentration of approx. 250 $\mu$M are shown in Figure 6.3. Similar CD signatures have been obtained at lower concentrations down to approx. 20 $\mu$M (data not shown).

As expected, "VW01" and "VW02" adopt a clear $\alpha$-helical conformation with two characteristic minima at 208 and 222 nm (see Figure 6.3 a and b). In contrast, "VW03" only exhibits a helical conformation at neutral and slightly basic pH, while characteristic random coil spectra with a minimum at 200 nm (see Figure 6.3 c) reveal a primarily unfolded conformation at pH 4.0 and 5.0. This was expected from the design and results from intramolecular coulomb repulsions between Lys and Arg residues at positions c, g, and f which are forming the pH switch. The OFF state of this pH switch at pH 7.4 and 8.0 on the other hand results in the formation of a defined helical structure. Consequently, "VW03" still exhibits a high propensity to assemble into *coiled coil* complexes, which makes it an ideal control peptide that, with a sequence that is comparable to that of "VW01" and "VW02", remains unfolded at acidic ESI conditions.

### 6.3.3   Oligomers in the gas phase

"VW01", "VW02Abz" (see footnote on page 117) (uncapped VW02 was not available at the time the experiments took place) as well as "VW03" were dissolved in buffered solution (10 mM $NH_4OFor$, pH 4.5) at a concentration of approx. 50 $\mu$M. Subsequently, the pH was controlled, adjusted if necessary and 40% (v/v) methanol was added. This solution was sprayed in the ESI source. The flowrate of the syringe pump was set to 8 $\mu$l/min, the ESI voltage to 3.7 kV, the cone voltage to 86 V and the temperature of the desolvation gas (N2, flowrate 200 L/h) to 60° C. For all molecules, monomer ions in various protonation states were observed. For "VW01" and "VW02Abz", monomer ions with 3, 4, and 5 excess protons and for "VW03", monomer ions with 4, 5, and 6 excess protons, were observed (however, the abundance of the "VW03" +6 ion was too low to allow recording of a spectrum). Dimer ions at the charge state +5 were observed in the case of "VW01" (see Figure 6.4) and "VW02Abz". Their relative intensity in comparison
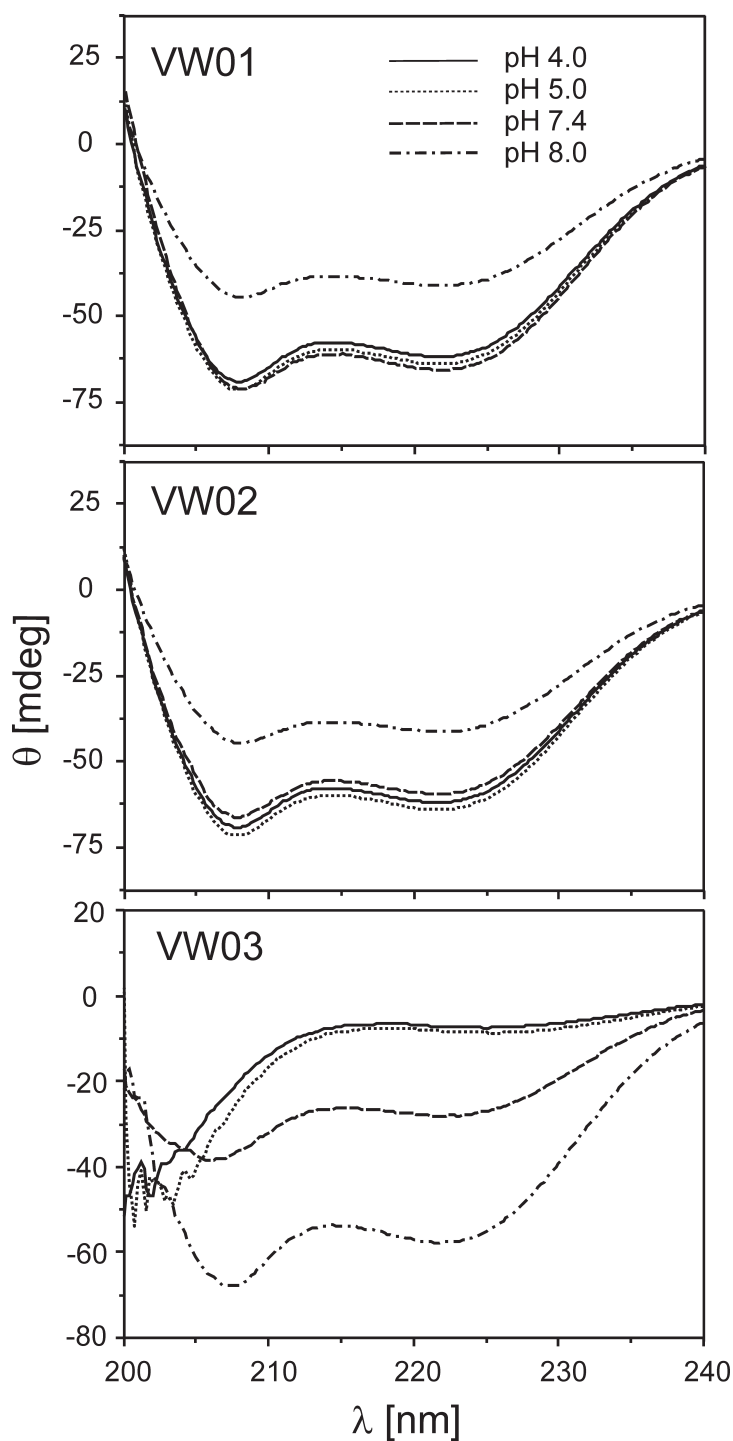
**Figure 6.3:** . CD spectra of approx. 250 $\mu$M peptide "VW01", "VW02", and "VW03" at pH 4.0, 5.0, 7.4, and 8.0. "VW01" and "VW02" exhibit a clear helical conformation at all investigated pH values while "VW03" remains unfolded at pH 4.0 and 5.0.

to the most intense monomer peak was in the range of 30%. Dimers were observed at concentrations as low as 20 $\mu$M, however no higher oligomers were found.
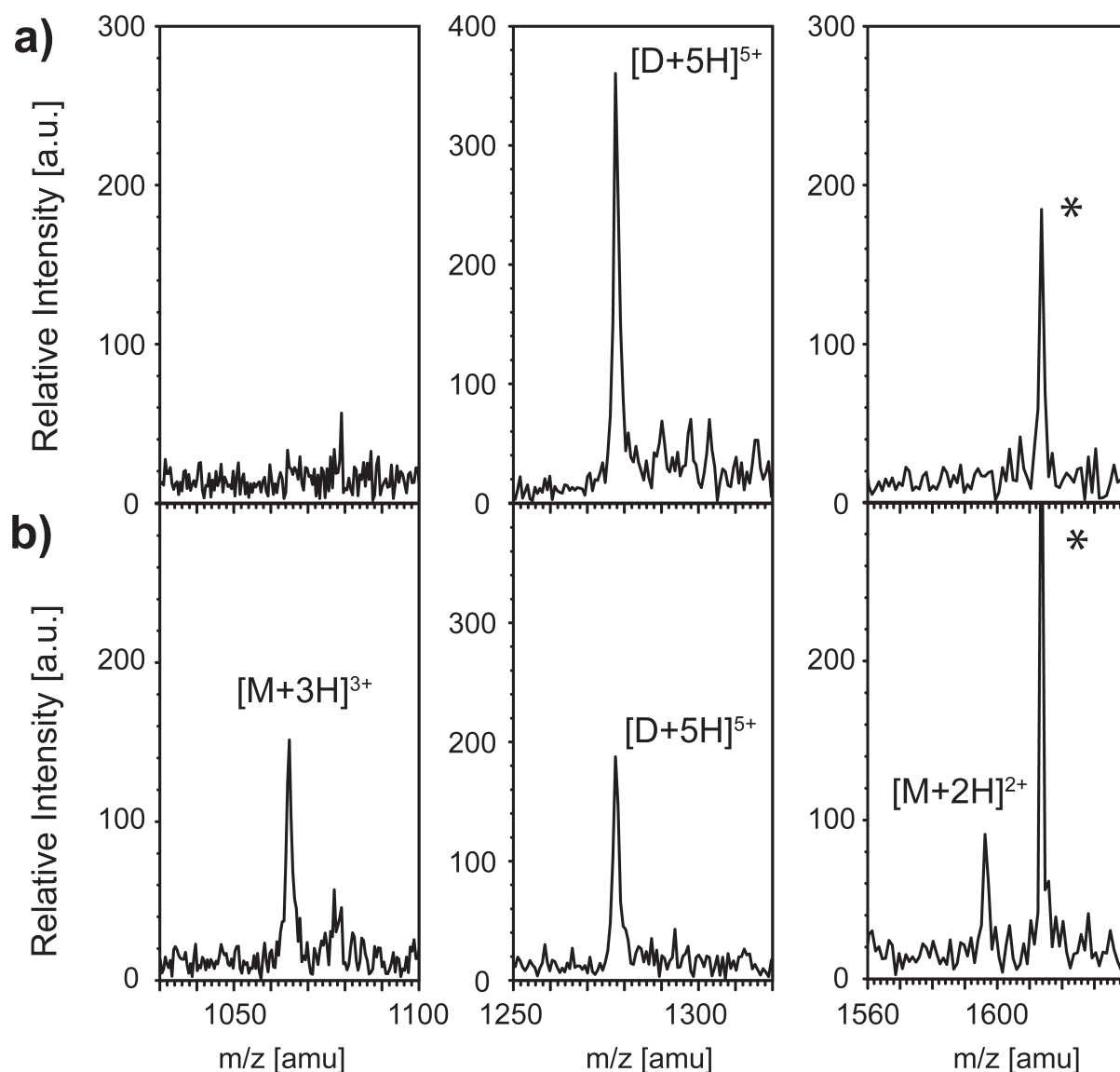
**Figure 6.4:** Mass spectra of swift-isolated "VW01" dimers (center, $[D + 5\,H]^{5+}$) after irradiation with FELIX at a) 5.5 $\mu$m and b) 6.1 $\mu$m. IR-MPD at 6.1 $\mu$m results in a fragmentation of the complex into the corresponding monomers (left, $[M + 3\,H]^{3+}$; right $[M + 2\,H]^{2+}$) while no dissociation occurs at 5.5 $\mu$m. Asterisks indicate electromagnetic interferences which occur sample-independently (see Section 1.4.4).

## 6.3.4   Photodissociation spectra of monomers

In order to record IR spectra, all ions, except the m/z ions of interest, are ejected by employing stored waveform inverse Fourier transform (SWIFT) excitation. The remaining ions are then irradiated by FELIX, and the fragmentation yield is recorded as a function of IR wavelength. Because of the roughly indicate resolution, it cannot be fully excluded that the observed monomer ions are slightly superimposed by the corresponding even-charged dimers. However, due to the much greater relative intensities of monomers in comparison to dimers it can be assumed that the majority of ions is indeed in the monomer state.

Figure 6.5 shows the acquired IR-MPD spectra for "VW01", "VW02Abz", and "VW03" monomer units for several charge states. Fragmentation occurs into several product ions, which were not assigned. In all spectra, two main peaks can be recognized. The peaks at higher wavenumbers fall into the C=O stretching (amide-I) vibration region while the peaks at lower wavenumbers are in a region, where N–H bending (amide-II) modes are expected to occur. In the top part, the spectra for "VW01" having the charge states +3, +4, and +5 are shown. The peaks in the amide-I region are all centered near 1662 $cm^{-1}$ and their width seems to increase slightly with increasing charge state. At the blue side of this band, a shoulder near 1745 $cm^{-1}$ is observed, which is more pronounced in the spectrum of the +5 ion of "VW01" than in the spectra of the other two peptides. In the amide-II region of the "VW01" spectra, the peaks for the +3 and +4 species are virtually identical. For the +5 ion, however, the amide-II peak becomes more intense, while also shifting to lower frequency and broadening.

The spectra of "VW02Abz" in the charge states +3, +4, and +5 are shown in the center of Figure 6.5. Similar peak positions, widths and trends as in the spectra of "VW01" are observed. The main differences to the case of "VW01" is that for "VW02Abz" in the +3 charge state, the peak near 1745 $cm^{-1}$ is missing and that in the spectra of the +4 ion, the amide-I and -II bands are broader and slightly shifted.

For the "VW03" monomer, only the +4 and +5 charge states could be generated at ion densities sufficient to record IR spectra. These are shown in the lower part of Figure 6.5. The spectra of those ions are similar to the spectra of the lower charge states of "VW01" and "VW02Abz".

For all three VW monomers considered, the spectra appear very similar. One exception are the spectra of "VW01" and "VW02Abz" in the +5 charge states, for which the amide-II bands are red-shifted, as well as displaying increased intensities. Further, for all ions, with increasing charge state, the amide-I bands shows a slight broadening towards the lower-frequency side. Nonetheless, the average position of the amide-I band is the same for all three ions for their respective charge states. This suggests that all the monomer ions have similar structures.

### 6.3.5   Photodissociation spectra of dimers

For "VW01" and "VW02Abz", the dimer ions having a +5 charge are observed in quantities sufficient to allow for the recording of IR spectra. For "VW03", no gas-phase dimers could be observed at the used experimental condition. When the $[(VW01)_2 + 5H]^{5+}$ or $[(VW02)_2 + 5H]^{5+}$ are isolated and irradiated by FELIX at wavelengths that allow for photon absorption and subsequent fragmentation, the corresponding +3 and +2 monomer ions are observed as fragmentation products (see Figure 6.4). In Figure 6.6, the IR-MPD spectra of $[(VW01)_2 + 5H]^{5+}$ and $[(VW02Abz)_2 + 5H]^{5+}$ are shown. For comparison, the corresponding spectra of the +3 monomer ions (which are also shown in Figure 6.5) are shown with dashed lines. The spectra of $[(VW01)_2 + 5H]^{5+}$ and $[(VW02Abz)_2 + 5H]^{5+}$ show amide-I and -II resonances that occur at virtually the same position and have the same width. This suggests that the structures of both types of dimers are very similar. In both dimer spectra, the amide-I and -II resonances are broader and shifted, compared to those of the monomers. For the amide-I mode this shift is about 11 $cm^{-1}$ to the red, while for the amide-II band, a 23 $cm^{-1}$ shift to the blue is
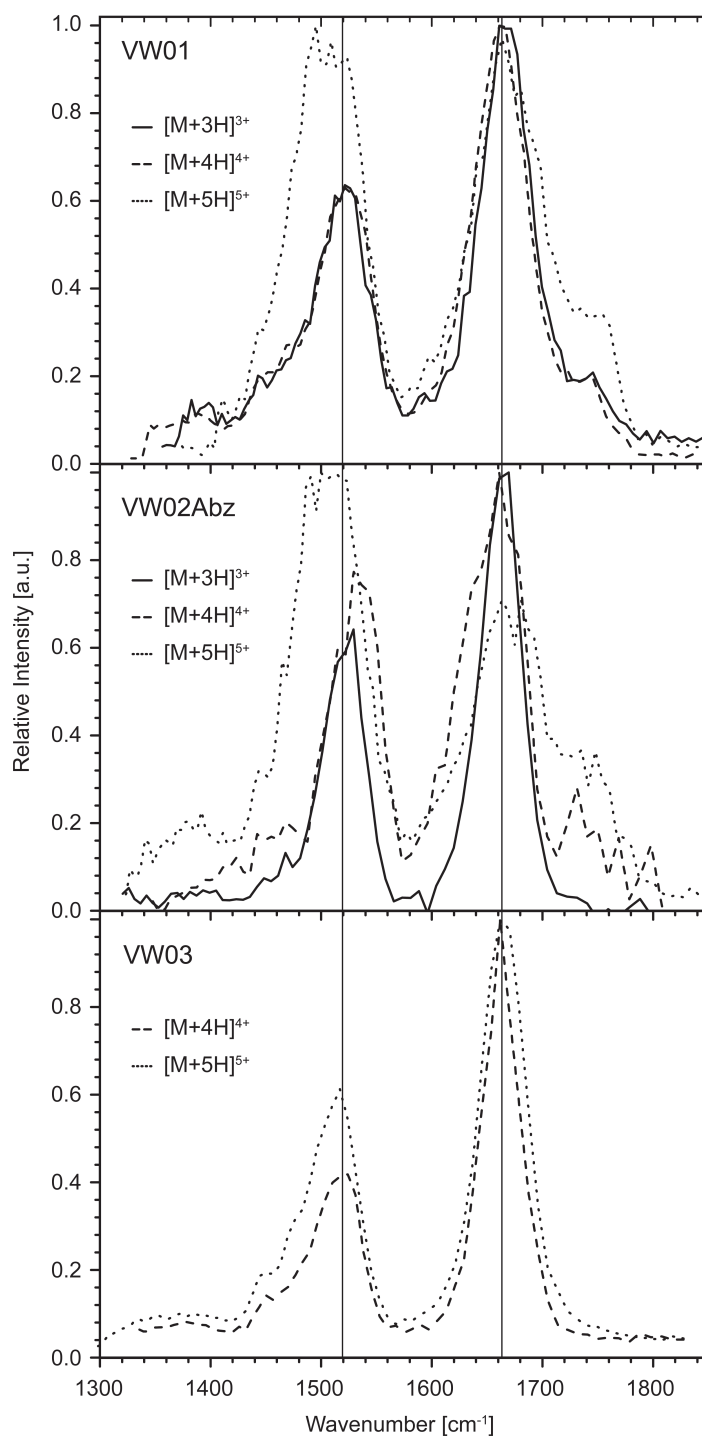
**Figure 6.5:** Mid-IR photodissociation spectra of "VW01", "VW02ABz", and "VW03" monomers. at charge state +3, +4, +5. For "VW03", charge state +3 was not observed.

observed. Different to the behavior of the monomers, some background fragmentation of the dimers can be observed at wavelengths that do not correspond to intense bands, such as the amide-I and -II resonances, thus resulting in an elevated baseline in their spectra.
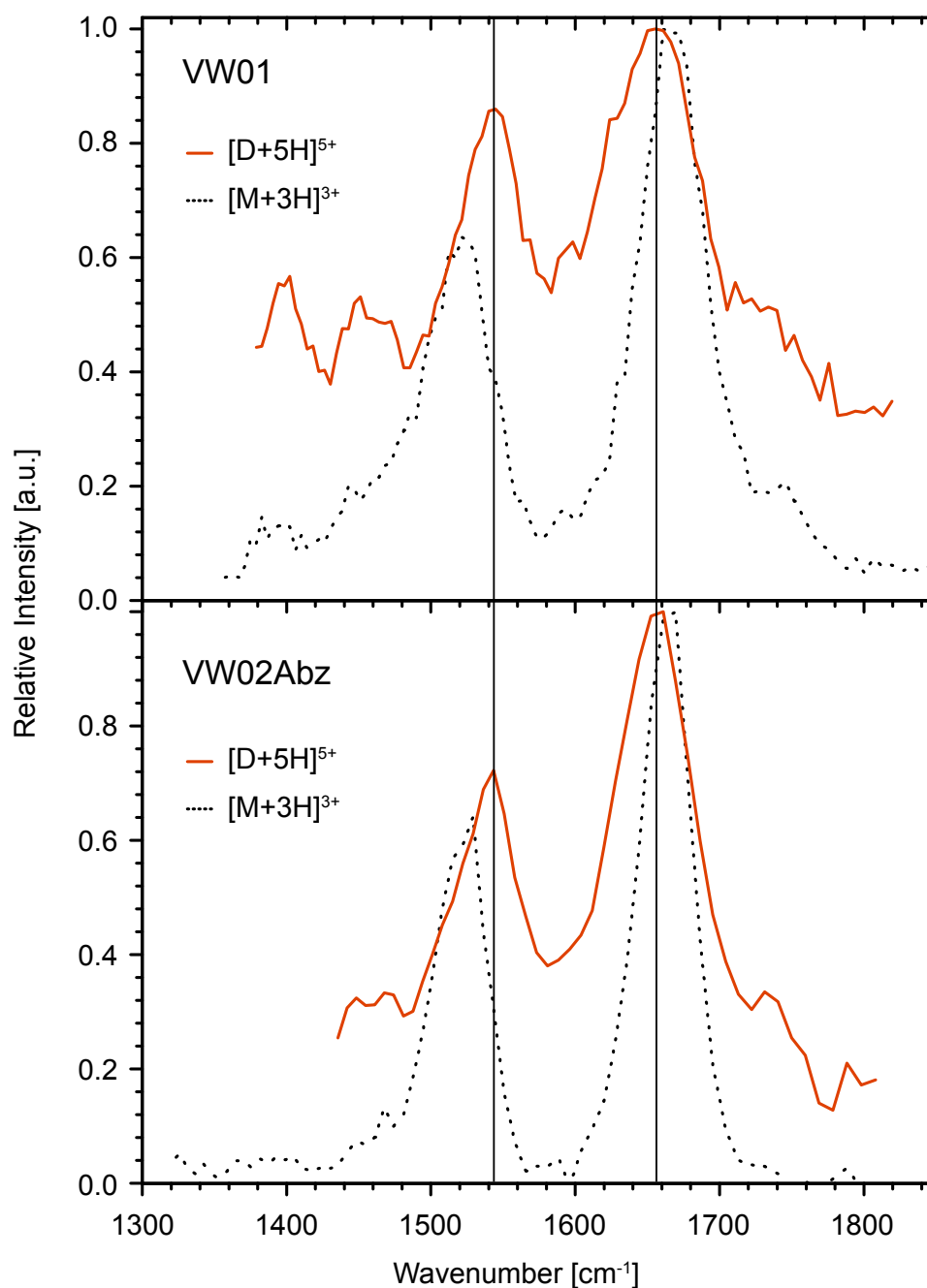
**Figure 6.6:** Mid-IR photodissociation spectra of "VW01" and "VW02Abz" dimers at charge state +5. The corresponding +3 monomers are shown for comparison purposes.

### 6.3.6 Structural implications for monomers and dimers

Based on the IR spectra, all monomer ions appear to have similar structures. For "VW03", intramolecular electrostatic interactions strongly destabilize the helical conformation and the solution data at comparable pH values as used in the gas-phase experiments reveal non-helical (presumably monomer) structures. It is thus likely that in the gas phase, the structure present for "VW03" is non-helical and more disordered, random-coil like. For "VW01" and "VW02Abz", intramolecular electrostatic interactions may support

helix formation even for monomers, however, the similarity of their spectra to those of "VW03" points to disordered random-coil structures for these peptides as well. With increasing charge state, the shoulder near 1750 cm$^{-1}$ is observed to become more prominent. A resonance at that position reveals the presence of non-hydrogen bonded C=O groups, either on the peptidic backbone, on the C-terminus or on a Glu side chain [282]. Likely additional protonation sites of the VW molecules are either the previously neutral basic side chains or the deprotonated COO$^-$ groups. Which of these processes occurs depends on the internal charge balance of the molecule. "VW01" has seven basic lysine residues, seven acidic glutamic acid residues as well as an acidic and a basic terminus. The amino acid content of "VW02Abz" is similar, except the N-terminus, which is capped with anthranilic acid. This cap also possesses a basic NH$_2$ group, however, its pK of 4.9 is significantly lower than that of a free N-terminus. In solution, at pH levels such as employed here, nearly all carboxylic acid groups of "VW01" should be deprotonated and all amino groups protonated, resulting in an overall uncharged, or slightly positively charged molecule. For "VW02Abz", the amino group at the anthranilic acid will be only partially protonated. When such an internal charge distribution persists in the gas phase, excess protons will be added to the COO$^-$ groups. This is in line with the observation that the intensity of the 1750 cm$^{-1}$ mode increases with the charge state. The situation is different in case of "VW03". Here, four additional arginine residues are present. It is thus likely that additional protons (up to a charge state of +7) will be added to the lysine or arginine side chains. In agreement with that, no 1750 cm$^{-1}$ mode is present in the spectra.

In the dimer spectra, the amide-I mode is shifted to the red and the amide-II mode to the blue, compared to the monomer spectra. The similarity between the two dimer spectra and the dissimilarity to the monomer spectra suggests that both dimers have comparable structures that are distinctively different from those of the monomers. However, a clear structural assignment on basis of the absolute band positions is difficult. To date, only a very limited amount of peptides and proteins with defined secondary structure have been characterized by gas-phase IR spectroscopy. Unlike in solution there are no reference values that clearly reveal a certain conformation. Nevertheless, the general tendencies known from the condensed phase [5] are also valid in the gas phase. Considering this, the observed shifts in the amide-I and -II positions are consistent with an increase in hydrogen bonding. Hydrogen bonds lead to a removal in electron density from the carbonyls oxygens, which in turn reduces the C=O stretch frequency (*i.e.*, red-shift of the amide-I band). Conversely, amide-II band are blue-shifted as a result of hydrogen bonding interactions. This can be rationalized by the fact that a hydrogen bond introduces "stiffness" into the N–H bending mode, thus requiring more energy for this motion. In solution, especially the amide-I mode strongly depends on the underlying secondary structure and red-shifts up to 50 cm$^{-1}$ are observed for H-bond-rich $\beta$-sheet proteins and protein aggregates [5]. The correlation between secondary structure and position of the amide-II band is less straightforward in comparison to the amide-I band [5]. Nevertheless, perceptible band shifts are also observed for this frequency region in solution. In the presented gas-phase spectra, perceptible band shifts in both amide modes are evident. According to the previously mentioned condensed phase principles, this provides evidence for an elevated hydrogen bond content in the dimer ions. Band broadening in case of the dimer ions furthermore

supports this hypothesis [283]. Therefore, it is likely that at least a fraction of the helical conformation is retained in the gas phase. Unlike in solution, the band shift is more pronounced for the amide-II band. It remains elusive if this is a unique behavior of the investigated peptides or a general difference between condensed-phase and gas-phase IR spectroscopy of proteins.

In the condensed phase, intramolecular as well as interhelical interactions strongly support the *coiled coil* structure. Those interactions include electrostatic (salt-bridge) interactions, hydrogen bonding as well as hydrophobic interactions between the central leucine residues. In the condensed phase, charged parts of the molecule will be tightly solvated by polar solvents. This solvation causes an effective steric and electrostatic shielding and reduces the relative importance of the electrostatic interactions. In the gas phase, no solvent shielding is present and the electrostatics will become the most important type of interaction. For "VW01" and "VW02Abz", such interactions will still stabilize a *coiled coil* conformation, but as those interactions become much more important compared to the hydrophobic interactions, it is not clear if the hydrophobic core will be conserved in the gas phase. Indeed, this is what is observed in preliminary MD simulations, where the leucine residues are found on the outside of a *coiled coil* structure[2]. In conclusion, it is likely that the spectra we observe in Figure 6.6 result from helix-rich *coiled coil*-like assemblies in the gas phase, however structural differences to the condensed phase situation are also expected.

## 6.4 Conclusions

We present here the IR spectroscopic characterization of three *de novo* designed $\alpha$-helical *coiled coil* peptides and their dimeric complexes in the gas phase. As shown by CD spectroscopy, two of these peptides assemble into non-covalent $\alpha$-helical *coiled coil* complexes at acidic pH, while the third peptide remains fully unfolded and monomeric at these conditions. Comparable mid-IR photodissociation spectra have been obtained for monomer ions of all three peptides, which points to a similar random coil-like conformation in the gas phase. Non-covalently associated dimer ions have been observed for peptides one and two, which, in comparison to monomers, exhibit considerably different photodissociation IR spectra. The position of the amide-I (C=O stretch) mode is shifted 11 cm$^{-1}$ to the red, while the amide-II band (N–H bend) is shifted 23 cm$^{-1}$ to the blue. This suggests an increased H-bond content in the dimers, which provides evidence that at least a certain fraction of the condensed phase helical structure is also retained in the gas-phase *coiled coil* dimer ions.

---

[2]Unpublished results

# Chapter 7

# Conclusions and future perspectives

In this Chapter, a summary of the results together with a discussion of the possibilities of infrared spectroscopy on gas-phase biomolecules for investigations of their structure as investigated in this thesis will be presented. Subsequently, an outlook on how to improve the infrared spectroscopy technique to determine structures of biomolecules in the gas phase is proposed.

## 7.1 Conclusions of the results obtained via IR spectroscopy on gas-phase biomolecules

In the condensed phase, infrared spectroscopy is widely used to investigate the presence and amount of certain secondary structure elements of proteins. In the wavelength range from 500 cm$^{-1}$ to 2000 cm$^{-1}$, the fingerprint region of a protein, observed vibrational band positions, in particular the amide-I band, allow for the assignment of structural features of molecules. Table 7.1 gives an overview of observed band positions assigned to several secondary structure elements as observed in the condensed phase. The aim of this thesis is to see in how far mid-IR spectroscopy can be used to obtain structural information on gas-phase biomolecules. For this, special attention is given to the amide-I and amide-II range in infrared spectra of peptides and proteins.

A summary of the observed amide-I and amide-II band positions for gas-phase biomolecules is given in Table 7.1. Remarks concerning expected structural motifs of the molecules are included, as well as observed conformations by other research groups for those molecules. The observed amide-I band position of some gas-phase biomolecules are graphically compared with observed amide-I band positions for different secondary structure elements in the condensed phase in Figure 7.1.

**Table 7.1:** Summary of observed amide-I and amide-II band positions for the molecules investigated in this thesis. Stated in the brackets are the band positions for several structural motifs as observed in the condensed phase, with the highest and lowest value being given. Table 1.1 in Section 1.3 shows the average values.

| Molecule | Amide-I [cm$^{-1}$] | Amide-II [cm$^{-1}$] | Remarks |
|---|---|---|---|
| Cyto c | | | $\alpha$-helix (solution) |
| 7H$^+$ | 1663 | 1535 | |
| 8H$^+$ | 1663 | 1529 | |
| 9H$^+$ | 1657 | 1534 | |
| 10H$^+$ | 1664 | 1548 | |
| 11H$^+$ | 1670 | 1544 | |
| 12H$^+$ | 1670 | 1543 | |
| 13H$^+$ | 1670 | 1544 | |
| 14H$^+$ | 1671 | 1542 | |
| 15H$^+$ | 1671 | 1543 | |
| 16H$^+$ | 1688 | 1524 | |
| *deut*-Cyto c | | | |
| 8D$^+$ | 1656 | 1419 | |
| 9D$^+$ | 1681 | 1450 | |
| 10D$^+$ | 1666 | 1437 | |
| 11D$^+$ | 1668 | 1433 | |
| 12D$^+$ | 1657 | 1455 | |
| 13D$^+$ | 1661 | 1456 | |
| 14D$^+$ | 1656 | 1448 | |
| 15D$^+$ | 1665 | 1456 | |
| 16D$^+$ | 1673 | 1429 | |
| 18D$^+$ | 1676 | 1423 | |
| 19D$^+$ | 1676 | 1425 | |
| Tendamistat | | | $\beta$-sheet (solution) |
| 5H$^+$ | 1664 | 1530 | |
| Ala$_n$H$^+$ | | | no $\alpha$-helix for n $\leq$ 20 |
| n=5 | 1704 | 1516 | |
| n=15 | 1670 | (1516) | |
| n=20 | 1662 | 1534 | |
| Ac-Ala$_{20}$X$^+$ | | | $\alpha$-helix with alkali ion |
| X=H | 1662 | 1534 | |
| X=K | 1680 | 1510 | |
| X=Na | 1675 | 1522 | |

Table 7.1 – Continued

| Molecule | Amide-I [cm$^{-1}$] | Amide-II [cm$^{-1}$] | Remarks |
|---|---|---|---|
| Ac-Ala$_n$LysH$^+$ | | | $\alpha$-helix for n $\geq$ 7 |
| n=5 | 1691 | 1521 | globular |
| n=10 | 1678 | 1524 | |
| n=15 | 1677 | 1527 | |
| n=19 | 1668 | 1527 | |
| Ac-LysAla$_{19}$H$^+$ | | | most likely dimer |
| | 1667 | 1521 | |
| Gram S | | | |
| 2H$^+$ | 1682 | 1501 | |
| 2D$^+$ | 1678 & (1640) | 1418 | |
| Gram S+crown | | | $\beta$-sheet model |
| 2H$^+$ | 1680 & 1653 | 1529 | amide-I strong |
| 2D$^+$ | 1681 & 1655 | N/A | |
| VW01 | | | M: VW01 |
| | | | D: VW01$_2$ |
| [M + 3 H]$^{3+}$ | 1662 | 1519 | random coil |
| [M + 4 H]$^{4+}$ | 1662 | 1519 | random coil |
| [M + 5 H]$^{5+}$ | 1662 | 1515 | random coil |
| [D + 5 H]$^{5+}$ | 1651 | 1544 | *coiled coil* |
| VW02Abz | | | M: VW02Abz |
| | | | D: VW02Abz$_2$ |
| [M + 3 H]$^{3+}$ | 1662 | 1520 | random coil |
| [M + 4 H]$^{4+}$ | 1662 | 1530 | random coil |
| [M + 5 H]$^{5+}$ | 1662 | 1510 | random coil |
| [D + 5 H]$^{5+}$ | 1651 | 1544 | *coiled coil* |
| VW03 | | | M: VW03 |
| [M + 4 H]$^{4+}$ | 1662 | 1519 | random coil |
| [M + 5 H]$^{5+}$ | 1662 | 1517 | random coil |
| Motif | | | condensed phase |
| $\alpha$-helix | [1648-1657] | N/A | |
| $\beta$-sheet | [1623-1641] | N/A | strong |
| | [1674-1695] | N/A | weak |
| turns | [1662-1686] | N/A | |
| | [1630-1700] | N/A | |
| random coil | [1642-1657] | N/A | |
| *coiled coil* | 1630 | N/A | in D$_2$O |

The measured amide-I band positions of the gas-phase molecules are compared to the band positions assigned to several different secondary structure elements of proteins in the condensed phase:
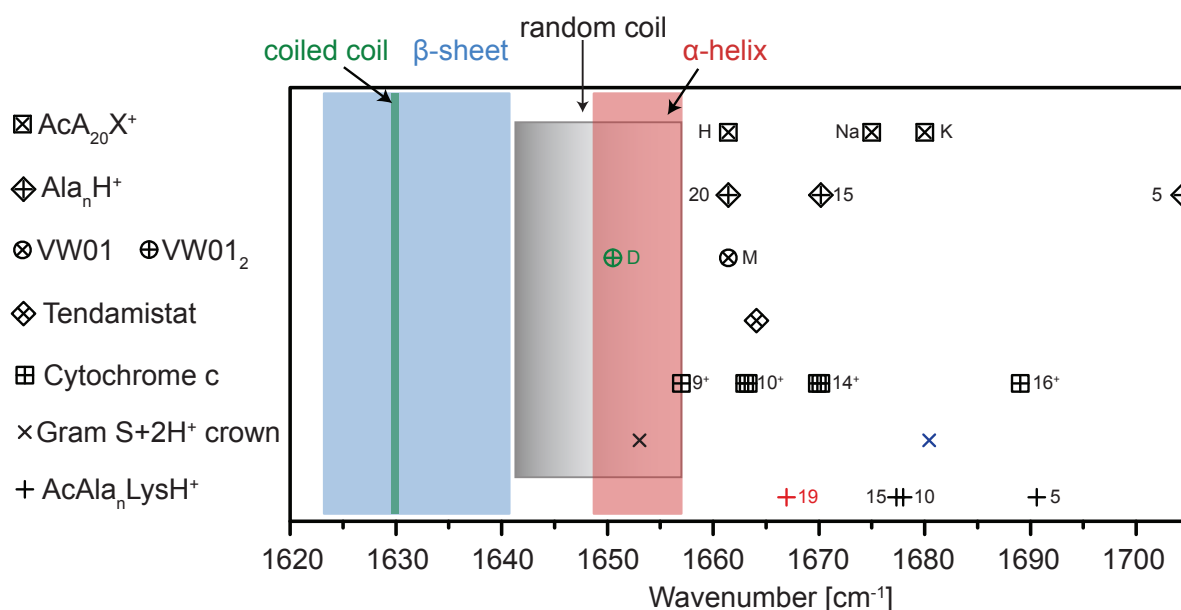
**Figure 7.1:** Graphical representation of the observed amide-I positions for some molecules investigated in this thesis. The colored areas mark the band positions of the secondary structural motif as found in the condensed phase: red → $\alpha$-helix, blue → $\beta$-sheet, green → *coiled coil*, and grey-shaded → random coil.

Polyalanine based peptides are investigated in the gas phase as potential $\alpha$-helical model systems. In previous experiments performed by other groups, it has been shown that polyalanine chains with a charged lysine residue next to the C-terminus are promising model systems for helical structured molecules in the gas phase.

The amide-I band position of a 19 alanine residue long peptide with a protonated lysine residue located next to the C-terminus is observed at 1668 cm$^{-1}$ and blue-shifted by about 14 cm$^{-1}$ with respect to the average band position found for $\alpha$-helical structures in the condensed phase. The measured value is also observed when the position of the protonated lysine residue is next to the N-terminus followed by 19 alanine residues. Ion mobility studies however, exclude the formation of an $\alpha$-helical motif for that particular molecule. On the other hand, the dimer $[(\mathrm{Ac-LysAla_{19}})_2 + 2\,\mathrm{H}]^{2+}$ is found to be in an $\alpha$-helical conformation in those studies. It is possible that instead of the monomer ion $\mathrm{Ac-LysAla_{19}} + \mathrm{H}^+$ the dimer ion was present in the infrared experiments, which could explain the identical observed values independently of the location of the lysine residue for both molecules.

Quantum chemical calculations performed on the systems $\mathrm{Ac-Ala_n Lys} + \mathrm{H}^+$ with n = 5, 10, 15, 19 underline the presence of a structure with an $\alpha$-helical character for n ≥ 10, with excellent agreement between the obtained theoretical spectra with the experimental ones. For n = 5, theoretical results suggest a more globular conformation, which is also indicated in the experimental spectrum: The amide-I mode is blue-shifted with respect to the observed position for the longer molecules, thus pointing to a weaker hydrogen bonding pattern and, as a consequence, a different conformation for the shortest molecule.

The amide-I band position is found to be blue-shifted for the short pure alanine residue peptide $Ala_5H^+$ with respect to longer alanine based peptides, for example $Ala_{20}H^+$, which also indicates a different structural motif between these two molecules. Ion mobility measurements suggest that on solely alanine residue based protonated peptides in the gas phase, at least up to this length, no $\alpha$-helical motif is present but rather a random globular or random coiled structure. If this is the case, the measured amide-I band positions for this molecules correspond to a random globular/coil of the backbone. Ion mobility measurements on alkali metal ion containing polyalanine chains show that the helicity of the complex is greatly enhanced by the metal ion with respect to the protonated counterpart. The amide-I band position observed for alanine peptides with an alkali ion attached is similar to that of $Ac-Ala_nLys + H^+$ with n = 10, 15, 19 thus possibly indicating an $\alpha$-helical motif for these molecules.

The observed amide-I band position for the doubly protonated gas-phase $\beta$-sheet model system gramicidin S complexed with crown ether molecules is observed at around 1680 cm$^{-1}$, blue-shifted in comparison to the position observed in the polyalanine peptides. It is found by calculations that solely protonated gramicidin S is not a good model system for a $\beta$-sheet structure in the gas phase due to strong interactions of the backbone C=O groups with charged amino acid side chains. These interactions from ornithine are reduced by capping the residue with a 18-crown-6 molecule. Two bands, appearing at 1633 cm$^{-1}$ (strong) and 1684 cm$^{-1}$ (weak) respectively, might indicate the presence of $\beta$-sheet structures of proteins in the condensed phase. In the infrared spectrum of capped gas-phase gramicidin S, in addition to the band observed at 1680 cm$^{-1}$ a intense band is present at 1653 cm$^{-1}$, blue-shifted with respect to the low frequency band observed in solution $\beta$-sheet structures. However, comparing theory with experiment shows that the band near 1680 cm$^{-1}$ mainly result from vibrations of four C=O groups involved in an antiparallel $\beta$-sheet motif while the band near 1653 cm$^{-1}$ results from vibrations of C=O groups present in the turn region of the cyclic peptide. This observation is in contrast to the band positions as found for that motif in the condensed phase and appears strongly blue-shifted.

It has been shown that the dimer of a model peptide artificially designed to possess a *coiled coil* structure can be transferred intact into the gas phase. Infrared spectroscopy on the protonated dimer as well as the protonated monomer were performed as part of the work of this thesis. The monomer is assumed to possess a random coil structure but is only observed in the gas phase. The dimer is designed such that a *coiled coil* structure is present in the condensed phase. The amide-I band position for this *coiled coil* model system in the gas phase is observed at 1651 cm$^{-1}$ and for the monomer at 1662 cm$^{-1}$. In the condensed phase, when $D_2O$ is used as a solvent, an amide-I band position at 1630 cm$^{-1}$ is assigned to a *coiled coil* structure of the molecule.

In the condensed phase, it is difficult to differentiate between a random coil structure and an $\alpha$-helical structure, as the amide-I band position is found at almost identical wavenumbers in the spectra of both motifs. In investigations on the model peptides VW01, VW02Abz, and VW03, which are designed to possess a random coil structure in the monomeric form as well as on $Ala_{20}H^+$, the amide-I band position is found at 1662

$cm^{-1}$, only slightly red-shifted compared to the position of the $\alpha$-helical model system $Ac{-}Ala_{19}Lys + H^+$.

Applying the results obtained from investigations of the above described model systems leads to the following conclusion:
In the spectra of charge state 7+ to 10+ of cytochrome c, the amide-I band position is observed at about 1663 $cm^{-1}$, thus suggesting a random coiled or $\alpha$-helical structure of the backbone. The blue-shift of the band observed for the higher charge states of cytochrome c is in agreement with a more elongated structural motif due to charge repulsion. The similarity of the amide-I band position observed in the protein tendamistat to the low charged cytochrome c molecules might also indicate a more random coiled structure of the molecule. It seems therefore, as if the $\beta$-sheet character of the backbone of tendamistat is not conserved in the gas phase at this charge state. That is not to say that it is impossible to conserve secondary structural elements in the gas phase, as has been demonstrated by the dimer of the model system VW01, a non-covalent bound protein complex.

In conclusion, it appears as if the spectral fingerprint for an $\alpha$-helical system in the gas phase is in the range of $\approx 1667$ $cm^{-1}$ to $\approx 1678$ $cm^{-1}$ of the amide-I band. The range, which can possibly assigned to a random coil conformation of the peptide begins slightly lower at $\approx 1662$ $cm^{-1}$ and ends at $\approx 1670$ $cm^{-1}$. Random globular structures can eventually be assigned to a amide-I band position of $\approx 1691$ $cm^{-1}$ to $\approx 1704$ $cm^{-1}$. It seems plausible that a distinction between random structures and $\alpha$-helical motifs is difficult in the gas phase as it is in the condensed phase. The strong blue-shift of the C=O vibrations involved in an antiparallel $\beta$-sheet motif in the gas phase with respect to the condensed phase might be caused by interactions with the environment. The amide-I band position value for the condensed phase is deduced for $\beta$-sheet sections that are embedded in larger proteins, that interact strongly with solvent or that are part of partially aggregated species. Therefore, to obtain the corresponding amide-I band position for gas-phase species it is likely that bigger complexes must be investigated to obtain a $\beta$-sheet structure with a strong spectral signature. A vibrational signature of the amide-I band at 1651 $cm^{-1}$ might correspond to a *coiled coil* structure of the molecule. It must be noted, that a *coiled coil* motif consists of two $\alpha$-helices, which are non-covalently bound. As a consequence, it might be that the measured gas-phase values are reflecting the helical character of the structure. However, the dimer is the biggest model system investigated in this thesis and the observed band positions indicate a stronger hydrogen bonding network compared to that of the other small molecules. It seems as if the amide-I band position observed in the gas-phase model systems is blue-shifted with respect to the corresponding condensed phase value for almost all structural motifs. The trend of the observed amide-I band positions for the different structural elements in the condensed phase is, on the other hand, reflected in the observations of the gas-phase model systems, if the assignment of the secondary structure elements of the molecules is correct. In the gas phase, not necessarily all C=O groups are involved in hydrogen bonding networks thus leading to the presence signatures of free C=O oscillators in the amide-I region. This might be additionally a reason, why the amide-I band appears often blue-shifted in the gas phase with respect to the condensed phase.

There are many challenges associated with applying those values to determine accurately the structure of the proteins investigated in this thesis. The molecules are produced by harsh electrospray ionization conditions such as a highly acidic solution, high temperatures of the desolvation gas, and high voltages of the entrance cone. As a consequence, high charge states are found in the gas phase and the conformations of those charge states presumably differ from the structures found in the condensed phase. This is due to Coulomb repulsion of the charges, with the consequential changes to more elongated structures. Thus, it is of importance to investigate the lower charge states that correspond more closely to the natural occurring charge state in the condensed phase and to compare those results with the values obtained here. Charge repulsion in general has a greater impact to structural folding in the gas phase than in the condensed phase. No solvent molecules can compensate for exposed charged groups by capping and thus shielding the charges. Different polar interactions together with the absence of bonding partners for hydrogen bonds in the gas phase with respect to the condensed phase leads to different structures especially for small molecules. The effect is not so distinct for macromolecules in contrast to small molecules, where the structure is mostly stabilized by interactions with the solvent molecules, and affects only the exposed groups. Furthermore, the dielectric constant $\epsilon = \epsilon_0 \cdot \epsilon_r$ with $\epsilon_0 = 8.85 \times 10^{-12} \mathrm{As/Vm}$ and the relative dielectric constant $\epsilon_r$. $\epsilon_r$ is 1 in vacuum and $\approx 80$ in water at $20°$. The difference in the dielectric constant between gas phase and condensed phase might lead in general to a relative shift of the peak positions for the vibrations of chemical groups due to the different environment.

Furthermore, in gas-phase investigations, more than one conformer of the protein can be present. The measured infrared spectra then correspond to a sum of all infrared spectra of the respective structures. An additional problem is the temperature of the investigated samples, which are thermalized to about 300 K, the ambient temperature. As a consequence, broad bands are measured due to the population of higher vibrational states than only the ground vibrational state of the molecule. Internal dynamics accompanying the relatively high temperature cause additional broadening of the vibrational modes.

To accurately assign amide-I band position to secondary structural elements in the gas phase, experimental conditions need to be improved and the conformational space of the molecular model system needs to be better controlled. In the next Section, several proposals are made as how the spectroscopic methods can be improved to possibly distinguish between those different structure elements.

## 7.2   Future perspectives

To allow for structural insights into large, complex biomolecular systems in the gas phase using infrared spectroscopy, further improvements of the spectroscopic techniques are needed and conditions such as temperature and charge state of the investigated systems need to be controllable and optimized. In this outlook, several techniques

to optimize structural investigation on gas-phase biomolecules are presented and potentially significant future research projects are introduced.

### 7.2.1   Experiments on cold molecules using helium nano-droplets

Increasing control over the temperature of the investigated molecules is of great interest. It has been shown recently, that high resolution spectra of gas-phase biomolecules can be obtained by spectroscopic investigations on cold biomolecules [48] allowing structural assignment of investigated molecules. Two spectra of the cyclic decapeptide gramicidin S are shown in Figure 7.2. An infrared spectrum of the molecule at room temperature obtained by IR-MPD is shown on the left hand side and the spectra of the same molecule in an ion trap kept at 6 K is shown on the right.
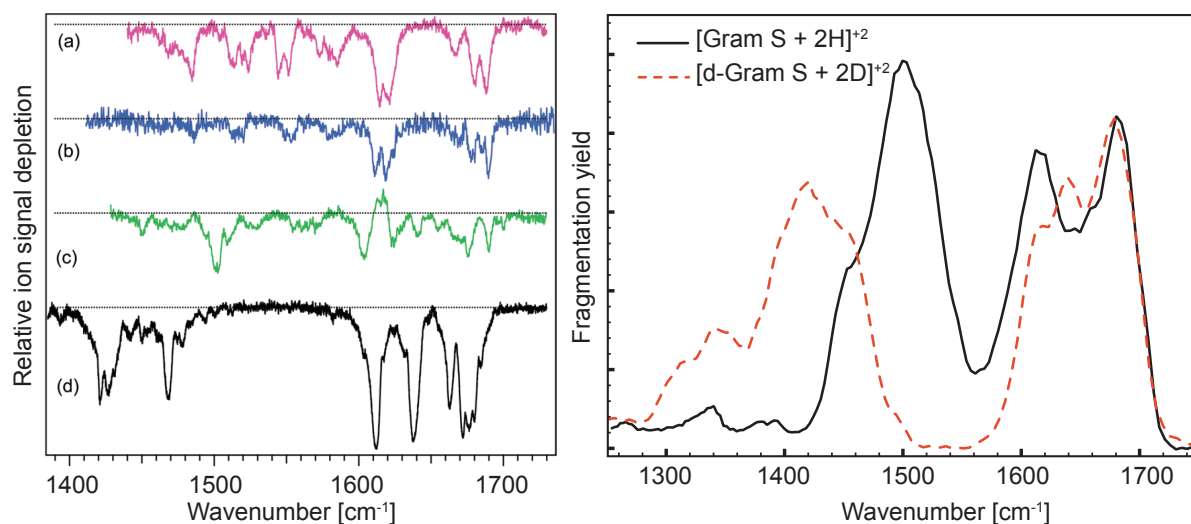


**Figure 7.2:** Left: Experimental IR-UV spectra of $[GramS + 2H]^{2+}$ at 6 K. Shown are spectra of different conformers, details can be found in the article by Nagornova *et al.* [48]. The spectrum is shown with permission of the authors. Right: IR-MPD spectrum of $[GramS + 2H]^{2+}$ and $[d-GramS + 2D]^{2+}$ at room temperature, see Chapter 5 for details

The comparison clearly shows that more structural information can be obtained by performing infrared spectroscopy on cold molecular systems. Different than using collisions of the molecule with cold low density helium gas where thermalization occurs over a relatively long time scale as performed in the experiments by Nagornova *et al.*, it is also possible to achieve cooling by deposition of the molecule inside a helium nano-droplet. The ion thermalizes rapidly due to the close contact with the helium environment, which is kept at 0.37 K by evaporative cooling. The helium itself is only weakly interacting with the doped ion and is optically transparent from the deep ultraviolet to the far infrared region of the electromagnetic spectrum, providing an ideal matrix for spectroscopic experiments [284, 285]. Recently, it has been shown that even proteins can incorporated into helium droplets [286]. The setup for this experiment is shown in Figure 7.3.
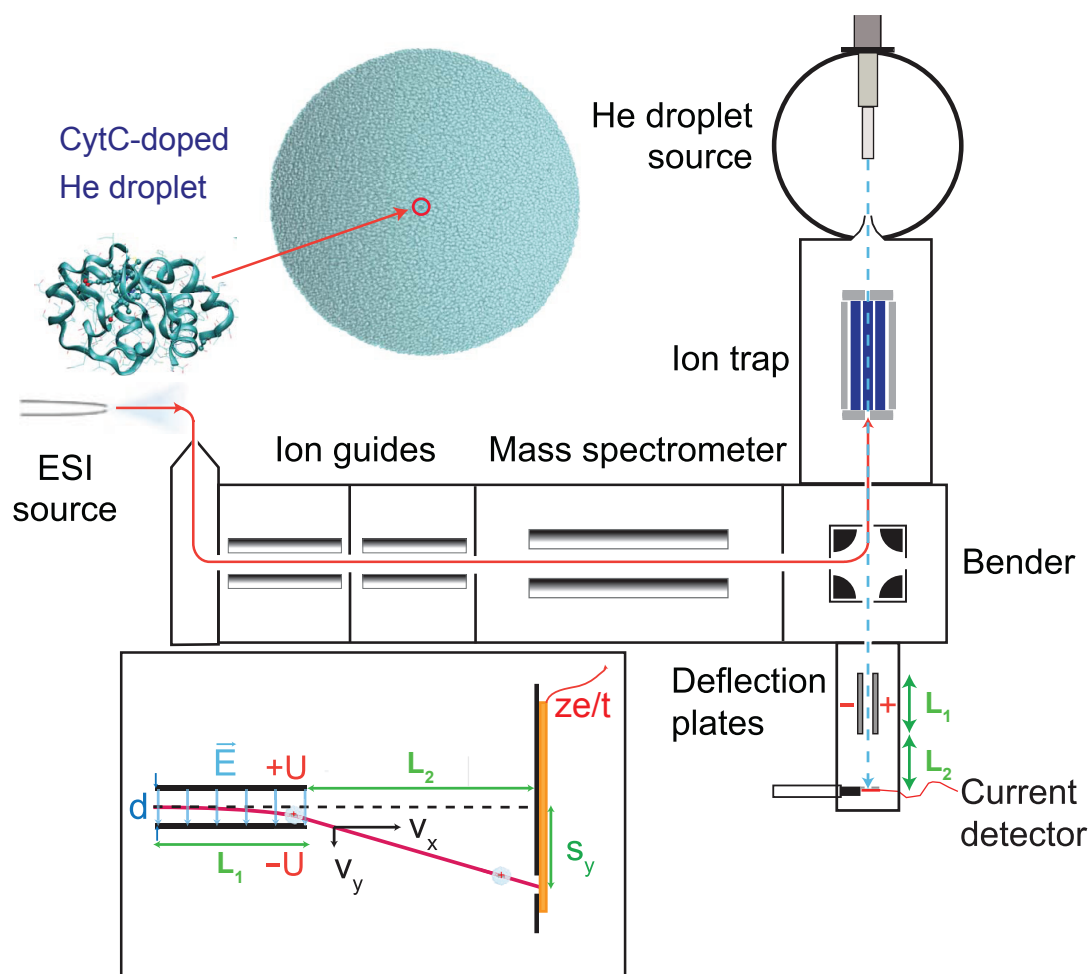
**Figure 7.3:** Shown is a scheme of the current experimental setup for embedding gas-phase biomolecular ions into helium nano-droplets with the possibility to deflect the charged cluster to allow for determination of the helium nano-droplet mass. Further details are given in the article by Bierau *et al.* [286].

In this experiment, ions are produced via electrospray ionization, mass-to-charge selected in a quadrupole, and trapped in a linear hexapole ion trap. Helium droplets are generated by expansion of helium gas at a pressure between 10 bar and 50 bar via a pulsed valve kept between 6 K and 20 K at 4 Hz following a design by Vilesov *et al.* [287]. The helium nano-droplets produced this way pass through the ion-filled trap and can capture ions. Due to the high kinetic energy of the droplets, ions can leave the trap when inside the droplets. After one minute, the trap content is depleted by about 50% and the helium nano-droplet source is switched off to allow for reloading the linear hexapole with fresh ions; the pick up is then repeated. Ion doped helium droplets are measured with an ion detector, which is similar to a Faraday cage and is mounted about 180 cm below the trap exit. Time of flight profiles between the opening of the pulsed valve and the measurement of current at the detector can be measured as a function of helium nano-droplet source parameters as shown in Figure 7.4. Details are given in the published article and a highly extensive introduction on the incorporation of ions into helium nano-droplets will be given in the doctoral thesis of Frauke Bierau [288]. Further

infrared spectroscopy experiments on cold ions embedded in helium nano-droplets are planned.
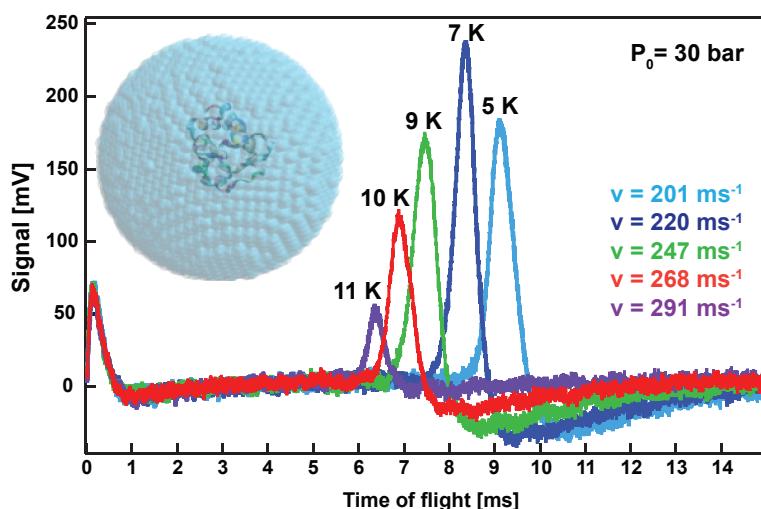


**Figure 7.4:** Time of flight distribution of protonated cytochrome c (q = 14) doped helium nano-droplets. The pressure of the helium behind the pulsed valve is 30 bar, the temperature of the valve varies between 5 K and 11 K as indicated in the Figure. With increasing temperature, the doped nano-droplets get faster. A considerable portion of the peak width can be attributed to the time response of the current amplifier.

## 7.2.2   Conformer preselection with the ion mobility technique

Often different conformers are present in the gas phase. Ion mobility experiments allow for determination of the average size of an ion by measuring its drift velocity through an inert buffer gas filled cell by applying a weak electric field, which pulls the molecule through that buffer. The drift time depends on the cross section of the ion, thus providing an insight into its overall structure. When different conformers with different overall structures are present, this method additionally allows for temporal and spatial separation of different structural families and therefore the preselection of conformers. It is possible to combine this ion mobility preselector with an experimental setup to perform infrared spectroscopy on structure separated samples, as shown in Figure 7.5.

In a setup currently being constructed at the FHI, ions are produced by nano electrospray. Before the ions can enter the high vacuum regime of the apparatus, they have to pass the drift region with a pressure of about 10 mbar – 100 mbar (usually rare gas, such as helium). Ions are injected into this region via a radio frequency (RF) ion funnel, which allows for pre-accumulation of the molecules and for the generation of short ion packets. The drift tube itself consists of coated conductive glass tubes facilitating the application of a constant voltage gradient along the axis. The tube length, determining the resolution of the separation, is in the order of 1 m but is easily expandable. At the end of the drift region, an RF ion funnel bunches the ions, which are spread out by diffusion, and transfers them into the high vacuum regime of the apparatus. By applying temporally appropriate voltages to, for example, the quadrupole bender, it
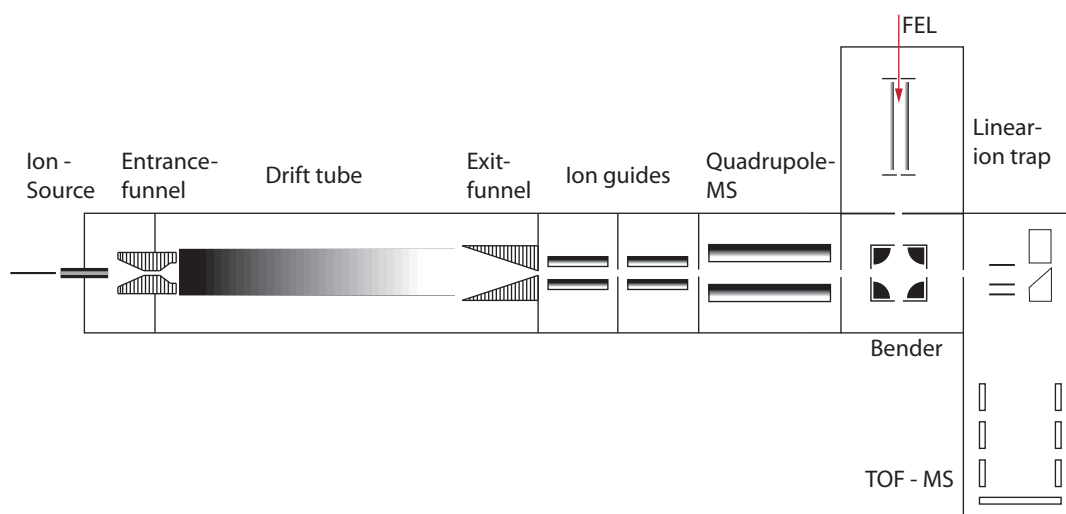
**Figure 7.5:** Proposed setup for combining ion mobility separation with (infrared) spectroscopy. For further details, see text.

is possible to inject only one structural and mass selected family into the trap for the purpose of performing spectroscopic investigations. By combining this method with the helium nano-droplet source, it should be possible to perform experiments on cold, structurally preselected molecules.

## 7.2.3 Soft ionization methods

For structural investigations of gas-phase biomolecules, it is critical that soft ionization techniques are used for the production of the ions in order to conserve subtle structural features of the condensed-phase conformation. Furthermore, trying to reproduce the charge state of the molecule that it would possess in its natural environment, the condensed phase, could be helpful to study relevant structures in the gas phase.

The electrospray conditions of the FT ICR apparatus in the Netherlands (see Chapter 1, Section 1.4.4) are rather harsh and can lead to breakdowns of the native structure. A promising alternative would be the use of nano electrospray ionization techniques. One advantage is that only very small sample volumes are required (ml versus $\mu$l), allowing for the investigation of precious protein samples. Secondly, the use of solvents to lower the surface tension as required in the conventional electrospray process can be avoided. The solvents such as methanol can influence the protein structure already in solution prior to investigation, which could lead to denatured conformations of the molecule. Finally, the desolvation process with the nano electrospray technique has been found to be more gentle and efficient for ionizing structurally intact biomolecular complexes [27] compared with conventional electrospray; as an example, no heated desolvation gas is needed. Thus, coupling a nano electrospray source to the existing setup could improve the production of structurally intact biomolecular ions for spectroscopic observations.

As pointed out in Chapter 3, it is of importance to investigate lower charge states of

proteins compared to the investigated charges states presented in this thesis. Charge reduction of proteins can be achieved by selective, non-covalent attachment and subsequent detachment of crown-ether molecules with high proton affinity to the sample [289].

All the proposed techniques in this Section on how to improve vibrational spectroscopy on gas-phase biomolecules are in preparation and a future opens up, full with exciting experiments for structural investigations on systems with biological relevance.

# Bibliography

[1] A. Miller and J. Tanner. *Essentials of Chemical Biology - Structure and Dynamics of Biological Macromolecules*. John Wiley & Sons Ltd, England (2008).

[2] http://www.britannica.com/EBchecked/topic/62958/Jons-Jacob-Berzelius.

[3] J. C. Kendrew. Myoglobin and the structure of proteins. nobelprize.org (1962).

[4] K. Wüthrich. NMR Studies of Structure and Function of Biological Macromolecules. nobelprize.org (2002).

[5] A. Barth and C. Zscherp. What vibrations tell us about proteins. *Quart. Rev. Biophys.*, **35**, 369–430 (2002).

[6] G. Fasman. *Circular Dichroism and the Conformational Analysis of Biomolecules*. Plenum Press, New Yorck (1996).

[7] R. B. Corey and L. Pauling. Molecular Models of Amino Acids, Peptides, and Proteins. *Rev. Sci. Instrum.*, **24**, 621–627 (1953).

[8] D. P. Barondeau, C. J. Kassmann, J. A. Tainer, and E. D. Getzoff. The case of the missing ring: Radical cleavage of a carbon-carbon bond and implications for GFP chromophore biosynthesis. *J. Am. Chem. Soc.*, **129**, 3118–3126 (2007).

[9] C. Tanford. The hydrophobic effect and the organization of living matter. *Science*, **200**, 1012–1018 (1978).

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

[11] D. Voet, J. G. Voet, and C. W. Pratt. *Lehrbuch der Biochemie*. Wiley-VCH Verlag GmbH & Co. KGaA (2002).

[12] J. S. Richardson. The Anatomy and Taxonomy of Protein Structure. http://kinemage.biochem.duke.edu/teaching/anatax/ (1981).

[13] J. N. Onuchic, Z. Luthey-Schulten, and P. G. Wolynes. Theory of Protein Folding: The Energy Landscape Perspective. *Annu. Rev. Phys. Chem.*, **48**, 545–600 (1997).

[14] J. N. Onuchic and P. G. Wolynes. Theory of protein folding. *Curr. Opin. Struct. Biol.*, **14**, 70–75 (2004).

[15] R. Jaenicke. Folding and association of proteins. *Prog. Biophys. Mol. Biol.*, **49**, 117–237 (1987).

[16] L. Pauling, R. B. Corey, and H. R. Branson. The structure of proteins: two hydrogen-bonded helical configurations of the polypeptide chain. *Proc. Natl. Acad. Sci.*, **37**, 205–211 (1951).

[17] L. Pauling and R. B. Corey. Configurations of polypeptide chains with favored orientations around single bonds: two new pleated sheets. *Proc. Natl. Acad. Sci.*, **37**, 729–740 (1951).

[18] kindly provided by M. Rossi.

[19] kindly provided by K. Pagel.

[20] kindly provided by G. von Helden.

[21] W. G. J. Hol. The role of the $\alpha$-helix dipole in protein function and structure. *Prog. Biophys. Mol. Biol.*, **45**, 149–195 (1985).

[22] J. T. Yang. Optical rotatory dispersion of polypeptides and proteins. *Tetrahedron*, **13**, 143–165 (1961).

[23] C. Krittanai and W. C. Johnson. The Relative Order of Helical Propensity of Amino Acids Changes With Solvent Environment. *Proteins: Struct., Funct., Genet.*, **39**, 132–141 (2000).

[24] F. H. C. Crick. The Packing of $\alpha$-Helices: Simple Coiled-Coils. *Acta Crystallogr.*, **6**, 689–697 (1953).

[25] B. T. Ruotolo and C. V. Robinson. Aspects of native proteins are retained in vacuum. *Curr. Opin. Chem. Biol.*, **10**, 402–408 (2006).

[26] Q. Y. Wu, S. van Orden, X. H. Cheng, R. Bakhtiar, and R. D. Smith. Characterization of Cytochrome c Variants with High-Resolution FTICR Mass Spectrometry: Correlation of Fragmentation and Structure. *Anal. Chem.*, **67**, 2498–2509 (1995).

[27] J. L. P. Benesch, B. T. Ruotolo, D. A. Simmons, and C. V. Robinson. Protein complexes in the gas phase: Technology for structural genomics and proteomics. *Chem. Rev.*, **107**, 3544–3567 (2007).

[28] M. Rossi. *Ab initio study of secondary structure formation in gas-phase peptides*. Ph.D. thesis, Fritz-Haber-Institut der Max-Planck-Gesellschaft (2011).

[29] M. Yamashita and J. B. Fenn. Electrospray Ion-Source. Another Variation on the Free-Jet Theme. *J. Phys. Chem.*, **88**, 4451–4459 (1984).

[30] S. Chapman. Carrier mobility spectra of spray electrified liquids. *Phys. Rev.*, **52**, 184–190 (1937).

[31] M. Dole, L. L. Mack, R. L. Hines, R. C. Mobley, L. D. Ferguson, and M. B. Alice. Molecular Beams of Macroions. *J. Chem. Phys.*, **49**, 2240–2249 (1968).

[32] J. B. Fenn, M. Mann, C. K. Meng, S. F. Wong, and C. M. Whitehouse. Electrospray Ionization for Mass-Spectrometry of Large Biomolecules. *Science*, **246**, 64–71 (1989).

[33] J. B. Fenn. Electrospray Wings for Molecular Elephants. `nobelprize.org` (2002).

[34] L. Rayleigh. On the equilibrium of liquid conducting masses charged with electricity. *Philos. Mag.*, **14**, 184 – 186 (1882).

[35] P. Kebarle. A brief overview of the present status of the mechanisms involved in electrospray mass spectrometry. *J. Mass Spectrom.*, **35**, 804–817 (2000).

[36] P. Kebarle and U. H. Verkerk. Electrospray: from ions in solution to ions in the gas phase, what we know now. *Mass Spectrom. Rev.*, **28**, 898–917 (2009).

[37] M. Karas and F. Hillenkamp. Laser Desorption Ionization of Proteins with Molecular Masses Exceeding 10000 Daltons. *Anal. Chem.*, **60**, 2299–2301 (1988).

[38] K. Tanaka. The Origin of Macromolecule Ionization by Laser Irradiation. `nobelprize.org` (2002).

[39] P. Morse. Diatomic molecules according to the wave mechanics. II. Vibrational levels. *Phys. Rev.*, **34**, 57–64 (1929).

[40] N. B. Colthup, L. H. Daly, and S. E. Wiberley. *Introduction to Infrared and Raman Spectroscopy*. Academic Press, London, UK (1990).

[41] W. C. Reisdorf, Jr. and S. Krimm. Infrared Amide I' Band of the Coiled Coil. *Biochemistry*, **35**, 1383–1386 (1996).

[42] A. Barth. Infrared spectroscopy of proteins. *Biochim. Biophys. Acta*, **1767**, 1073–1101 (2007).

[43] N. C. Polfer and J. Oomens. Vibrational Spectroscopy of Bare and Solvated Ionic Complexes of Biological Relevance. *Mass Spectrom Rev.*, **28**, 468–494 (2009).

[44] M. A. Duncan. Frontiers in the spectroscopy of mass-selected molecular ions. *Int. J. Mass Spectrom.*, **200**, 545–569 (2000).

[45] E. J. Bieske and O. Dopfer. High-resolution spectroscopy of cluster ions. *Chem. Rev.*, **100**, 3963–3998 (2000).

[46] T. Rizzo, Y. Park, L. Peteanu, and D. Levy. The Electronic Spectrum of the Amino Acid Tryptophan in the Gas Phase. *J. Chem. Phys.*, **84**, 2534–2541 (1986).

[47] A. Abo-Riziq, B. O. Crews, M. P. Callahan, L. Grace, and M. S. de Vries. Spectroscopy of isolated gramicidin peptides. *Angew. Chem. - Int. Ed.*, **45**, 5166–5169 (2006).

[48] N. Nagornova, T. Rizzo, and O. Boyarkin. Highly Resolved Spectra of Gas-Phase Gramicidin S: A Benchmark for Peptide Structure Calculations. *J. Am. Chem. Soc.*, **132**, 4040–4041 (2010).

[49] M. Gerhards, C. Unterberg, and A. Gerlach. Structure of a beta-sheet model system in the gas phase: Analysis of the C=O stretching vibrations. *Phys. Chem. Chem. Phys.*, **4**, 5563–5565 (2002).

[50] M. Mons, I. Dimicoli, F. Piuzzi, B. Tardivel, and M. Elhanine. Tautomerism of the DNA Base Guanine and Its Methylated Derivatives as Studied by Gas-Phase Infrared and Ultraviolet Spectroscopy. *J Phys Chem A*, **106**, 5088–5094 (2002).

[51] M. Putter, G. von Helden, and G. Meijer. Mass selective infrared spectroscopy using a free electron laser. *Chem. Phys. Lett.*, **258**, 118–122 (1996).

[52] G. von Helden, I. Holleman, G. M. H. Knippels, A. F. G. van der Meer, and G. Meijer. Infrared resonance enhanced multiphoton ionization of fullerenes. *Phys. Rev. Lett.*, **79**, 5234–5237 (1997).

[53] A. Fielicke, A. Kirilyuk, C. Ratsch, J. Behler, M. Scheffler, G. von Helden, and G. Meijer. Structure Determination of Isolated Metal Clusters via Far-Infrared Spectroscopy. *Phys. Rev. Lett.*, **93**, 023401 (2004).

[54] K. R. Asmis, M. Brümmer, C. Kaposta, G. Santambrogio, G. von Helden, G. Meijer, K. Rademann, and L. Wöste. Mass-selected infrared photodissociation spectroscopy of $V_4O_{10}^+$. *Phys. Chem. Chem. Phys.*, **4**, 1101–1104 (2002).

[55] J. Bakker, L. Mac Aleese, G. Meijer, and G. von Helden. Fingerprint IR spectroscopy to probe amino acid conformations in the gas phase. *Phys. Rev. Lett.*, **91**, 203003 (2003).

[56] J. Lemaire, P. Boissel, M. Heninger, G. Mauclaire, G. Bellec, H. Mestdagh, A. Simon, S. Le Caer, J. M. Ortega, F. Glotin, and P. Maitre. Gas Phase Infrared Spectroscopy of Selectively Prepared Ions. *Phys. Rev. Lett.*, **89**, 273002 (2002).

[57] J. J. Valle, J. R. Eyler, J. Oomens, D. T. Moore, A. F. G. van der Meer, G. von Helden, G. Meijer, C. Hendrickson, A. Marshall, and G. Blakney. Free electron laser-Fourier transform ion cyclotron resonance mass spectrometry facility for obtaining infrared multiphoton dissociation spectra of gaseous ions. *Rev. Sci. Instrum.*, **76**, 023103 (2005).

[58] J. Oomens, B. G. Sartakov, G. Meijer, and G. von Helden. Gas-phase infrared multiple photon dissociation spectroscopy of mass-selected molecular ions. *Int. J. Mass spectrom.*, **254**, 1–19 (2006).

[59] J. G. Black, E. Yablonovitch, N. Bloembergen, and S. Mukamel. Collisionless Multiphoton Dissociation of $SF_6$: Statistical Thermodynamic Process. *Phys. Rev. Lett.*, **38**, 1131–1134 (1977).

[60] M. Okumura, L. I. Yeh, and Y. T. Lee. The vibrational predissociation spectroscopy of hydrogen cluster ions. *J. Chem. Phys.*, **83**, 3705–3706 (1985).

[61] J. Oomens, A. J. A. van Roij, G. Meijer, and G. von Helden. Gas-phase infrared photodissociation spectroscopy of cationic polyaromatic hydrocarbons. *Astrophys. J.*, **542**, 404–410 (2000).

[62] M. Okumura, L. I. Yeh, J. D. Myers, and Y. T. Lee. Infrared spectra of the cluster ions $H_7O_3^+ \cdot H_2$ and $H_9O_4^+ \cdot H_2$. *J. Chem. Phys.*, **85**, 2328–2329 (1986).

[63] H. P. Freund and T. M. Antonson, Jr. *Principles of Free electron Lasers*. Chapman & Hall, London, second edition (1996).

[64] kindly provided by J. Oomens.

[65] D. Oepts, A. F. G. van der Meer, and P. W. van Amersfoort. The Free-Electron-Laser User Facility Felix. *Infrared Phys. Technol.*, **36**, 297–308 (1995).

[66] A. G. Marshall, C. L. Hendrickson, and G. S. Jackson. Fourier transform ion cyclotron resonance mass spectrometry: A primer. *Mass Spectrom. Rev.*, **17**, 1–35 (1998).

[67] A. G. Marshall, T.-C. L. Wang, and T. L. Ricca. Tailored Excitation for Fourier Transform Ion Cyclotron Resonance Mass Spectrometry. *J. Am. Chem. Soc.*, **107**, 7893–7897 (1985).

[68] http://magnet.fsu.edu/~midas/.

[69] T. H. Mize, I. Taban, M. Duursma, M. Seynen, M. Konijnenburg, A. Vijftigschild, C. V. Doornik, G. V. Rooij, and R. M. A. Heeren. A modular data and control system to improve sensitivity, selectivity, speed of analysis, ease of use, and transient duration in an external source FTICR-MS. *Int. J. Mass Spectrom.*, **235**, 243–253 (2004).

[70] R. C. Haddon, A. F. Hebard, M. J. Rosseinsky, D. W. Murphy, S. J. Duclos, K. B. Lyons, B. Miller, J. M. Rosamilia, R. M. Fleming, A. R. Kortan, S. H. Glarum, A. V. Makhija, A. J. Muller, R. H. Eick, S. M. Zahurak, R. Tycko, G. Dabbagh, and F. Thiel. Conducting films of $C_{60}$ and $C_{70}$ by alkali-metal doping. *Nature*, **350**, 320–322 (1991).

[71] A. F. Hebard, M. J. Rosseinsky, R. C. Haddon, D. W. Murphy, S. H. Glarum, T. T. M. Palstra, A. P. Ramirez, and A. R. Kortan. Superconductivity at 18 K in potassium-doped $C_{60}$. *Nature*, **350**, 600–601 (1991).

[72] K. Holczer, O. Klein, S. M. Huang, R. B. Kaner, K. J. Fu, R. L. Whetten, and F. Diederich. Alkali-fulleride superconductors: synthesis, composition, and diamagnetic shielding. *Science*, **252**, 1154–1157 (1991).

[73] M. J. Rosseinsky, A. P. Ramirez, S. H. Glarum, D. W. Murphy, R. C. Haddon, A. F. Hebard, T. T. M. Palstra, A. R. Kortan, S. M. Zahurak, and A. V. Makhija. Superconductivity at 28 K in $Rb_xC_{60}$. *Phys. Rev. Lett.*, **66**, 2830–2832 (1991).

[74] A. Y. Ganin, Y. Takabayashi, Y. Z. Khimyak, S. Margadonna, A. Tamai, M. J. Rosseinsky, and K. Prassides. Bulk superconductivity at 38 K in a molecular system. *Nature*, **7**, 367–371 (2008).

[75] C. C. Chancey and M. C. M. O'Brien. *The Jahn-Teller Effect in $C_{60}$ and Other Icosahedral Complexes*. Princeton University Press, Princeton, NJ (1997).

[76] F. Negri, G. Orlandi, and F. Zerbetto. Quantum Chemical Investigation of Franck-Condon and Jahn Teller Activity in the Electronic Spectra of Buckminsterfullerene. *Chem. Phys. Lett.*, **144**, 31–37 (1988).

[77] N. Koga and K. Morokuma. Ab initio MO study of the $C_{60}$ anion radical: the Jahn-Teller distortion and electronic structure. *Chem. Phys. Lett.*, **196**, 191–196 (1992).

[78] W. H. Green, S. M. Gorun, G. Fitzgerald, P. W. Fowler, A. Ceulemans, and B. C. Titeca. Electronic structures and geometries of $C_{60}$ anions via density functional calculations. *J. Phys. Chem.*, **100**, 14892–14898 (1996).

[79] D. R. Lawson, D. L. Feldheim, C. A. Foss, P. K. Dorhout, C. M. Elliott, C. R. Martin, and B. Parkinson. Near-IR Absorption Spectra for the Buckminsterfullerene Anions - an Experimental and Theoretical Study. *J. Electrochem. Soc.*, **139**, L68–L71 (1992).

[80] O. Gunnarsson, H. Handschuh, P. S. Bechthold, B. Kessler, G. Gantefor, and W. Eberhardt. Photoemission spectra of $C_{60}^-$ - Electron-phonon coupling, Jahn-Teller effect, and superconductivity in the fullerides. *Phys. Rev. Lett.*, **74**, 1875–1878 (1995).

[81] V. C. Long, J. L. Musfeldt, K. Kamáras, A. Schilder, and W. Schütz. Far-infrared study of the Jahn-Teller-distorted $C_{60}$ monoanion in $C_{60}$-tetraphenylphosphoniumiodide. *Phys. Rev. B*, **58**, 14338–14348 (1998).

[82] S. Tomita, J. U. Andersen, E. Bonderup, P. Hvelplund, B. Liu, S. Brøndsted Nielsen, U. V. Pedersen, J. Rangama, K. Hansen, and O. Echt. Dynamic Jahn-Teller effects in isolated $C_{60}^-$ studied by near-infrared spectroscopy in a storage ring. *Phys. Rev. Lett.*, **94**, 53002 (2005).

[83] G. Klupp, K. Kamáras, N. M. Nemes, C. M. Brown, and J. Leão. Static and dynamic Jahn-Teller effect in the alkali metal fulleride salts A(4)C(60) (A = K, Rb, Cs). *Phys. Rev. B*, **73**, 085415 (2006).

[84] V. C. Long, E. C. Schundler, G. B. Adams, J. B. Page, W. Bietsch, and I. Bauer. Jahn-Teller distortion of $C_{60}^-$ in $(Ph_4As)_2ClC_{60}$: $C_{2h}$ versus $C_i$ symmetry. *Phys. Rev. B*, **75**, 125402 (2007).

[85] T. Pichler, R. Winkler, and H. Kuzmany. Equilibrium phases in K-doped and Rb-doped $C_{60}$ from *in situ* infrared reflectivity measurements. *Phys. Rev. B*, **49**, 15879–15889 (1994).

[86] P.-M. Allemand, G. Srdanov, A. Koch, K. Khemani, F. Wudl, Y. Rubin, F. Diederich, M. M. Alvarez, S. Anz, and R. L. Whetten. The unusual electron spin resonance of fullerene $C_{60}^-$. *J. Am. Chem. Soc.*, **113**, 2780–2781 (1991).

[87] A. Pénicaud, A. Peréz-Benítez, R. Gleason V., E. Muñoz P., and R. Escudero. Electrocrystallizing $C_{60}$ - Synthesis, Single-Crystal X-Ray Structure, And Magnetic (Esr, Squid) Characterization Of $[(C_6H_5)_4P]_2[C_{60}][I]_X$. *J. Am. Chem. Soc.*, **115**, 10392–10393 (1993).

[88] P. A. Limbach, L. Schweikhard, K. A. Cowen, M. T. McDermott, A. G. Marshall, and J. V. Coe. Observation of the doubly charged gasphase fullerene anions $C_{60}^{2-}$ and $C_{70}^{2-}$. *J. Am. Chem. Soc.*, **113**, 6795–6798 (1991).

[89] O. Hampe, M. Neumaier, M. N. Blom, and M. M. Kappes. On the generation and stability of isolated double negatively charged fullerenes. *Chem. Phys. Lett.*, **354**, 303–309 (2002).

[90] G. von Helden, I. Holleman, G. Meijer, and B. Sartakov. Excitation of $C_{60}$ using a chirped free electron laser. *Opt. Expr.*, **4**, 46–52 (1999).

[91] A. Bekkerman, E. Kolodney, G. von Helden, B. Sartakov, D. van Heijnsbergen, and G. Meijer. Infrared multiphoton ionization of superhot $C_{60}$: Experiment and model calculations. *J. Chem. Phys.*, **124**, 184312 (2006).

[92] H. Ajie, M. M. Alvarez, S. J. Anz, R. D. Beck, F. Diederich, K. Fostiropoulos, D. R. Huffman, W. Krätschmer, Y. Rubin, K. E. Schriver, D. Sensharma, and R. L. Whetten. Characterization of the Soluble All-Carbon Molecules $C_{60}$ and $C_{70}$. *J. Phys. Chem.*, **94**, 8630–8633 (1990).

[93] D. S. Bethune, G. Meijer, W. C. Tang, and H. J. Rosen. The vibrational Raman spectra of purified solid films of $C_{60}$ and $C_{70}$. *Chem. Phys. Lett.*, **174**, 219–222 (1990).

[94] F. Cataldo, S. Iglesias-Groth, and A. Manchado. Low and High Temperature Infrared Spectroscopy of $C_{60}$ and $C_{70}$ Fullerenes. *Fullerenes, Nanotubes, Carbon Nanostruct.*, **18**, 224–235 (2010).

[95] D. S. Bethune, G. Meijer, W. C. Tang, H. J. Rosen, W. G. Golden, H. Seki, C. A. Brown, and M. S. de Vries. Vibrational raman and infrared-spectra of chromatographically separated $C_{60}$ and $C_{70}$ fullerene clusters. *Chem. Phys. Lett.*, **179**, 181–186 (1991).

[96] R. A. Jishi, M. S. Dresselhaus, G. Dresselhaus, K. A. Wang, P. Zhou, A. M. Rao, and P. C. Eklund. Vibrational mode frequencies in $C_{70}$. *Chem. Phys. Lett.*, **206**, 187–192 (1993).

[97] W. Zhao, L. Q. Chen, Y. X. Li, N. Zhao, Y. Z. Huang, Z. X. Zhang, H. T. Wang, P. X. Ye, and Z. X. Zhao. Spectroscopic studies of the interaction of $C_{60}$ and $C_{70}$ films with metal substrates. *Spectrochim Acta A*, **50**, 1759–1767 (1994).

[98] J. M. Treubig and P. R. Brown. Analysis of $C_{60}$ and $C_{70}$ fullerenes using high-performance liquid chromatography-Fourier transform infrared spectroscopy. *J. Chromatogr., A*, **960**, 135–142 (2002).

[99] J. Fulara, M. Jakobi, and J. P. Maier. Electronic spectra of the $C_{70}$ molecule and $C_{70}^{+}$, $C_{70}^{-}$ ions in neon matrices. *Chem. Phys. Lett.*, **206**, 203–209 (1993).

[100] H. Hase and Y. Miyatake. Electronic spectra of $C_{70}$ anions produced in $\gamma$-irradiated organic glasses at 77 K. *Chem. Phys. Lett.*, **215**, 141–143 (1993).

[101] T. Kodama, T. Kato, T. Moriwaki, H. Shiromaru, and Y. Achiba. Laser Study on the Resonance-Enhanced Multiphoton Electron Detachment (REMPED) Processes for $C_{60}^{-}$ and $C_{70}^{-}$. *J. Phys. Chem.*, **98**, 10671–10673 (1994).

[102] V. N. Semkin, N. V. Drichko, A. V. Talysin, A. Graja, S. Krol, D. N. Konarev, and R. N. Lyubovskaya. Infrared and visible spectra of the complexes of fullerene-70 and its salt. *Synth. Met.*, **93**, 207–212 (1998).

[103] X. B. Wang, H. K. Woo, X. Huang, M. M. Kappes, and L. S. Wang. Direct Experimental Probe of the On-Site Coulomb Repulsion in the Doubly Charged Fullerene Anion $C_{70}^{2-}$. *Phys. Rev. Lett.*, **96**, 143002 (2006).

[104] B. Concina, M. Neumaier, O. Hampe, and M. M. Kappes. Photodetachment of fullerene monoanions and dianions in a Penning trap: Probes of delayed electron emission and associated activation barriers. *Int. J. Mass. Spectrom.*, **252**, 110–116 (2006).

[105] N. C. Polfer and J. Oomens. Reaction products in mass spectrometry elucidated with infrared spectroscopy. *Phys. Chem. Chem. Phys.*, **9**, 3804–3817 (2007).

[106] R. N. Rosenfeld, J. M. Jasinski, and J. I. Brauman. Infrared multi-photon electron detachment from the benzyl anion. *J. Chem. Phys.*, **71**, 1030–1031 (1979).

[107] J. D. Steill and J. Oomens. Action Spectroscopy of Gas-Phase Carboxylated Anions by Multiple Photon IR Electron Detachment/Attachment. *J. Phys. Chem. A*, **113**, 4941–4964 (2009).

[108] R. Ahlrichs, M. Bär, M. Häser, H. Horn, and C. Kölmel. Electronic Structure Calculations on Workstation Computers: The Program System TURBOMOLE. *Chem. Phys. Lett.*, **162**, 165–169 (1989).

[109] O. Treutler and R. Ahlrichs. Efficient Molecular Numerical Integration Schemes. *J. Chem. Phys.*, **102**, 346–354 (1995).

[110] K. Eichkorn, F. Weigend, O. Treutler, and R. Ahlrichs. Auxiliary basis sets for main row atoms and transition metals and their use to approximate Coulomb potentials. *Theor. Chem. Acc.*, **97**, 119–124 (1997).

[111] A. Schäfer, H. Horn, and R. Ahlrichs. Fully Optimized Contracted Gaussian Basis Sets for Atoms Li to Kr. *J. Chem. Phys.*, **97**, 2571–2577 (1992).

[112] V. Schettino, M. Pagliai, and G. Cardini. The infrared and Raman spectra of fullerene $C_{70}$. DFT calculations and correlation with $C_{60}$. *J. Phys. Chem. A*, **106**, 1815–1823 (2002).

[113] W. Krätschmer, K. Fostiropoulos, and D. R. Huffman. The infrared and ultraviolet-absorption spectra of laboratory-produced carbon dust: evidence for the presence of the $C_{60}$ molecule. *Chem. Phys. Lett.*, **170**, 167–170 (1990).

[114] L. Nemes, R. S. Ram, P. F. Bernath, F. A. Tinker, M. C. Zumwalt, L. D. Lamb, and D. R. Huffman. Gas-phase infrared-emission spectra of $C_{60}$ and $C_{70}$ - temperature-dependent studies. *Chem. Phys. Lett.*, **218**, 295–303 (1994).

[115] K. A. Wang, A. M. Rao, P. C. Eklund, M. S. Dresselhaus, and G. Dresselhaus. Observation of higher-order infrared modes in solid $C_{60}$ films. *Phys. Rev. B*, **48**, 11375–11380 (1993).

[116] T. Pichler, M. Matus, and H. Kuzmany. Electron-Vibrational Mode-Coupling in $K_3C_{60}$ from IR-Transmittance and Reflectivity. *Solid State Commun.*, **86**, 221–225 (1993).

[117] M. C. Martin, D. Koller, and L. Mihaly. Insitu infrared transmission study of Rb-and K-doped fullerenes. *Phys. Rev. B*, **47**, 14607–14610 (1993).

[118] J. P. Merrick, D. Moran, and L. Radom. An evaluation of harmonic vibrational frequency scale factors. *J. Phys. Chem. A*, **111**, 11683–11700 (2007).

[119] R. A. Jishi, R. M. Mirie, M. S. Dresselhaus, G. Dresselhaus, and P. C. Eklund. Force-constant model for the vibrational modes in $C_{70}$. *Phys. Rev. B*, **48**, 5634–5642 (1993).

[120] L. Pergamon Press (editor). *Report of the Commission on Enzymes of the International Union of Biochemistry* (1961).

[121] R. Lemberg and J. Barrett (editors). *Cytochromes*. Academic Press INC. (London) LTD (1973).

[122] X. S. Liu, C. N. Kim, J. Yang, R. Jemmerson, and X. Wang. Induction of apoptotic program in cell-free extracts: Requirement for dATP and cytochrome c. *Cell*, **86**, 147–157 (1996).

[123] X. J. Jiang and X. D. Wang. Cytochrome C-Mediated Apoptosis. *Annu. Rev. Biochem.*, **73**, 87–106 (2004).

[124] E. Margoliash, H. Tuppy, E. L. Smith, and G. Kreil. The Complete Amino-acid Sequence. *Nature*, **192**, 1125–1127 (1961).

[125] R. E. Dickerson, M. L. Kopka, J. Weinzierl, J. Varnum, D. Eisenberg, and E. Margoliash. Location of the Heme in Horse Heart Ferricytochrome c by X-Ray Diffraction. *J. Biol. Chem.*, **242**, 3015–3018 (1967).

[126] H. A. Harbury, J. R. Cronin, M. W. Fanger, T. P. Hettinger, A. J. Murphy, Y. P. Myer, and S. N. Vinogradov. Complex formation between methionine and a heme peptide from cytochrome c. *Proc. Natl. Acad. Sci.*, **54**, 1658–1664 (1965).

[127] C. C. McDonald, W. D. Phillips, and S. N. Vinogradov. Proton Magnetic Resonance Evidence for Methionine-Iron Coordination in Mammalian-type Ferrocytochrome c. *Biochem. Biophys. Res. Commun.*, **36**, 442–449 (1969).

[128] A. Desbois. Resonance Raman spectroscopy of c-type cytochromes. *Biochimie*, **76**, 693–707 (1994).

[129] M. F. Jarrold. Peptides and proteins in the vapor phase. *Annu. Rev. Phys. Chem.*, **51**, 179–207 (2000).

[130] `http://www.uniprot.org/uniprot/P00004#section_seq` (2010).

[131] L. Vértesy, V. Oeding, R. Bender, K. Zepf, and G. Nesemann. Tendamistat (HOE 467), a tight-binding $\alpha$-amylase inhibitor from *streptomyces tendae* 4158 - Isolation, biochemical properties. *Eur. J. Biochem.*, **141**, 505–512 (1984).

[132] H. Aschauer, L. Vértesy, and G. Braunitzer. The sequence of the $\alpha$-amylase inhibitor Hoe-467 A (alpha-amylase inactivator Hoe-467 A) from *Streptomyces tendae* 4158. *Hoppe-Seyler's Z. Physiol. Chem.*, **362**, 465–467 (1981).

[133] A. D. Kline and K. Wüthrich. Secondary Structure of the $\alpha$-Amylase Polypeptide Inhibitor Tendamistat from *Streptomyces tendae* Determined in Solution by $^1$H Nuclear Magnetic Resonance. *J. Mol. Biol.*, **183**, 503–507 (1985).

[134] A. D. Kline, W. Braun, and K. Wüthrich. Determination of the Complete Three-dimensional Structure of the $\alpha$-Amylase Inhibitor Tendamistat in Aqueous Solution by Nuclear Magnetic Resonance and Distance Geometry. *J. Mol. Biol.*, **204**, 675–724 (1988).

[135] J. W. Pflugrath, G. Wiegand, R. Huber, and L. Vértesy. Crystal Structure Determination, Refinement and the Molecular-Model of the $\alpha$-Amylase Inhibitor Hoe-467 A. *J. Mol. Biol.*, **189**, 383–386 (1986).

[136] V. König, L. Vértesy, and T. R. Schneider. Structure of the $\alpha$-amylase inhibitor tendamistat at 0.93 Å. *Acta Crystallogr., Sect. D: Biol. Crystallogr.*, **59**, 1737–1743 (2003).

[137] H. M. Berman, J. Westbrook, Z. Feng, G. Gilliland, T. N. Bhat, H. Weissig, I. N. Shindyalov, and P. E. Bourne. The Protein Data Bank. *Nucleic Acids Res.*, **28**, 235–242 (2000).

[138] D. Frishman and P. Argos. Knowledge-based protein secondary structure assignment. *Proteins: Struct., Funct., Genet.*, **23**, 566–579 (1995).

[139] Y. O. Kamatari, T. Konno, M. Kataoka, and K. Akasaka. The Methanol-induced Globular and Expanded Denatured States of Cytochrome c: A Study by CD Fluorescence, NMR And Small-angle X-ray Scattering. *J. Mol. Biol.*, **259**, 512–523 (1996).

[140] L. Konermann and D. J. Douglas. Acid-Induced Unfolding of Cytochrome c at Different Methanol Concentrations: Electrospray Ionization Mass Spectrometry Specifically Monitors Changes in the Tertiary Structure. *Biochemistry*, **36**, 12296–12302 (1997).

[141] A. T. Iavarone and E. R. Williams. Mechanism of charging and supercharging molecules in electrospray ionization. *J. Am. Chem. Soc.*, **125**, 2319–2327 (2003).

[142] S. H. Lomeli, S. Yin, R. R. Ogorzalek Loo, and J. A. Loo. Increasing Charge While Preserving Noncovalent Protein Complexes for ESI-MS. *J. Am. Soc. Mass Spectrom.*, **20**, 593–596 (2009).

[143] S. H. Lomeli, I. X. Peng, S. Yin, R. R. Ogorzalek Loo, and J. A. Loo. New Reagents for Increasing ESI Multiple Charging of Proteins and Protein Complexes. *J. Am. Soc. Mass Spectrom.*, **21**, 127–131 (2010).

[144] R. Grandori. Origin of the conformation dependence of protein charge-state distributions in electrospray ionization mass spectrometry. *J. Mass Spectrom.*, **38**, 11–15 (2003).

[145] V. J. Nesatyy and M. J.-F. Suter. On the conformation-dependent neutralization theory and charging of individual proteins and their non-covalent complexes in the gas phase. *J. Mass Spectrom.*, **39**, 93–97 (2004).

[146] J. Oomens, N. Polfer, D. T. Moore, L. van der Meer, A. G. Marshall, J. R. Eyler, G. Meijer, and G. von Helden. Charge-state resolved mid-infrared spectroscopy of a gas-phase protein. *Phys. Chem. Chem. Phys.*, **7**, 1345–1348 (2005).

[147] T. M. Watson and J. D. Hirst. Influence of electrostatic environment on the vibrational frequencies of proteins. *J. Phys. Chem. A*, **107**, 6843–6849 (2003).

[148] S. N. Timasheff and G. D. Fasman (editors). *Structure and Stability of Biological Macromolecules*. Marcel Dekker (1969).

[149] S. J. McClellan and E. I. Franses. Adsorption of bovine serum albumin at solid/aqueous interfaces. *Colloids Surf., A*, **260**, 265–275 (2005).

[150] A. Barth. The infrared absorption of amino acid side chains. *Prog. Biophys. Mol. Biol.*, **74**, 141–173 (2000).

[151] E. G. Bendit. Infrared Absorption of Tyrosine Side Chains in Proteins. *Biopolymers*, **5**, 525–533 (1967).

[152] C. F. Correia, P. O. Balaj, D. Scuderi, P. Maitre, and G. Ohanessian. Vibrational signatures of protonated, phosphorylated amino acids in the gas phase. *J. Am. Chem. Soc.*, **130**, 3359–3370 (2008).

[153] A. Dong, P. Huang, and W. S. Caughey. Protein Secondary Structures in Water from Second-Derivative Amide I Infrared Spectra. *Biochemistry*, **29**, 3303–3308 (1990).

[154] J. O. Speare and T. S. Rush III. Ir Spectra of Cytochrome c Denatured with Deuterated Guanidine Hydrochloride Show Increase in $\beta$ Sheet. *Biopolymers*, **72**, 193–204 (2003).

[155] G. von Helden. *Investigations of the Structure and Energetics of Gas Phase Cluster Ions Using Ion Chromatography*. Ph.D. thesis, University of California Santa Barbara (1994).

[156] G. von Helden, M. T. Hsu, P. R. Kemper, and M. T. Bowers. Structures of carbon cluster ions from 3 to 60 atoms: Linears to rings to fullerenes. *J. Chem. Phys.*, **95**, 3835–3837 (1991).

[157] K. B. Shelimov and M. F. Jarrold. "Denaturation" and Refolding of Cytochrome c *in Vacuo*. *J. Am. Chem. Soc.*, **118**, 10313–10314 (1996).

[158] D. E. Clemmer and M. F. Jarrold. Ion Mobility Measurements and their Applications to Clusters and Biomolecules. *J. Mass Spectrom.*, **32**, 577–592 (1997).

[159] B. C. Bohrer, S. I. Merenbloom, S. L. Koeniger, A. E. Hilderbrand, and D. E. Clemmer. Biomolecule Analysis by Ion Mobility Spectrometry. *Annu. Rev. Anal. Chem.*, **1**, 293–327 (2008).

[160] K. B. Shelimov, D. E. Clemmer, R. R. Hudgins, and M. F. Jarrold. Protein Structure *in Vacuo*: Gas-Phase Confirmations of BPTI and Cytochrome c. *J. Am. Chem. Soc.*, **119**, 2240–2248 (1997).

[161] E. R. Badman, S. Myung, and D. E. Clemmer. Evidence for Unfolding and Refolding of Gas-Phase Cytochrome c Ions in a Paul Trap. *J. Am. Soc. Mass Spectrom.*, **16**, 1493–1497 (2005).

[162] T. D. Wood, R. A. Chorush, F. M. Wampler III, D. P. Little, P. B. O'Connor, and F. W. McLafferty. Gas-phase folding and unfolding of cytochrome c cations. *Proc. Natl. Acad. Sci.*, **92**, 2451–2454 (1995).

[163] F. W. McLafferty, Z. Q. Guan, U. Haupts, T. D. Wood, and N. L. Kelleher. Gaseous Conformational Structures of Cytochrome c. *J. Am. Chem. Soc.*, **120**, 4732–4740 (1998).

[164] `http://www.uniprot.org/`.

[165] F. He, C. L. Hendrickson, and A. G. Marshall. Unequivocal determination of metal atom oxidation state in naked heme proteins: Fe(III)myoglobin, Fe(III)cytochrome c, Fe(III)cytochrome b5, and Fe(III)cytochrome b5 . *J. Am. Soc. Mass Spectrom.*, **11**, 120–126 (2000).

[166] C. Kapota, J. Lemaire, P. Maitre, and G. Ohanessian. Vibrational signature of charge solvation vs salt bridge isomers of sodiated amino acids in the gas phase. *J. Am. Chem. Soc.*, **126**, 1836–1842 (2004).

[167] `http://webbook.nist.gov/`.

[168] C. A. Swenson and R. Formanek. Infrared Study of Poly-L-Proline in Aqueous Solution. *J. Phys. Chem.*, **71**, 4073–4077 (1967).

[169] R. Wu and T. B. McMahon. An Investigation of Protonation Sites and Conformations of Protonated Amino Acids by IRMPD Spectroscopy. *ChemPhysChem*, **9**, 2826–2835 (2008).

[170] R. Linder, M. Nispel, T. Häber, and K. Kleinermanns. Gas-phase FT-IR-spectra of natural amino acids. *Chem. Phys. Lett.*, **409**, 260–264 (2005).

[171] R. Wu and T. B. McMahon. Infrared multiple photon dissociation spectra of proline and glycine proton-bound homodimers. Evidence for zwitterionic structure. *J. Am. Chem. Soc.*, **129**, 4864–4865 (2007).

[172] R. Wu and T. B. McMahon. Stabilization of the Zwitterionic structure of proline by an alkylammonium ion in the gas phase. *Angew. Chem., Int. Ed.*, **46**, 3668–3671 (2007).

[173] R. Tantipolphan, T. Rades, and N. J. Medlicott. Insights into the structure of protein by vibrational spectroscopy. *Curr. Pharm. Anal.*, **4**, 53–68 (2008).

[174] P. Y. Chou and G. D. Fasman. Prediction of Protein Conformation. *Biochemistry*, **13**, 222–245 (1974).

[175] S. H. White and W. C. Wimley. Membrane Protein Folding and Stability: Physical Principles. *Annu. Rev. Biophys. Biomol. Struct.*, **28**, 319–365 (1999).

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

[177] Z. S. Shi, C. A. Olson, G. D. Rose, R. L. Baldwin, and N. R. Kallenbach. Polyproline II structure in a sequence of seven alanine residues. *Proc. Natl. Acad. Sci.*, **99**, 9190–9195 (2002).

[178] A. Wada. The alpha-helix as an electrical macro-dipole. *Adv. Biophys.*, **9**, 1–63 (1976).

[179] F. Eisenmenger, U. H. E. Hansmann, S. Hayryan, and C. K. Hu. [SMMP] A modern package for simulation of proteins. *Comput. Phys. Commun.*, **138**, 192–212 (2001).

[180] Y. Wei, W. Nadler, and U. H. E. Hansmann. On the helix-coil transition in alanine based polypeptides in gas phase. *J. Chem. Phys.*, **126**, 204307 (2007).

[181] W. G. J. Hol, P. T. van Duijnen, and H. J. C. Berendsen. The $\alpha$-helix dipole and the properties of proteins. *Nature*, **273**, 443–446 (1978).

[182] R. R. Hudgins and M. F. Jarrold. Helix Formation in Unsolvated Alanine-Based Peptides: Helical Monomers and Helical Dimers. *J. Am. Chem. Soc.*, **121**, 3494–3501 (1999).

[183] J. A. Stearns, C. Seaiby, O. V. Boyarkin, and T. R. Rizzo. Spectroscopy and conformational preferences of gas-phase helices. *Phys. Chem. Chem. Phys.*, **11**, 125–132 (2009).

[184] R. B. Merrifield. Solid Phase Peptide Synthesis. I. Synthesis of a Tetrapeptide. *J. Am. Chem. Soc.*, **85**, 2149–2154 (1963).

[185] D. A. Wellings and E. Atherton. Standard Fmoc Protocols. *Methods Enzymol.*, **289**, 44–67 (1997).

[186] E. P. L. Hunter and S. G. Lias. Evaluated Gas Phase Basicities and Proton Affinities of Molecules: An Update. *J. Phys. Chem. Ref. Data*, **27**, 413–656 (1998).

[187] B. Lucas, G. Grégoire, J. Lemaire, P. Maître, J. M. Ortega, A. Rupenyan, B. Reimann, J. P. Schermann, and C. Desfrançois. Investigation of the protonation site in the dialanine peptide by infrared multiphoton dissociation spectroscopy. *Phys. Chem. Chem. Phys.*, **6**, 2659–2663 (2004).

[188] S. Suzuki, T. Ohshima, N. Tamiya, K. Fukushima, T. Shimanouchi, and S. I. Mizushima. Infrared spectra of deuterated alpha-amino acids $NH_3^+$CDRCOO$^-$ assignment of the absorption bands of $\alpha$-alanine. *Spectrochim. Acta*, **15**, 969–976 (1959).

[189] R. C. Dunbar, J. D. Steill, and J. Oomens. Conformations and vibrational spectroscopy of metal-ion/polylalanine complexes. *Int. J. Mass Spectrom.*, **297**, 107–115 (2010).

[190] R. Wu and T. B. McMahon. Infrared Multiple Photon Dissociation Spectroscopy as Structural Confirmation for GlyGlyGlyH$^+$ and AlaAlaAlaH$^+$ in the Gas Phase. Evidence for Amide Oxygen as the Protonation Site. *J. Am. Chem. Soc.*, **129**, 11312–11313 (2007).

[191] M.-P. Gaigeot. Infrared spectroscopy of the alanine dipeptide analog in liquid water with DFT-MD. Direct evidence for P-II/beta conformations. *Phys. Chem. Chem. Phys.*, **12**, 10198–10209 (2010).

[192] T. D. Vaden, T. S. J. A. de Boer, J. P. Simons, L. C. Snoek, S. Suhai, and B. Paizs. Vibrational Spectroscopy and Conformational Structure of Protonated Polyalanine Peptides Isolated in the Gas Phase. *J. Phys. Chem. A*, **112**, 4608–4616 (2008).

[193] R. Wieczorek and J. J. Dannenberg. $\alpha$-Helical Peptides Are Not Protonated at the N-Terminus in the Gas Phase. *J. Am. Chem. Soc.*, **126**, 12278–12279 (2004).

[194] M. Goodman, I. Listowsky, Y. Masuda, and F. Boardman. Conformational Aspects of Polypeptides .VIII. Helical Assignments via Far Ultraviolet Absorption Spectra and Optical Activity. *Biopolymers*, **1**, 33–42 (1963).

[195] V. A. Jaravine, A. T. Alexandrescu, and S. Grzesiek. Observation of the closing of individual hydrogen bonds during TFE-induced helix formation in a peptide. *Protein Sci.*, **10**, 943–950 (2001).

[196] P. Wallimann, R. J. Kennedy, J. S. Miller, W. Shalongo, and D. S. Kemp. Dual Wavelength Parametric Test of Two-State Models for Circular Dichroism Spectra of Helical Polypeptides: Anomalous Dichroic Properties of Alanine-Rich Peptides. *J. Am. Chem. Soc.*, **125**, 1203–1220 (2003).

[197] R. R. Hudgins, Y. Mao, M. A. Ratner, and M. F. Jarrold. Conformations of Gly$_n$H$^+$ and Ala$_n$H$^+$ Peptides in the Gas Phase. *Biophys. J.*, **76**, 1591–1597 (1999).

[198] M. Kohtani, B. S. Kinnear, and M. F. Jarrold. Metal-Ion Enhanced Helicity in the Gas Phase. *J. Am. Chem. Soc.*, **122**, 12377–12378 (2000).

[199] M. Kohtani, M. F. Jarrold, S. Wee, and R. A. J. O'Hair. Metal Ion Interactions with Polyalanine Peptides. *J. Phys. Chem. B*, **108**, 6093–6097 (2004).

[200] I. A. Kaltashov and C. Fenselau. Stability of Secondary Structural Elements in a Solvent-Free Environment: The $\alpha$ Helix. *Proteins*, **27**, 165–170 (1997).

[201] L. G. Presta and G. D. Rose. Helix Signals in Proteins. *Science*, **240**, 1632–1641 (1988).

[202] J. W. Seale, R. Srinivasan, and G. D. Rose. Sequence determinants of the capping box, a stabilizing motif at the N-termini of $\alpha$-helices. *Protein Sci.*, **3**, 1741–1745 (1994).

[203] R. R. Hudgins, M. A. Ratner, and M. F. Jarrold. Design of Helices That Are Stable in Vacuo. *J. Am. Chem. Soc.*, **120**, 12974–12975 (1998).

[204] M. F. Jarrold. Helices and sheets in vacuo. *Phys. Chem. Chem. Phys.*, **9**, 1659–1671 (2007).

[205] M. Kohtani and M. F. Jarrold. Water Molecule Adsorption on Short Alanine Peptides: How Short Is the Shortest Gas-Phase Alanine-Based Helix? *J. Am. Chem. Soc.*, **126**, 8454–8458 (2004).

[206] M. Kohtani, T. C. Jones, J. E. Schneider, and M. F. Jarrold. Extreme Stability of an Unsolvated $\alpha$-Helix. *J. Am. Chem. Soc.*, **126**, 7420–7421 (2004).

[207] J. A. Stearns, O. V. Boyarkin, and T. R. Rizzo. Spectroscopic Signatures of Gas-Phase Helices: Ac-Phe-(Ala)$_5$-Lys-H$^+$ and Ac-Phe-(Ala)$_{10}$-Lys-H$^+$. *J. Am. Chem. Soc.*, **129**, 13820–13821 (2007).

[208] J. A. Stearns, O. V. Boyarkin, and T. R. Rizzo. Effects of N-terminus substitution on the structure and spectroscopy of gas-phase helices. *Chimia*, **62**, 240–243 (2008).

[209] M. Rossi, V. Blum, P. Kupser, G. von Helden, F. Bierau, K. Pagel, G. Meijer, and M. Scheffler. Secondary Structure of Ac-Ala$_n$-LysH$^+$Polyalanine Peptides (n= 5,10,15) in Vacuo: Helical or Not? *J. Phys. Chem. Lett.*, **1**, 3465–3470 (2010).

[210] V. Blum, R. Gehrke, F. Hanke, P. Havu, V. Havu, X. Ren, K. Reuter, and M. Scheffler. Ab initio molecular simulations with numeric atom-centered orbitals. *Comput. Phys. Commun.*, **180**, 2175–2196 (2009).

[211] J. P. Perdew, K. Burke, and M. Ernzerhof. Generalized Gradient Approximation Made Simple. *Phys. Rev. Lett.*, **77**, 3865–3868 (1996).

[212] A. Tkatchenko and M. Scheffler. Accurate Molecular Van Der Waals Interactions from Ground-State Electron Density and Free-Atom Reference Data. *Phys. Rev. Lett.*, **102**, 073005 (2009).

[213] J. B. Pendry. Reliability factors for LEED calculations. *J. Phys. C: Solid State Phys.*, **13**, 937–944 (1980).

[214] G. von Helden, T. Wyttenbach, and M. T. Bowers. Conformation of Macromolecules in the Gas Phase: Use of Matrix-Assisted Laser Desorption Methods in Ion Chromatography. *Science*, **267**, 1483–1485 (1995).

[215] T. Wyttenbach and M. T. Bowers. Intermolecular interactions in biomolecular systems examined by mass spectrometry. *Annu. Rev. Phys. Chem.*, **58**, 511–533 (2007).

[216] H. Oh, K. Breuker, S. K. Sze, Y. Ge, B. K. Carpenter, and F. W. McLafferty. Secondary and tertiary structures of gaseous protein ions characterized by electron capture dissociation mass spectrometry and photofragment spectroscopy. *Proc. Natl. Acad. Sci. USA*, **99**, 15863–15868 (2002).

[217] M. Gerhards, C. Unterberg, A. Gerlach, and A. Jansen. $\beta$-sheet model systems in the gas phase: Structures and vibrations of Ac-Phe-NHMe and its dimer (Ac-Phe-NHMe)$_2$. *Phys. Chem. Chem. Phys.*, **6**, 2682–2690 (2004).

[218] A. Gerlach, C. Unterberg, H. Fricke, and M. Gerhards. Structures of Ac-Trp-OMe and its dimer (Ac-Trp-OMe)$_2$ in the gas phase: influence of a polar group in the side-chain. *Mol. Phys.*, **103**, 1521–1529 (2005).

[219] H. Fricke, A. Gerlach, and M. Gerhards. Structure of a beta-sheet model system in the gas phase: Analysis of the fingerprint region up to 10 $\mu$m. *Phys. Chem. Chem. Phys.*, **8**, 1660–1662 (2006).

[220] I. Compagnon, J. Oomens, G. Meijer, and G. von Helden. Mid-infrared spectroscopy of protected peptides in the gas phase: A probe of the backbone conformation. *J. Am. Chem. Soc.*, **128**, 3592–3597 (2006).

[221] H. Fricke, A. Funk, T. Schrader, and M. Gerhards. Investigation of secondary structure elements by IR/UV double resonance spectroscopy: Analysis of an isolated beta-sheet model system. *J. Am. Chem. Soc.*, **130**, 4692–4698 (2008).

[222] C. S. Colley, S. R. Griffiths-Jones, M. W. George, and M. S. Searle. Do interstrand hydrogen bonds contribute to beta-hairpin peptide stability in solution? IR analysis of peptide folding in water. *Chem. Commun.*, pages 593–594 (2000).

[223] R. N. A. H. Lewis, E. J. Prenner, L. H. Kondejewski, C. R. Flach, R. Mendelsohn, R. S. Hodges, and R. N. McElhaney. Fourier transform infrared spectroscopic studies of the interaction of the antimicrobial peptide gramicidin S with lipid micelles and with lipid monolayer and bilayer membranes. *Biochemistry*, **38**, 15193–15203 (1999).

[224] K. Pagel, P. Kupser, F. Bierau, N. C. Polfer, J. D. Steill, J. Oomens, G. Meijer, B. Koksch, and G. von Helden. Gas-phase IR spectra of intact alpha-helical coiled coil protein complexes. *Int. J. Mass Spectrom.*, **283**, 161–168 (2009).

[225] P. Dugourd, R. Antoine, G. Breaux, M. Broyer, and M. F. Jarrold. Entropic stabilization of isolated beta-sheets. *J. Am. Chem. Soc.*, **127**, 4675–4679 (2005).

[226] S. L. Bernstein, N. F. Dupuis, N. D. Lazo, T. Wyttenbach, M. M. Condron, G. Bitan, D. B. Teplow, J. E. Shea, B. T. Ruotolo, C. V. Robinson, and M. T. Bowers. Amyloid-beta protein oligomerization and the importance of tetramers and dodecamers in the aetiology of Alzheimer's disease. *Nat. Chem.*, **1**, 326–331 (2009).

[227] G. F. Gause and M. G. Brazhnikova. Gramicidin S and its use in the treatment of infected wounds. *Nature*, **154**, 703 (1944).

[228] R. L. M. Synge. 'Gramicidin S': Over-all Chemical Characteristics and Amino-acid Composition. *Biochem. J.*, **39**, 363–367 (1945).

[229] D. C. Hodgkin and B. M. Oughton. Possible molecular models for gramicidin-s and their relationship to present ideas of protein structure. *Biochem. J.*, **65**, 752–756 (1957).

[230] S. E. Hull, R. Karlsson, P. Main, M. M. Woolfson, and E. J. Dodson. The Crystal Structure of a Hydrated Gramicidin S Urea Complex. *Nature*, **275**, 206–207 (1978).

[231] K. Yamada, M. Unno, K. Kobayashi, H. Oku, H. Yamamura, S. Araki, H. Matsumoto, R. Katakai, and M. Kawai. Stereochemistry of protected ornithine side chains of gramicidin S derivatives: X-ray crystal structure of the bis-boc-tetra-N-methyl derivative of gramicidin S. *J. Am. Chem. Soc.*, **124**, 12684–12688 (2002).

[232] E. M. Krauss and S. I. Chan. Intramolecular Hydrogen-Bonding in Gramicidin S .2. Ornithine. *J. Am. Chem. Soc.*, **104**, 6953–6961 (1982).

[233] Y. Xu, I. P. Sugár, and N. R. Krishna. A variable target intensity-restrained global optimization (VARTIGO) procedure for determining three-dimensional structures of polypeptides from NOESY data: Application to gramicidin-s. *J. Biomol. NMR*, **5**, 37–48 (1995).

[234] M. Kawai, T. Yamamoto, K. Yamada, M. Yamaguchi, S. Kurobe, H. Yamamura, S. Araki, Y. Butsugan, K. Kobayashi, R. Katakai, K. Saito, and T. Nakajima. Stereochemistry of protected ornithine side chains in gramicidin S derivatives and their resistance to N-methylation. *Lett. Pept. Sci.*, **5**, 5–12 (1998).

[235] A. C. Gibbs, L. H. Kondejewski, W. Gronwald, A. M. Nip, R. S. Hodges, B. D. Sykes, and D. S. Wishart. Unusual $\beta$-sheet periodicity in small cyclic peptides. *Nat. Struct. Bio.*, **5**, 284–288 (1998).

[236] D. S. Gross and E. R. Williams. Structure of gramicidin S $(M+H+X)^{2+}$ ions (X=Li, Na, K) probed by proton transfer reactions. *J. Am. Chem. Soc.*, **118**, 202–204 (1996).

[237] S. E. Rodriguez-Cruz, J. S. Klassen, and E. R. Williams. Hydration of Gas-Phase Gramicidin S $(M+2H)^{2+}$ Ions Formed by Electrospray: The Transition From Solution to Gas-Phase Structure. *J. Am. Soc. Mass Spectrom.*, **8**, 565–568 (1997).

[238] B. T. Ruotolo, C. C. Tate, and D. H. Russell. Ion mobility-mass spectrometry applied to cyclic peptide analysis: Conformational preferences of gramicidin S and linear analogs in the gas phase. *J. Am. Soc. Mass Spectrom.*, **15**, 870–878 (2004).

[239] P. Wadhwani, S. Afonin, M. Ieromino, J. Buerck, and A. S. Ulrich. Optimized protocol for synthesis of cyclic gramicidin S: Starting amino acid is key to high yield. *J.Org. Chem.*, **71**, 55–61 (2006).

[240] R. R. Julian and J. L. Beauchamp. The unusually high proton affinity of aza-18-crown-6 ether: Implications for the molecular recognition of lysine in peptides by lariat crown ethers. *J. Am. Soc. Mass Spectrom.*, **13**, 493–498 (2002).

[241] TURBOMOLE V6.0 2009, a development of University of Karlsruhe and Forschungszentrum Karlsruhe GmbH, 1989-2007, TURBOMOLE GmbH, since 2007; available from `http://www.turbomole.com`.

[242] D. Case, T. Darden, T. Cheatham III, C. Simmerling, J. Wang, R. Duke, R. Luo, K. Merz, D. Pearlman, M. Crowley, R. Walker, W. Zhang, B. Wang, S. Hayik, A. Roitberg, G. Seabra, K. Wong, F. Paesani, X. Wu, S. Brozell, V. Tsui, H. Gohlke, L. Yang, C. Tan, J. Mongan, V. V. Hornak, G. Cui, P. Beroza, D. Mathews, C. Schafmeister, W. Ross, and P. Kollman. AMBER 9, University of California, San Francisco (2006).

[243] P. Kollman, R. Dixon, W. Cornell, T. Fox, C. Chipot, and A. Pohorille. *Computer Simulation of Biomolecular Systems*, volume 3. Elsevier (1997).

[244] A. Mitsutake, Y. Sugita, and Y. Okamoto. Generalized-ensemble algorithms for molecular simulations of biopolymers. *Biopolymers*, **60**, 96–123 (2001).

[245] A. Scott and L. Radom. Harmonic vibrational frequencies: An evaluation of Hartree-Fock, Moller-Plesset, quadratic configuration interaction, density functional theory, and semiempirical scale factors. *J. Phys. Chem.*, **100**, 16502–16513 (1996).

[246] J. Loo. Studying noncovalent protein complexes by electrospray ionization mass spectrometry. *Mass Spectrom. Rev.*, **16**, 1–23 (1997).

[247] R. H. H. van den Heuvel and A. J. R. Heck. Native protein mass spectrometry: from intact oligomers to functional machineries. *Curr. Opin. Chem. Biol.*, **8**, 519–526 (2004).

[248] A. J. R. Heck and R. H. H. van den Heuvel. Investigation of intact protein complexes by mass spectrometry. *Mass Spectrom. Rev.*, **23**, 368–389 (2004).

[249] A. Ashcroft. Recent developments in electrospray ionisation mass spectrometry: noncovalently bound protein complexes. *Nat. Prod. Rep.*, **22**, 452–464 (2005).

[250] J. L. P. Benesch and C. Robinson. Mass spectrometry of macromolecular assemblies: preservation and dissociation. *Curr. Opin. Struct. Biol.*, **16**, 245–251 (2006).

[251] F. Sobott, H. Hernandez, M. G. McCammon, M. A. Tito, and C. V. Robinson. A tandem mass spectrometer for improved transmission and analysis of large macromolecular assemblies. *Anal. Chem.*, **74**, 1402–1407 (2002).

[252] M. S. Wilm and M. Mann. Electrospray And Taylor-Cone Theory, Doles Beam Of Macromolecules At Last. *Int. J. Mass spectrom.*, **136**, 167–180 (1994).

[253] N. P. Barrera, N. D. Bartolo, P. J. Booth, and C. V. Robinson. Micelles protect membrane complexes from solution to vacuum. *Science*, **321**, 243–246 (2008).

[254] G. Siuzdak, B. Bothner, M. Yeager, C. Brugidou, C. M. Fauquet, K. Hoey, and C. M. Chang. Mass spectrometry and viral analysis. *Chem. Biol.*, **3**, 45–48 (1996).

[255] B. T. Ruotolo, J. L. P. Benesch, A. M. Sandercock, S. J. Hyung, and C. V. Robinson. Ion mobility-mass spectrometry analysis of large protein complexes. *Nat. Protoc.*, **3**, 1139–1152 (2008).

[256] B. T. Ruotolo, K. Giles, I. Campuzano, A. M. Sandercock, R. H. Bateman, and C. V. Robinson. Evidence for macromolecular protein rings in the absence of bulk water. *Science*, **310**, 1658–1661 (2005).

[257] S. Bernstein, T. Wyttenbach, A. Baumketner, J. Shea, G. Bitan, D. Teplow, and M. Bowers. Amyloid beta-protein: Monomer structure and early aggregation states of A beta 42 and its Pro(19) alloform. *J. Am. Chem. Soc.*, **127**, 2075–2084 (2005).

[258] A. E. Counterman and D. E. Clemmer. Large anhydrous polyalanine ions: Evidence for extended helices and onset of a more compact state. *J. Am. Chem. Soc.*, **123**, 1490–1498 (2001).

[259] R. Hudgins and M. Jarrold. Conformations of unsolvated glycine-based peptides. *J. Phys. Chem. B*, **104**, 2154–2158 (2000).

[260] R. Sudha, M. Kohtani, and M. F. Jarrold. Non-covalent interactions between unsolvated peptides: Helical complexes based on acid-base interactions. *J. Phys. Chem. B*, **109**, 6442–6447 (2005).

[261] W. K. Surewicz, H. H. Mantsch, and D. Chapman. Determination of Protein Secondary Structure by Fourier Transform Infrared Spectroscopy: A Critical Assessment. *Biochemistry*, **32**, 389–394 (1993).

[262] M. F. Bush, J. Oomens, R. J. Saykally, and E. R. Williams. Alkali metal ion binding to glutamine and glutamine derivatives investigated by infrared action spectroscopy and theory. *J. Phys. Chem. A*, **112**, 8578–8584 (2008).

[263] N. C. Polfer, J. Oomens, and R. C. Dunbar. Alkali metal complexes of the Dipeptides PheAla and AlaPhe : IRMPD spectroscopy. *ChemPhysChem*, **9**, 579–589 (2008).

[264] N. C. Polfer, J. Oomens, S. Suhai, and B. Paizs. Infrared Spectroscopy and Theoretical Studies on Gas-Phase Protonated Leu-Enkephalin and its Fragments: Direct Experimental Evidence for the Mobile Proton. *J. Am. Chem. Soc.*, **129**, 5887–5897 (2007).

[265] T. D. Vaden, S. A. N. Gowers, T. S. J. A. de Boer, J. D. Steill, J. Oomens, and L. C. Snoek. Conformational Preferences of an Amyloidogenic Peptide: IR Spectroscopy of Ac-VQIVYK-NHMe. *J. Am. Chem. Soc.*, **130**, 14640–14650 (2008).

[266] E. Wolf, P. S. Kim, and B. Berger. MultiCoil: A program for predicting two- and three-stranded coiled coils. *Protein Sci*, **6**, 1179–1189 (1997).

[267] J. M. Mason and K. M. Arndt. Coiled coil domains: Stability, specificity, and biological implications. *ChemBioChem*, **5**, 170–176 (2004).

[268] Y. B. Yu. Coiled-coils: stability, specificity, and drug delivery potential. *Adv. Drug Delivery Rev.*, **54**, 1113–1129 (2002).

[269] D. N. Woolfson. The design of coiled-coil structures and assemblies. *Adv. Protein Chem.*, **70**, 79–112 (2005).

[270] K. Pagel, T. Vagt, T. Kohajda, and B. Koksch. From $\alpha$-helix to $\beta$-sheet - a reversible metal ion induced peptide secondary structure switch. *Org. Biomol. Chem.*, **3**, 2500–2502 (2005).

[271] K. Pagel and B. Koksch. Following polypeptide folding and assembly with conformational switches. *Curr. Opin. Chem. Biol.*, **12**, 730–739 (2008).

[272] M. G. Oakley and J. J. Hollenbeck. The design of antiparallel coiled coils. *Curr. Opin. Struct. Biol.*, **11**, 450–457 (2001).

[273] D. G. Gurnon, J. A. Whitaker, and M. G. Oakley. Design and characterization of a homodimeric antiparallel coiled coil. *J. Am. Chem. Soc.*, **125**, 7518–7519 (2003).

[274] K. Pagel, K. Seeger, B. Seiwert, A. Villa, A. E. Mark, S. Berger, and B. Koksch. Advanced approaches for the characterization of a de novo designed antiparallel coiled coil peptide. *Org. Biomol. Chem.*, **3**, 1189–1194 (2005).

[275] P. Burkhard, S. Ivaninskii, and A. Lustig. Improving coiled-coil stability by optimizing ionic interactions. *J. Mol. Biol.*, **318**, 901–910 (2002).

[276] M. Hoshino, N. Yumoto, S. Yoshikawa, and Y. Goto. Design and characterization of the anion-sensitive coiled-coil peptide. *Protein Sci.*, **6**, 1396–1404 (1997).

[277] A. Patriksson, E. Marklund, and D. van der Spoel. Protein structures under electrospray conditions. *Biochemistry*, **46**, 933–945 (2007).

[278] H. Wendt, E. Dürr, R. M. Thomas, M. Przybylski, and H. R. Bosshard. Characterization Of Leucine-Zipper Complexes By Electrospray-Ionization Mass-Spectrometry. *Protein Sci.*, **4**, 1563–1570 (1995).

[279] S. Witte, F. Neumann, U. Krawinkel, and M. Przybylski. Mass spectrometric identification of leucine zipper-like homodimer complexes of the autoantigen L7. *J. Biol. Chem.*, **271**, 18171–18175 (1996).

[280] K. Pagel, S. C. Wagner, K. Samedov, H. von Berlepsch, C. Böttcher, and B. Koksch. Random Coils, $\beta$-Sheet Ribbons, and $\alpha$-Helical Fibers: One Peptide Adopting Three Different Secondary Structures at Will. *J. Am. Chem. Soc.*, **128**, 2196–2197 (2006).

[281] K. Pagel, S. C. Wagner, R. R. Araghi, H. von Berlepsch, C. Böttcher, and B. Koksch. Intramolecular Charge Interactions as a Tool to Control the Coiled-Coil-to-Amyloid Transformation. *Chem. Eur. J.*, **14**, 11442–11451 (2008).

[282] J. M. Bakker, C. Plützer, I. Hünig, T. Häber, I. Compagnon, G. von Helden, G. Meijer, and K. Kleinermanns. Folding structures of isolated peptides as revealed by gas-phase mid-infrared spectroscopy. *ChemPhysChem*, **6**, 120–128 (2005).

[283] D. T. Moore, J. Oomens, L. van der Meer, G. von Helden, G. Meijer, J. Valle, A. G. Marshall, and J. R. Eyler. Probing the Vibrations of Shared, $OH^+O$-Bound Protons in the Gas Phase. *ChemPhysChem*, **5**, 740–743 (2004).

[284] S. Goyal, D. L. Schutt, and G. Scoles. Vibrational Spectroscopy of Sulfur-Hexafluoride Attached to Helium Clusters. *Phys. Rev. Lett.*, **69**, 933–936 (1992).

[285] M. Hartmann, R. E. Miller, J. P. Toennies, and A. Vilesov. Rotationally Resolved Spectroscopy of $SF_6$ in Liquid Helium Clusters: A Molecular Probe of Cluster Temperature. *Phys. Rev. Lett.*, **75**, 1566–1569 (1995).

[286] F. Bierau, P. Kupser, G. Meijer, and G. von Helden. Catching Proteins in Liquid Helium Droplets. *Phys. Rev. Lett.*, **105**, 133402 (2010).

[287] M. N. Slipchenko, S. Kuma, T. Momose, and A. F. Vilesov. Intense pulsed helium droplet beams. *Rev. Sci. Instrum.*, **73**, 3600–3605 (2002).

[288] F. Bierau. *Snatching charged (bio)molecules from a linear ion trap by liquid helium droplets.* Ph.D. thesis, Fritz-Haber-Institut der Max-Planck-Gesellschaft (2011).

[289] K. Pagel, S.-J. Hyung, B. T. Ruotolo, and C. V. Robinson. Alternate Dissociation Pathways Identified in Charge-Reduced Protein Complex Ions. *Anal. Chem.*, **82**, 5363–5372 (2010).

# Danksagung

Die hier vorliegende Arbeit entstand am Fritz-Haber-Institut der Max-Planck-Gesellschaft in der Abteilung Molekülphysik unter der Leitung von Herrn Prof. Dr. Gerard Meijer. An dieser Stelle möchte ich mich bei den Personen bedanken, die mich bei meiner Arbeit unterstützt haben und wertvolle Beiträge sowohl beruflicher als auch menschlicher Natur geleistet haben.

Zuerst möchte ich mich bei Gerard dafür bedanken, dass ich die Möglichkeit bekommen habe, sowohl meine Diplomarbeit als auch meine Doktorarbeit in seiner Abteilung durchführen zu können. Vielen Dank für Anregungen und ein offenes Ohr. Es ist ja auch immer eine Frage des Chefs, wie sich die Arbeitsatmosphäre und die allgemeine Stimmung in einer Abteilung entwickeln kann die in der Molekülphysik überaus angenehm ist. Ich finde es toll, dass so viele gemeinsame Veranstaltungen existieren wie zum Beispiel das Geburtstagskuchenessen, die Weihnachtsfeiern oder die Wandertage, wobei derjenige, welcher die Verbindung der lutheranischen Kultur mit der nördlichsten Weinanbautechnik wagte, besonders legendär ist.

Was wäre ein Doktorand ohne seinen "Kopromoter" (ich finde das deutsch Wort Betreuer irgendwie komisch...)? Einen besonderen Dank geht daher an den Gert von Helden, der diese Aufgabe bei mir übernommen hat und das auch schon während der Diplomarbeitszeit. Ich habe extrem viel lernen können und zwar vom Apparaturaufbau ("nen SHV Stecker brauchste? Nimm doch ne Büroklammer...") bis hin zur Semantik (conclusion kommt vor assumption, oder doch anders herum?). Gert kam immer mit vielen Ideen ins Labor was dazu führte, dass das Labor das "Bastlerlabor" wurde. Danke auch für die lustige Zeit in Holland nach nächtelanger Schicht (ich sag nur Kneipe gegenüber der blauen Brücke) und das Verbessern meines Schreibstils.

Mein nächster Dank geht an Gerts Arbeitsgruppe. Zuallererst an Frauke Bierau, mit der ich viel Zeit im Labor zubringen durfte aber auch sonst lustige Abende hatte (du weisst, der erhobene Zeigefinger in illustrer Runde...). Dir und Stephan Warnke, welcher meinen Musikgeschmack nicht verrohen läßt, danke ich auch für das Korrekturlesen der Arbeit. Der Apparatus ist nun bei Frank Filsinger in den besten Händen dem ich auch für so einige lustige Abende danke. Interessante und hilfreiche Diskussionen habe ich mit Kevin Pagel geführt, und viel über Moleküle aus Sicht eines Chemikers gelernt. Des Weiteren habe ich den Geschmack von Chicorée und Ale zu schätzen gelernt.

Andre Fielicke hatte das Vergnügen, mit mir in einem Büro zu sitzen und sich die ein oder andere Frage gefallen lassen zu dürfen. Andre: Biomoleküle sind großartig. Andreas Osterwalder hat die genialste Konferenz organisiert, an der ich teilgenommen habe. Gerne wieder. Das mit dem Business klappt auch noch. Henrik Haak hat sich geduldig meine Konstruktionsfragen angehört und konnte mir gute Ideen liefern. Wolfgang Erlebach hat dann geguckt, dass die Konstruktionen auch zur Maßarbeit

Berlin, im Januar 2011

# Publikationsliste

- P. Kupser, J. D. Steill, J. Oomens, G. Meijer, and G. von Helden. IR spectroscopy of gas-phase $C_{60}^-$. *Phys. Chem. Chem. Phys.*, **10**, 6862-6866 (2008).

- K. Pagel, P. Kupser, F. Bierau, N. C. Polfer, J. D. Steill, J. Oomens, G. Meijer, B. Koksch, and G. von Helden. Gasphase IR spectra of *coiled coil* protein complexes. *Int. J. Mass Spectrom.*, **283**, 161-168 (2009)

- P. Kupser, K. Pagel, J. Oomens, N. C. Polfer, B. Koksch, G. Meijer, and G. von Helden. Amide-I and -II Vibrations of the Cyclic $\beta$-Sheet Model Peptide Gramicidin S in the Gas Phase. *J. Am. Chem. Soc.*, **132**, 2085-2093 (2010).

- F. Bierau, P. Kupser, G. Meijer, and G. von Helden. Catching Proteins in Liquid Helium Droplets. *Phys. Rev. Lett.*, **105**, 133402 (2010).

- M. Rossi, V. Blum, P. Kupser, G. von Helden, F. Bierau, K. Pagel, G. Meijer, and M. Scheffler. Secondary Structure of $Ac-Ala_n-LysH^+$ Polyalanine Peptides ($n=$ 5,10,15) in Vacuo: Helical or Not? *J. Phys. Chem. Lett.*, **1**, 3465-3470 (2010).

# Lebenslauf

Der Lebenslauf ist aufgrund datenschutzrechtlicher Gründe in der Online-Ausgabe nicht verfügbar.