Stockholm University

Infrared spectroscopy:
a tool for protein characterization

Chenge Li

ISBN 978-91-7649-407-3, pages 1-49

Printed by Holmbergs, Malmö 2016
Distributor: Department of Biochemistry and Biophysics, Stockholm University

To my loved ones

Abstract

Infrared (IR) spectroscopy detects vibrations of molecules, for example, proteins. The absorption of the peptide group gives rise to 9 characteristic bands in the infrared region, named A, B, I-VII, with a decreasing energy or wavenumber (cm^{-1}). Among the 9 bands, amide I, which is mainly caused by C=O stretching vibration, is most sensitive to backbone structure and environment, and can therefore be used for structural analysis. In this thesis, a membrane protein sarcoplasmic reticulum Ca^{2+}-ATPase (SERCA1a) and a self-assembling peptide were studied by IR spectroscopy.

In the first two papers, IR spectroscopy was used to assess the quality of a recombinant SERCA1a. A yeast-based expression system was applied to express recombinant SERCA1a, and the reaction cycle as well as the structure was analyzed using IR spectroscopy. Different reaction intermediates were accumulated under different buffer conditions upon the release of ATP. The results showed that the recombinant protein exhibited similar IR features compared to the native protein. However, two SERCA1a preparations showed a difference around 1640 cm^{-1} in the amide I region. Using curve fitting, the band was assigned to β structure, and further investigation indicated that the difference in this region originated from protein aggregation. In the third paper, a co-fitting approach was tested and showed to be a more reliable method for structural analysis, and that it can be applied in biological IR spectroscopy. In the fourth paper, a peptide was computational designed and was predicted to self-assemble into amyloid fibrils, the formation of the fibril was confirmed with both electron microscopy and X-ray diffraction. IR spectroscopy was used to analyze further the structural details and the results supported our structural predication.

List of publications

I. Kumar S*, **Li C***, Montigny C, le Maire M, Barth A (2013) Conformational changes of recombinant Ca^{2+}-ATPase studied by reaction-induced infrared difference spectroscopy. *FEBS J*. 280:5398-407

II. **Li C**, Kumar S, Montigny C, le Maire M, Barth A (2014) Quality assessment of recombinant proteins by infrared spectroscopy. Characterisation of a protein aggregation related band of the Ca^{2+}-ATPase. *Analyst.* 139:4231-40

III. Baldassarre M, **Li C**, Eremina N, Goormaghtigh E, Barth A (2015) Simultaneous fitting of absorption spectra and their second derivatives for an improved analysis of protein infrared spectra. *Molecules.* 20:12599-622

IV. Kaltofen S, **Li C**, Huang PS, Serpell LC, Barth A, André I (2015) Computational de novo design of a self-assembling peptide with predefined structure. *J Mol Biol*. 427: 550-62

* These authors contributed equally

Abbreviations

ATR	Attenuated total reflection
Caged ATP	P^3-1-(2-nitro)-phenylethyladenosine 5'- triphosphate
DDM	n-dodecyl-β-D-maltopyranoside
FTIR	Fourier transform infrared
IR	Infrared
MIR	Mid infrared
SERCA	Sarco/endoplasmic reticulum calcium ATPase
SR	Sarcoplasmic reticulum

Table of contents

Abstract 5

List of publications 6

Abbreviations 7

Table of contents 8

1. Introduction 10

1.1 Biological membranes 10

1.2 Protein structure 11

1.3 Membrane proteins 12

1.4 P-type ATPases 13

1.4.1 Overview 13

1.4.2 Ca^{2+}-ATPase 13

1.5 Amyloid fibrils 17

1.5.1 Introduction 17

1.5.2 Characteristics of amyloid fibrils 18

1.5.3 Amyloid fibrils in Materials Science 19

2. Biophysical Methods 20

2.1 Infrared (IR) spectroscopy 20

2.2 Fourier transform infrared (FTIR) spectroscopy 22

2.3 IR spectroscopy of proteins 23

2.4 Sampling techniques 25

2.4.1 Transmission measurements 25

2.4.2 Reaction-induced difference FTIR spectroscopy 26

2.4.3 ATR FTIR spectroscopy 27

2.5 Resolution enhancement 28

2.6 Curve fitting 29

3. Aims and Objectives 31

4. Results and discussion 32
Paper I 32
Paper II 33
Paper III 34
Paper IV 35
5. Summary 36
6. Future plans 37
Acknowledgements 38
Sammanfattning på svenska 41
References 43

1. Introduction

1.1 Biological membranes

Biological membranes are essential to life as they are responsible for the compartmentalization that defines cells and organelles. In addition to that, membranes also play a role in cell division, biological reproduction, intracellular trafficking, as well as cell signaling[1].

The basic unit of a biological membrane is a lipid bilayer formed by lipids, which are molecules with a polar head group and one or more hydrophobic tails (Figure 1). The tails are usually fatty acids, which can differ in length and degree of saturation. Lipids organize into a bilayer with head groups along the surfaces and the acyl tails forming the hydrophobic core in between. This arrangement is driven by the hydrophobic effect where hydrophobic tails are shielded from the surrounding aqueous environment. The hydrophobicity of the core structure makes the membrane impermeable to ions and other hydrophilic substances, the transport of which requires channels and transporters.

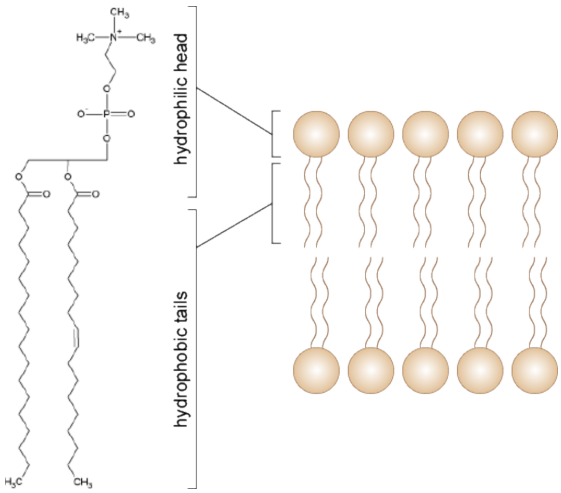

Figure 1: Structure of a lipid bilayer composed of phosphatidylcholine lipid[2].

1.2 Protein structure

Proteins are relatively large, compact, structurally complex molecules that are made from small molecules called amino acids. Amino acids share a common structure, which contains an amino group ($-NH_2$), a carboxyl group (-COOH), along with a side chain that is specific for each amino acid. The side chain can be hydrophobic, polar, positively or negatively charged, depending on its composition, structure and properties. In total, there are mainly 20 natural protein-building amino acids, which combine into peptide chains via peptide bonds (-CO-NH-). To be able to perform biological functions, proteins have to fold to defined and functional structures, driven by for example the hydrophobic effect, hydrogen bonding, ionic interactions and van der Waals forces.

There are four levels of protein structure, primary, secondary, tertiary and quaternary (Figure 2)[3]. The specific amino acid sequence of a protein is called primary structure. The backbone atoms of amino acid chains interact with each other through hydrogen bonding, forming stable patterns known as secondary structure. Most common secondary structures are α-helices and β-sheets. An α-helix is a coiled conformation where every –NH group of the backbone forms a hydrogen bond with a C=O group that is 4 residues away. In a β-sheet, the -NH group is hydrogen bonded with the C=O group from a nearby strand. Depending on the relative direction of the strands, β-sheets are further divided to parallel and anti-parallel structures. Tertiary structure refers to the three-dimensional structure of a monomeric protein and quaternary structure is used to describe proteins that are composed of multiple subunits.

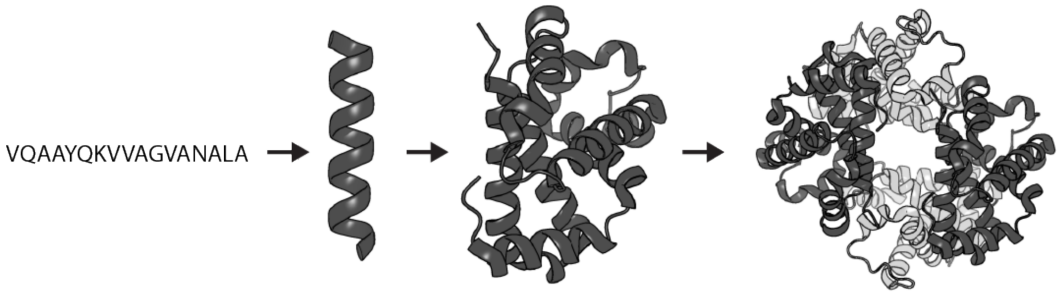

Figure 2: Levels of structure in proteins. From left to right: primary, secondary, tertiary and quaternary structure.

1.3 Membrane proteins

Membrane proteins are proteins that interact with membranes. Depending on how the protein interacts with membranes, it is classified as peripheral or integral membrane protein. Peripheral membrane proteins loosely associate at the surface of membranes and can be solubilized and removed from the membrane with a mild treatment. Integral membrane proteins are proteins that interact closely with nearby lipids and proteins, and cannot be removed without disrupting the membranes. Integral proteins perform a variety of roles in the cell such as serving as pumps, receptors and channels.

There are mainly two basic membrane-spanning structures for integral membrane proteins: α-helix-bundles and β-barrels (Figure 3)[4]. These two basic architectures reflect the need to form hydrogen bonds when buried in a hydrophobic lipid bilayer with no water around; the hydrogen bonds are formed within the protein backbone itself and α-helix-bundles and β-barrels are the two ways to do this[4-6]. Alpha-helical bundles have two or more transmembrane helices that span the membrane, and they are more or less uniformly hydrophobic. They are found in all cellular membranes and have diverse functions. β-barrels are cylindrical structures that form from several β-strands, they are less hydrophobic compared to an α-helical bundle, with polar and charged residue projecting into a central pore. They are only found in the outer membrane of Gram-negative bacteria, mitochondria and chloroplast, and they serve mostly as passive diffuse pores for small molecules [7].

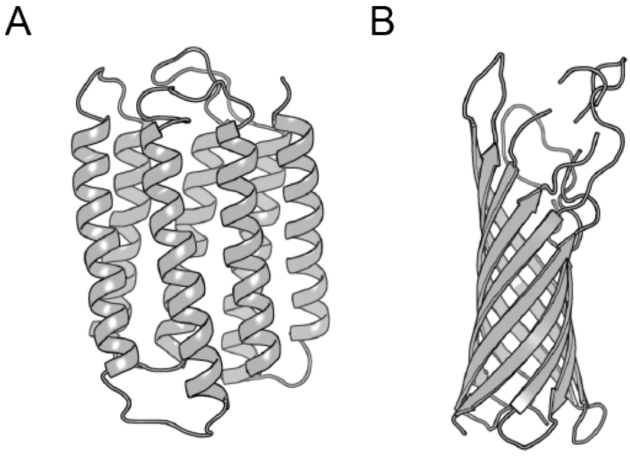

Figure 3: Membrane spanning structures (A) Bacteriorhodopsin (PDB 2BRD) has 7 transmembrane helices[8]; (B) OmpA (PDB 1BXW) has 8 transmembrane strands[9].

1.4 P-type ATPases

1.4.1 Overview

P-type ATPase family is a family of ion pumps that transfer cations across membranes using the energy released by ATP hydrolysis[10-12]. P-type ATPases are characterized by the formation of a phosphorylated enzyme intermediate that forms through the transfer of the γ-phosphate of ATP to a conserved aspartate residue. Na^+K^+-ATPase, a member of P-type ATPases, was first discovered by Jens Skou in 1957 from crab legs[13]. In subsequent years, other ion pumps with similar characteristics were found, and they mainly include Ca^{2+}, H^+, K^+, proton and heavy metal (e.g. Cu^+) transporting ATPases [14-17]. They play an important role in many fundamental processes in biology, including generation of membrane potentials, muscle relaxation and removal of toxic ions from cells[18-20].

1.4.2 Ca^{2+}-ATPase

Role in muscle relaxation

In muscle cells, free Ca^{2+} is stored, in mM concentrations, in a special form of endoplasmic reticulum (ER) called sarcoplasmic reticulum (SR). Sarco/endoplasmic reticulum Ca^{2+}-ATPase (SERCA) locates in the SR membrane and accounts for 70-80 % of the total protein in SR membrane[21]. The main function of SERCA is to pump Ca^{2+}, which is released to the cytosol during muscle contraction, against its gradient back to the SR lumen so that the muscle can relax.

There are more than 10 isoforms of SERCA, and they are encoded by three separate genes: *ATP2A1* (*SERCA1*), *ATP2A2* (*SERCA2*) and *ATP2A3* (*SERCA3*)[22-25]. The diversity of SERCA isoforms is enhanced through alternative splicing of the transcripts. Although sharing significant structural and sequence similarities, these isoforms exhibit strong tissue and development specificity. This thesis focuses on the role of SERCA1a in adult skeletal muscle.

The contractile force of a muscle is generated through the interaction of two proteins, myosin and actin. In muscle cells, molecules of myosin arrange and form structures called thick filaments. Molecules of globular actin, together with two regulatory proteins, troponin and tropomyosin form the thin filaments. Muscle contraction happens when the thin and thick filaments slide past each other.

Calcium ions are an essential controlling factor in muscle contraction, and the Ca^{2+} concentration is regulated by SERCA[26,27]. At rest, the two regulatory proteins, tropomyosin and troponin, bind to the actin thin filaments in a way that blocks the attachment of the myosin head group to actin[27]. When Ca^{2+} is released from the SR after a nerve impulse, the Ca^{2+} concentration increases about 100 fold. Calcium binds to troponin, causing a conformational change of the troponin-tropomyosin complex, which exposes the myosin-binding site on the thin filament. The filaments then slide past each other and the muscle contracts. As long as the binding sites remain exposed, the cycle is repeated. To relax the muscle, SERCA actively transports Ca^{2+} back into the SR, allowing troponin-tropomyosin back to its original conformation, covering the myosin-binding site, therefore ceasing the contraction cycle and relaxing the muscle. Because SERCA is abundant in the SR membrane, Ca^{2+} uptake only takes milliseconds, allowing a fast relaxation[22].

Mutations in *SERCA1* are reported to be associated with Brody´s disease. Dr. Brody described this disease first time in 1969[28]. The main syndrome is that patients suffer from muscle cramping and exercise-induced impairment of muscle relaxation due to decreased activity of SERCA1a. This disease, however, is genetically heterogeneous, which means mutations in other genes other than SERCA1 can also result in this disease[23,29].

Structure

SERCA1a is composed of a single polypeptide chain that contains 994 amino acids and has a molecular mass of 110 kDa[30,31]. The first three-dimensional (3D) structure of SERCA1a was solved by Toyoshima *et al.* in 2000[31]. The 3D structure shows a 10 helix (M1-M10) transmembrane domain and 3 globular domains extending to the cytoplasmic side: the N (nucleotide-binding) domain which has the ATP binding site;

the P (phosphorylation) domain where the phosphorylation site Asp351 is located; and the A (actuator) domain that alters Ca^{2+} affinity through structural changes.

Similar to other P-type ATPases, SERCA shares the classical E1/E2 theory, where E1 states have a high Ca^{2+} binding affinity and E2 states have a low Ca^{2+} binding affinity. In the E1 state, the high affinity Ca^{2+} binding sites are accessible from the cytoplasmic side, two Ca^{2+} bind to the E1 form (Ca_2E1) before the enzyme gets phosphorylated at Asp351 (Ca_2E1P). This is followed by a series of conformational changes to the low Ca^{2+} affinity form (E2P), Ca^{2+} is then released to the luminal side of the membrane, the cycle is completed when the enzyme undergoes dephosphorylation and returns to the E1 state (Figure 4).

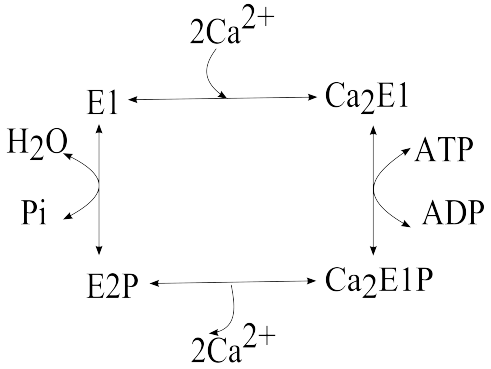

Figure 4: A simplified reaction scheme of the SERCA catalytic cycle[21,25].

SERCA1a has been extensively characterized, crystals were obtained in the presence and absence of Ca^{2+}, with different substrate and/or product analogues, possibly mimicking most of the main conformational states of the protein cycle and making it possible to follow the structural changes that take place during transport (Figure 5)[31-38].

In Ca_2E1, which represents the state after binding of two Ca^{2+} ions, the three cytoplasmic domains adopt an open arrangement with no evident relation between the N and P domain[31]. The two Ca^{2+} binding sites are located near the middle of the membrane bilayer and are 5.7 Å apart[31].

To mimic the ATP and Ca^{2+} bound state Ca_2E1ATP, ATP analogues, 5'-β,γ-methylene-triphosphate (AMPPCP) and adenosine 5'-β,γ-imido-triphosphate (AMPPNP), adenosine diphosphate plus aluminum fluoride (ADP-AlF_4^-) were used[32-

[34]. The conformations of SERCA1a crystalized in the presence of Ca^{2+} and different ATP analogues are similar and they show that binding of the nucleotide causes large and global changes in protein structure, especially in the cytoplasmic domains[34]. They are now reorganized to form a compact headpiece. The ATP analogues bring the N and P domains very close by directly bridging them. The N domain is 90^{o} inclined towards the P domain and the A domain rotates. The conformational changes pull the transmembrane helices M1 and M2, and close the cytosolic entrance for Ca^{2+}, thereby preventing backflow of Ca^{2+} ions[32]. After phosphorylation, the nucleotide no longer bridges the interface between the N and P domain.

In the E1P-E2P transition, the transmembrane helices undergo rearrangement, revealing a wide-open, funnel-shaped Ca^{2+} exit pathway, with Ca^{2+} binding residues now facing the lumen[34,39]. The change of conformation of the liganding residues reduces the Ca^{2+} binding affinity so that the two Ca^{2+} ions can efficiently be translocated[34].

All crystal structures in the absence of Ca^{2+} to mimic the E2 state are with inhibitors that stabilize the transmembrane domain, as the ATPase is unstable in the absence of Ca^{2+} [40]. The crystal structures with thapsigargin (TG) and 2,5-di-*tert*-butyl-1,4-dihydroxybenzene (BHQ) are essentially the same[35,41]. Three cytoplasmic domains gather together to form a single headpiece, the arrangement of which is distinctly different from that of the crystals mimicking Ca_2E1ATP. The release of phosphate induces a conformational change that closes the lumenal gate, and the top amphipathic part of M1 forms a part of a cytoplasmic access funnel leading to one of the Ca^{2+} ligands[42]. A crystal obtained with another inhibitor, cyclopiazonic acid (CPA), however, reveals a different headpiece orientation, which could mean that E2 has considerable plasticity[43]. From the E2 state, the enzyme progresses to E1, Ca^{2+} binds, and another cycle can start[19].

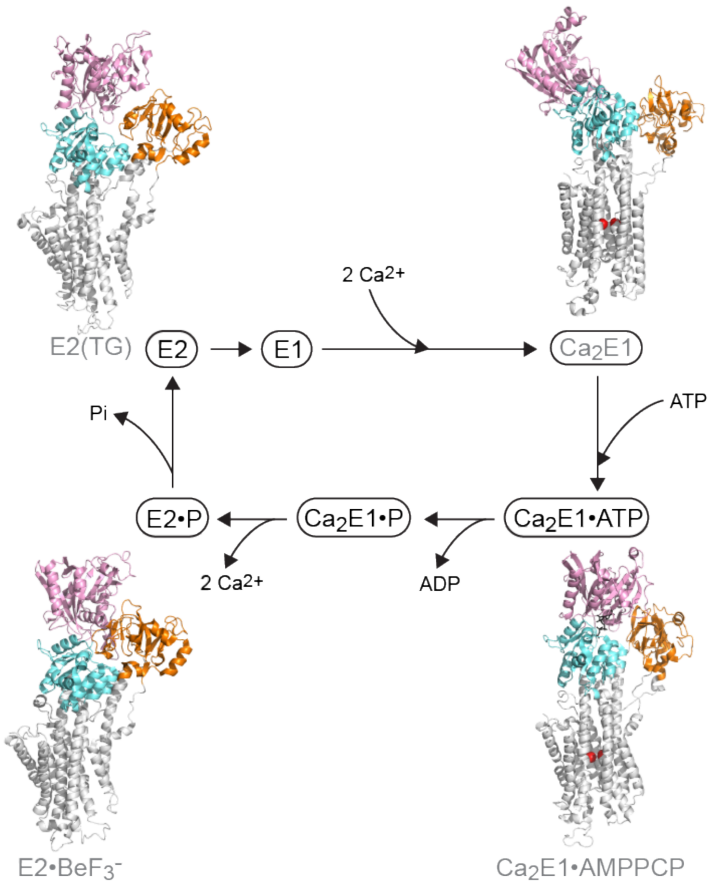

Figure 5: Comparison of SERCA1a structures representing key states of the reaction cycle, based on the PDB IDs: 1SU4 (Ca_2E1), 1VFP (Ca_2E1•AMPPCP), 3B9B (E2•BeF_3^-), and 1TWO (E2•TG)[31,33-35]. Shown with the N-domain in pink, A-domain in orange, P-domain in blue, the transmembrane domain in gray and Ca^{2+} as red spheres.

1.5 Amyloid fibrils

1.5.1 Introduction

The term "amyloid" was first used in the 1800´s by botanist Schleiden to describe starch, and then by pathologist Virchow to describe the small round deposits in the nervous system that showed the same iodine color reaction as starch. The assignment regarding the latter shifted from carbohydrates to proteins when the high nitrogen content was established later[44-46].

Proteins or peptides can convert from its soluble form to insoluble, highly ordered amyloid fibrils, the deposition of which is central to the pathology of more than 20 diseases *in vivo* including Alzheimer´s disease and Type 2 diabetes[47-51]. The

traditional histopathological definition of an amyloid is an extracellular, proteinaceous deposit characterized by apple green birefringence (Figure 6B) stained with Congo red and viewed under polarized light. A broader biophysical definition includes any polypeptide that polymerizes to form a cross-β structure (see below), whether *in vivo* or *in vitro*[52].

1.5.2 Characteristics of amyloid fibrils

Amyloidogenic proteins show no obvious sequence similarity, nor do their native structures look alike. However, the structures of amyloid fibrils are remarkably similar. Under an electron microscope, amyloid fibrils are of micrometer length, linear, non-branching, with a diameter between 70 and 120 Å (Figure 6A). The specific binding of Congo red to an amyloid produces an apple-green color under polarized light (Figure 6B)[53,54]. X-ray diffraction shows that amyloid fibrils share a typical cross-β pattern. This pattern was first discovered in 1968, and it consists of a strong meridional reflection at 4.7 Å and a weak equatorial reflection at 10 Å (Figure 6C)[55]. This is characteristic for a structure with β-sheets parallel to the fibril axis, and the strands forming those β-sheets perpendicular to the axis[56]. The intense meridional diffraction around 4.7 Å derives from the ordered hydrogen-bonded β-strands that are perpendicular to the fibril axis, and it reflects the spacing between strands from the same β-sheet; the weaker equatorial reflection around 10 Å reflects the spacing between β-sheets in the amyloid fibrils[48,57-59].

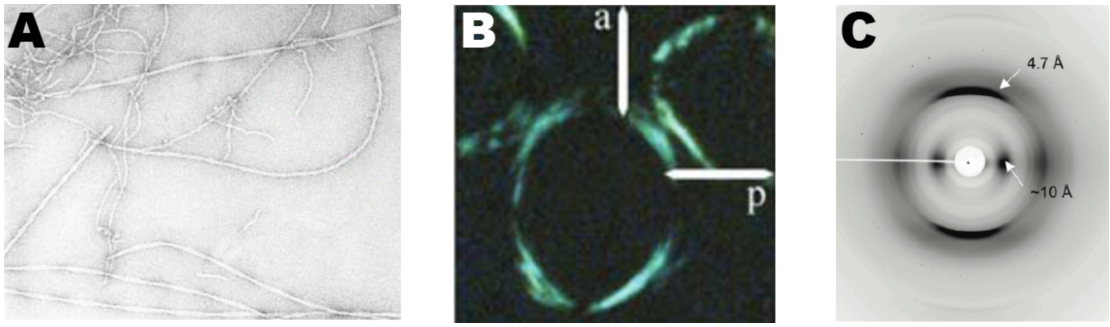

Figure 6: Images of amyloid fibrils (A) transmission electron microscope image of amyloid β-protein (Aβ) fibrils[60]; (B) Sections of kidney stained with Congo Red in polarized light. Amyloid appears green against a black background[53]; (C) X-ray diffraction pattern of synthetic amyloid fibrils made from Aβ peptide[46].

1.5.3 Amyloid fibrils in Materials Science

Amyloid fibrils are not only extremely stable and resistant to degradation, but their nanoscale dimensions, efficient self-assembly into a well-defined structure make them a good candidate for functional nanomaterial construction[46,61-63]. Amyloid fibrils have been reported to function as nanowires elements. Self-assembling of a Phe-Phe peptide was used to form a template for silver nanowires[64]. Coated with metals, fibrils show a similar conductive property as solid conducing wires[65]. Amyloid fibrils also have potential as a biofunctional material by conjugating it with functional proteins or enzymes[62]. Fusing cytochrome to a self-assembling peptide was shown to form an amyloid fibril with functional cytochrome on the surface [66].

2. Biophysical Methods

2.1 Infrared (IR) spectroscopy

Spectroscopy studies the wavelength-dependent interaction of molecules or atoms with electromagnetic radiation. In some experiments, the electromagnetic radiation behaves like particles called photons. Each photon has a wavelength-dependent energy. The energy can be described using the Bohr equation $E=h\nu$, where h is Planck´s constant ($h= 6.63 \times 10^{-34}$ J s) and ν is the frequency of radiation.

There is an enormous difference between the energies of photons from different parts of the spectrum (Figure 7). For example, photons in the UV and visible range have such a high energy and they can induce electronic transitions, while photons in the lower frequency range have less energy, for example, photons in the IR region can only induce changes in molecular vibrations or rotations.

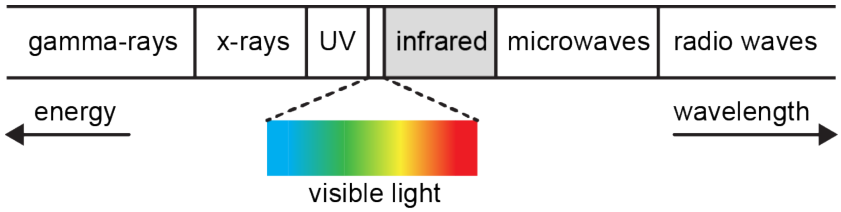

Figure 7: Wavelength ranges of the electromagnetic spectrum. Energy is proportional to the frequency of radiation, and inversely proportional to the wavelength.

Most of the vibrational energies fall into the IR region of the spectrum, the typical wavelength of which is between 0.78 and 250 µm. In IR spectroscopy, wavenumber is normally used instead of wavelength because it has the advantage of being proportional to the energy. The unit of wavenumber is cm^{-1}. The IR spectrum is further divided into three regions: the near IR region, which corresponds to a wavelength between 0.78 to 2.5 µm (12820-4000 cm^{-1}), the mid IR region, which corresponds to a wavelength between 2.5 and 25 µm (4000-400 cm^{-1}), as well as the far IR region, which has a wavelength between 25 and 250 µm (400-40 cm^{-1}).

In the simplest case, the 2-atomic oscillator, the vibrational frequency ν of the oscillator follows the expression

$$\nu = \frac{1}{2\pi}\sqrt{\frac{k}{\mu}}$$

where k is the force constant and μ is the reduced mass. A larger mass and/or a weaker bond give rise to a lower vibrational frequency, thus a lower wavenumber (Figure 8).

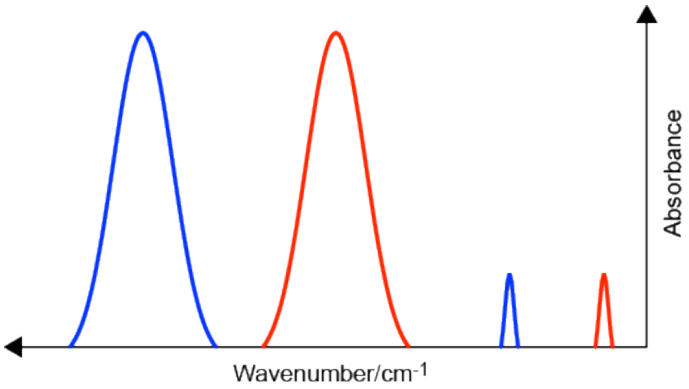

Figure 8: Effect of molecular mass on the vibrational frequency. Blue: smaller mass, Red: larger mass.

In a molecule with several atoms, different types of vibrations are distinguished. For example, stretching vibrations corresponds to bond length changes, and bending vibrations correspond to bond angle changes (Figure 9). Vibrations from neighboring groups are often coupled if the frequencies are similar.

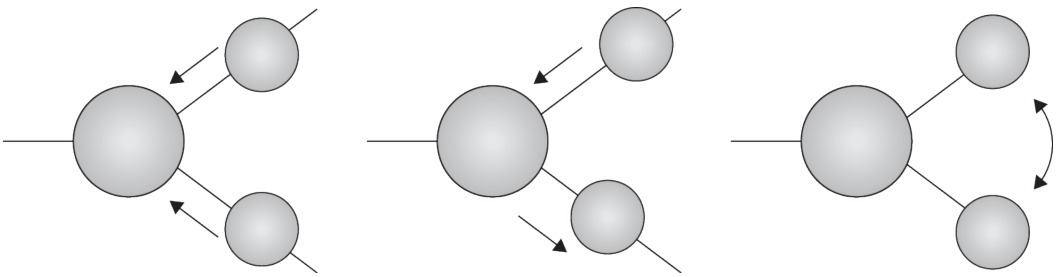

Figure 9: Examples of vibrations. From left to right: symmetric stretching vibration, antisymmetric stretching vibration, bending vibration.

2.2 Fourier transform infrared (FTIR) spectroscopy

In a traditional dispersive IR measurement, a diffraction grating is used to separate a single wavelength of light and directs it to the detector. Each spectral element is measured one at a time, therefore it is a slow process. A major breakthrough in IR technology was the introduction of FTIR spectrometers; the way how the sample spectrum is recorded is fundamentally different from dispersive instruments in how the IR beam gets from the source to the detector (Figure 10)[67]. The first commercial FTIR instruments became available in the late 1960s and today practically all IR spectrometers are FTIR instruments[68]. Their major advantages over traditional, dispersive IR instruments include that they provide a high-speed data collection due to simultaneous measurement over the whole wavelength range, a much better signal to noise ratio because more energy reaches the detector, and frequency precision is ensured by the use of a laser[67].

A typical FTIR spectrometer consists of an IR-source, a laser, an interferometer and a detector (Figure 10). The interferometer contains a beam splitter and two mirrors. The beam from the IR-source is first split into two parts, which are then reflected by two different mirrors. Half of the beam meets the mirror that is fixed at a certain distance away from the splitter and gets reflected. The other half of the beam meets a moving mirror and gets reflected. 50% of the beam reflected by both mirrors is recombined at the beam splitter, and then passes through the sample and reaches the detector, while the other 50% is reflected back to the light source. When the two beams reflected by the mirrors recombine, they have travelled different distances, which leads to constructive or deconstructive interference, which creates an interferogram. The interferogram is back-transformed into an IR spectrum by the computer. Using a laser with a defined wavelength (normally a helium-neon laser) ensures the accuracy of the determination of wavelength.

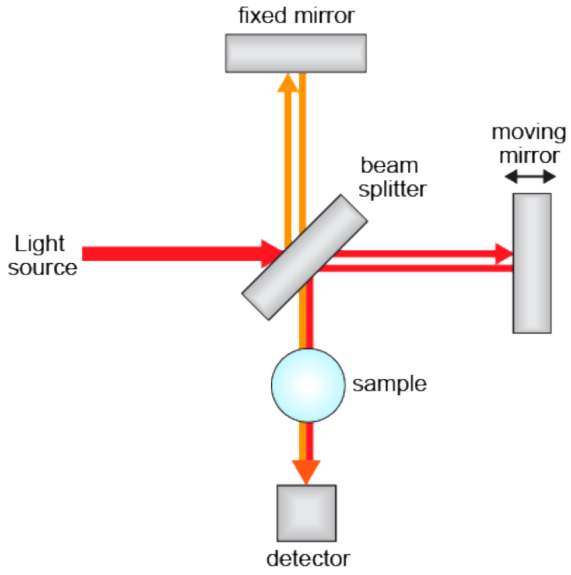

Figure 10: A scheme of a FTIR spectrometer [69]

The most common detectors for IR are Deuterated Triglycine Sulfate (DTGS) and Mercury Cadmium Telluride (MCT) detectors[70]. DTGS detectors are stable and operate at room temperature, with the drawback of relative slow response times. MCT detectors, on the other hand, must be cooled with liquid nitrogen to work properly, but offer better sensitivity and have an about 10 times faster response time compared to DTGS detectors, so that spectra can be recorded much faster.

2.3 IR spectroscopy of proteins

IR spectroscopy has been used to study biological molecules, for example proteins or peptides. It can not only provide information about protein structure, environment of the protein backbone and side chains, but also be used for dynamic studies.

The absorption of a peptide group gives rise to 9 characteristic bands named amide A, B, I though VII (Table 1), among which amide I and II are the ones most related to protein backbone structure[71-73]. The amide I band mainly stems from the absorption of the C=O stretching vibration (approximately 80 %), together with small contributions from for example the C-N stretching vibration. It has the most intense absorption and is the prime subject for protein structural analysis. The amide I band is normally found in the 1700-1600 cm^{-1} region and its exact position is determined by the geometry of the backbone and by the strength of hydrogen bonding. The amide II band occurs in

the 1580-1510 cm^{-1} range, where the N-H bending vibrations and the C-N stretching vibrations couple to form the amide II normal modes that absorb in this region. Although also sensitive to backbone, the relationship between structure and band positions is more complex than for the amide I band and therefore not used much for structural analysis.

Table 1: Characteristic infrared bands[73,74]

	Approximate wavenumbers (cm^{-1})	Description
Amide A	3270-3310	NH streching
Amide B	3030-3100	NH streching
Amide I	1600-1700	C=O streching
Amide II	1510-1580	CN streching, NH bending
Amide III	1200-1400	CH streching, NH bending
Amide IV	625-767	OCN bending
Amide V	640-800	Out-of-plane NH bending
Amide VI	537-606	Out-of-plane C=O bending
Amide VII	200	Skeletal torsion

A large number of experimental and theoretical work on synthetic polypeptides and data from mode analysis of numerous structures are the foundations for establishing the basis of secondary structure determination (Table 2)[73,75]. The average wavenumber of an α-helical structure in the amide I region is around 1654 cm^{-1}. The exact number is affected by many factors, for example, the number of residues involved in the helical structure and the neighboring environment. For a β-sheet structure, the average wavenumber of the main component band is around 1630 cm^{-1}, with the antiparallel β structure having an additional higher wavenumber component around 1684 cm^{-1}, the intensity of which is about 1/10 of that of the lower wavenumber component. The exact wavenumbers are influenced by strand length, number of strands, orientation as well as other factors.

The amide I region is overlapped with the O-H bending vibration from H_2O. Therefore the IR measurements often require a higher protein concentration (usually above 10 mg/ml), or prepare the sample in D_2O, of which the bending vibration of O-D is around 1200 cm^{-1} and there is little overlap of D_2O absorption and amide I band[76,77].

Table 2: Spectral ranges of several secondary structures in the amide I region[78,79]

Secondary structure	Band positions in H_2O /(cm^{-1})	
	average	range
α-helix	1654	1648-1657
β-sheet	1633	1623-1641
	1684	1674-1695
Turns	1672	1662-1686
Random coil	1654	1642-1657

The most important operational advantage of IR spectroscopy is that spectrum of almost any biological material can be obtained in a wide range of conditions[80,81]. Proteins can be studied in solution, organic solvents, detergent micelles, lipid membranes, and crystals, etc[81]. It also has the advantage that it is technically simple and inexpensive, requires small protein amounts, and that the size of the protein is not important.

2.4 Sampling techniques

2.4.1 Transmission measurements

In mid IR (MIR) spectroscopy for studying proteins, BaF_2 and CaF_2 cuvettes are commonly used because they have no or little absorbance in the mid IR region. Our cuvettes consist of two windows, a flat window and a window with a trough in the middle, together forming a well with a certain optical pathlength, which can vary a lot. For aqueous samples, short pathlength windows (around 5 μm) are preferred, because H_2O, whose absorption overlaps with the amide I region, has a high extinction coefficient. The short pathlength lowers the absorption of water, but also limits the intensity of the bands from the protein amide I region and the signal-to-noise ratio at a given protein concentration[82]. Using D_2O, the pathlength can be increased to around 50 μm. Because of the small cuvette volume, the sample volume is often quite small, 1-5 μl. However, the measurements often require a high protein concentration.

2.4.2 Reaction-induced difference FTIR spectroscopy

Proteins normally contain a large number of vibrations, which makes the IR spectrum very crowed and complicated. One strategy to overcome this problem is to use difference spectroscopy, where only the groups that participate in a reaction are observed. This is particularly suited to investigate the molecular mechanism of a protein reaction. Normally a protein sample is prepared in state A and the absorbance spectrum of this state is recorded. A reaction is then initiated, the protein proceeds to state B and another absorbance spectrum is recorded. The resulting difference spectrum will originate only from the groups that participate in the reaction, while the others will cancel out and are not be visible in the difference spectrum (Figure 11). Therefore, difference spectroscopy exhibits details of the reaction mechanism despite a large background absorption. In the difference spectrum, negative bands are characteristic of the initial state A and positive bands reflect state B during or after the reaction.

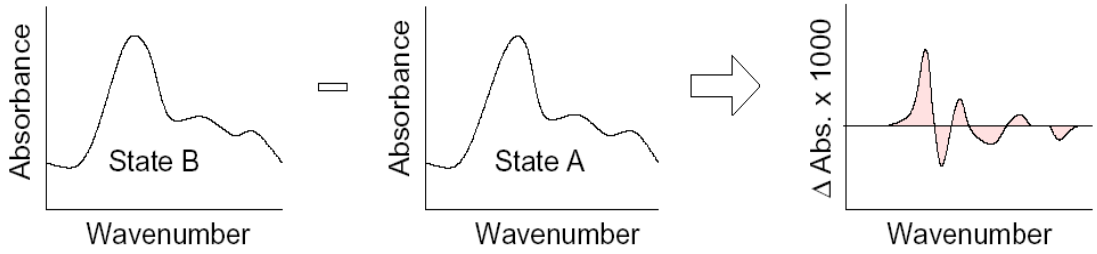

Figure 11: A difference spectrum is generated by subtracting a spectrum in state A from a spectrum in state B.

The reaction between the two protein states can be initiated by methods like temperature and pressure jump, light-induced difference spectroscopy, attenuated total reflection (ATR) with buffer exchange, and the photolytic release of effector substances from caged compounds[79]. A description of these methods can be found in several reviews[83-89]. The use of a photosensitive caged compound is applied here to trigger a reaction directly in the cuvette. The dominant class of caged compounds are 2-nitrobenzyl derivatives[90,91]. Upon illumination in the near-UV spectral range (300-350 nm), the compound of interest is released, which generates a concentration jump of the free effector substance (Figure 12). The binding of the substance to a protein

and the subsequent conformational changes alter the infrared spectrum. In the study of proteins, caged nucleotides such as caged ADP and caged ATP, caged Ca^{2+} and caged proton have been used[92-96].

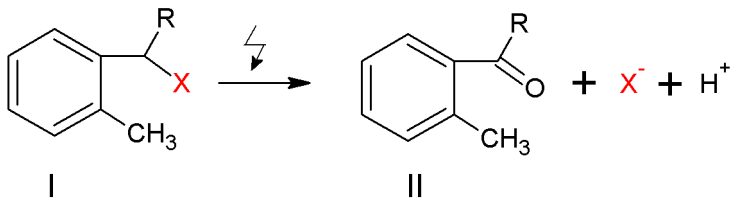

Figure 12: Photolysis of caged compounds. Caged compound I undergoes an internal redox reaction upon UV illumination, giving rise to a 2-nitrosobenzaldehyde related byproduct II and the photo-released effector molecule X^- and H^{+}[90].

Although a difference spectrum can contain a wealth of relevant information on the mechanism of a protein, it is often very difficult to interpret. This is because the assignment of each band to an individual group is not straightforward. Assignment of IR bands to specific chemical bonds can be achieved, for example by site-directed mutagenesis and by isotopic labeling studies[79]. Site-directed mutagenesis is a powerful approach, ideally an IR signal due to a specific amino acid is missing when it is selectively replaced. Isotopic labeling uses the mass effect on the vibration frequencies (Figure 8) and introduces band shifts, which can help identify the absorption of labeled groups. It is a simple but yet powerful tool to regard the spectra as a characteristic fingerprint of the conformational change without attempting a molecular interpretation. The signature of a conformational change in the spectrum can then be used to detect and define transient states of a protein.

2.4.3 ATR FTIR spectroscopy

An alternative to transmission spectroscopy is the use of ATR (Figure 13). In an ATR experiment, a sample is placed on top of a crystal that has a higher refraction index, for example, diamond, ZnSe or Ge. A beam of IR light passes through the crystal in a way that it is totally reflected at the interface between sample and crystal. Upon reflection at the interface, an evanescent wave is formed and it penetrates into the

sample, carrying information about the sample. IR radiation is further reflected one or several times in the crystal before reaching the detector.

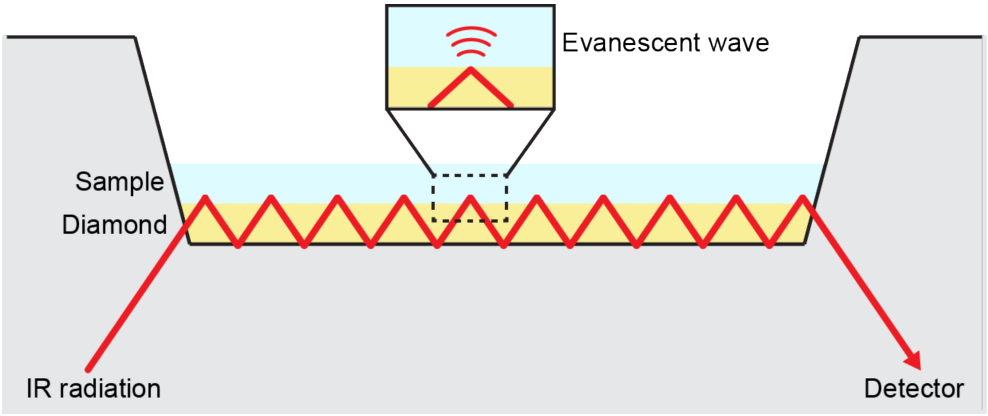

Figure 13: Scheme of a typical ATR setup.

The advantage of ATR over transmission methods is that the sample conditions can be manipulated directly since one side of the crystal is accessible. For example, ligands can be directly added or the pH can be changed easily. For protein studies, a thin protein film is often formed on top of the crystal, normally simply by drying. The advantage of drying is that a clear protein spectrum with good signal to noise ratio can be obtained because there is no interference with water absorption. The drawback of the drying procedure is that in some cases, proteins will change their structure or function upon drying.

2.5 Resolution enhancement

The amide I region of proteins is composed of multiple bands which are characteristic for secondary structures, they overlap with each other and form a broad unstructured band. One way to better separate overlapping bands is to calculate the *n* th derivative (mostly commonly used is the second derivative) of an absorption spectrum. The line width of the second derivative band is smaller than that of the original absorption band, therefore the second derivative can be used to resolve overlapping bands (Figure 14). The minima of the second derivative give the positions of the overlapping components. Note that this is not a real improvement of the spectrum resolution, but an increase of the degree of separation by narrowing the bandwidth of individual components for easier visualization[97].

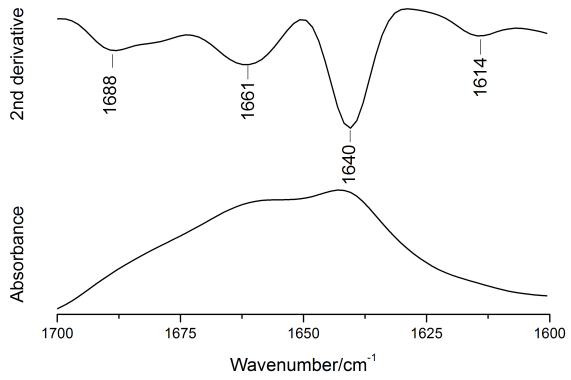

Figure 14: Spectrum enhanced by second derivative. 2nd derivative (top) and absorption (bottom) spectrum of Ribonuclease A.

2.6 Curve fitting

To estimate the secondary structure content of a protein, not only the position, but also the intensity and width of component bands underneath a complex band profile need to be characterized. Curve fitting is an approach that decomposes the broad unstructured band into its component bands, allowing to obtain information about the band parameters. The first comprehensive curve fitting study of a series of proteins of known structures was reported in 1986 by Byler and Susi and the agreement with X-ray data was remarkble[98]. In this approach, component bands are placed at positions that were determined from the peak positions of a resolution enhanced spectrum, while the intensity, line shape (Gauss or Lorentz), and band width are allowed to vary in order to have a minimal difference between the experimental spectrum and the fitted spectrum[98]. Curve fitting now is the most common approach to obtain secondary structure information. The results of curve fitting, due to large number of adjustable parameters that are involved in the procedure, are not necessarily unique.

The standard fitting approach is to fit absorption spectra. However, a good match between the fitted and experimental spectra is no guarantee for a good match between their resolution-enhanced spectra.

An alternative approach is to fit the resolution-enhanced spectrum instead[99]. In this approach, the derivatives of component bands are fitted to the experimental derivative spectrum. This approach has also been questioned mainly for two reasons. Firstly, resolution-enhanced spectra are more affected by instrumental noise and water vapor contamination[97,100,101], but this problem can be eliminated by technical solutions such as purging the instrument with dry air. Secondly, the spectral derivation does not preserve the relative intensities of the absorption band, broader bands are much less visible compared to narrow bands[101,102].

3. Aims and Objectives

The first part of the thesis focuses on the study of SERCA1a using IR spectroscopy. SERCA1a is a member of the P-type ATPases and can serve as a model for P-type ATPases. Although crystal structures of most of the main states of SERCA1a are now available, there are still questions not answered. For example, how is the communication between the phosphorylation site and the Ca^{2+} binding sites when they are so far away (50Å), and what is the mechanism of proton countertransport? To answer those questions, study of SERCA1a mutants is of great interest. The aim of this project is to express and purify SERCA1a mutants in our lab and study it with time resolved FTIR spectroscopy.

An important application of IR spectroscopy is to study the secondary structure of proteins/peptides. The absorption spectrum is normally fitted with its component bands, the position of which is decided from second derivative analysis. Bandwidth, intensity and shape are allowed to vary to minimize the differences between the experimental and the fitted spectra. This approach was used in our ATPase study (Paper II), but because of the lack of constraints, the fitting result commonly are not unique, therefore we felt the need to improve the approach in order to have a higher reliability (Paper III).

4. Results and discussion

Paper I

Conformational changes of recombinant Ca^{2+}-ATPase studied by reaction-induced infrared difference spectroscopy

Kumar S*, **Li C***, Montigny C, le Maire M, Barth A

FEBS J. 280:5398-407 (2013)

* These authors contributed equally

Recombinant Ca^{2+}-ATPase was expressed in *Saccharomyces cerevisiae* with a biotin-acceptor domain and purified using avidin affinity chromatography. The purified protein was then reconstituted into vesicles. Different intermediate states of both recombinant and native rabbit SERCA1a were accumulated using different buffer conditions and their formation from the original state (Ca_2E1) was followed by reaction-induced IR spectroscopy. The difference spectra show that recombinant and native SERCA1a share the same features regarding the formation of intermediates in the catalytic cycle, which means that the recombinant protein behaves in the same way as the native protein with our method, and it can be applied, together with mutations, to study the important role of individual amino acids. The enzyme activity of the recombinant protein, however, is about half the value of the native SERCA1a.

Paper II

Quality assessment of recombinant proteins by infrared spectroscopy. Characterisation of a protein aggregation related band of the Ca^2-ATPase.

Li C, Kumar S, Montigny C, le Maire M, Barth A

Analyst. 139:4231-40 (2014)

Although recombinant and native SERCA1a behave the same regarding the formation of intermediates from the original state (Ca_2E1), the spectrum showed a different secondary structure of those two preparations in a region close to 1640 cm^{-1}. Instead of having a shoulder at 1642 cm^{-1} and a weak band at 1632 cm^{-1} as in the native protein, the recombinant protein had a single strong band at 1631 cm^{-1}. By using curve fitting, together with the information that is available from the crystal structure, the shoulder band at 1642 cm^{-1} was assigned to β-sheet structure. The origin of the difference in that region was investigated further. The native protein was incubated at room temperature and the structural change during incubation was followed. The result showed that longer incubation time made the IR spectrum more similar between the native and the recombinant protein. This indicates that the change of secondary structure is possibly caused by aggregation. A similar structural change was also observed when a protein sample was repeatedly dried on the ATR crystal at low ion concentrations. Another support for the aggregation hypothesis came from an experiment where the native protein was solubilized in detergent, and the soluble protein and protein aggregates were separated by high-speed centrifugation. The results show that the soluble protein has the same IR features as the native protein, while the protein aggregates had the same band characteristics as the recombinant protein.

Paper III

Simultaneous fitting of absorption spectra and their second derivatives for an improved analysis of protein infrared spectra.

Baldassarre M, **Li C**, Eremina N, Goormaghtigh E, Barth A

Molecules. 20:12599-622 (2015)

The amide I band often shows a complex, unresolved line shape. In order to identify and characterize the component bands underneath the amide I band, curve fitting is commonly applied. In this paper, we introduce co-fitting, which simultaneously fits an absorption spectrum and its second derivative. To test the effect of co-fitting, both synthetic spectra as well as the spectra from ribonuclease A, pyruvate kinase and aconitase were used. The results show that co-fitting can lead to considerable improvements in the accuracy and reliability of band parameters, and thus provides a better base for secondary structure quantification.

Paper IV

Computational de novo design of a self-assembling peptide with predefined structure.

Kaltofen S, **Li C**, Huang PS, Serpell LC, Barth A, André I

J Mol Biol. 427: 550-562 (2015)

Protein and peptide self-assembly is a powerful tool for engineering of new biomolecules. In this paper, a computational design principle was developed for a self-assembling peptide with predefined structure. The design protocol was to use a βαβ building block to self-assemble into a double sheet structure with layers of α-helix on the side. The formation of fibrils was conformed by electron microscopy. X-ray fiber diffraction showed that the fibril had the typical cross-β diffraction pattern common for amyloid fibrils. The structure was further checked using infrared and circular dichroism (CD) spectroscopy. The results support that the peptide adopts the designed structure.

5. Summary

The work presented in this thesis shows how one can use IR spectroscopy to study protein structure and dynamics. Compared to other biophysical methods like CD, NMR, Fluorescence, and X-ray diffraction, it has the advantage that it is an easy and fast technique, it requires small protein amounts, and the subject is not limited by size.

Papers I and II show the use of IR spectroscopy on a membrane protein, Ca^{2+}-ATPase (SERCA1a). In order to get a better understanding of the transition between different protein states, IR studies of protein mutants are of great interest. In order to do that, a yeast based expression system was established. The SERCA gene is connected with a sequence encoding a biotin-acceptor domain, which allows its auto-biotinylation during expression; the protein was then purified by a single avidin affinity chromatography purification process before reconstitution into a lipid membrane. The kinetics and structure of the protein product were studied by IR spectroscopy. The recombinant protein showed similar structural changes in the reaction cycle compared to native rabbit SERCA1a, however, the structure showed a slight difference in the amide I region of the absorption spectrum. Further investigation concluded that this structure differences originated from protein aggregation.

Paper III explores an improved curve fitting method for secondary structure analysis with IR spectroscopy. The conventional way is to fit only the absorption spectrum, whereas here we introduce an approach named co-fitting that fits the absorption and the resolution enhanced spectrum simultaneously. We tested this approach with both a synthetic spectrum, as well as 3 protein spectra with mixed α/β structure, pyruvate kinase, ribonuclease A and aconitase. The results show that co-fitting leads to a more reliable analysis of the protein secondary structure.

Paper IV shows that IR spectroscopy can be also used to study peptides. A synthetic peptide was designed and predicted to be able to self assemble into amyloid fibrils. Both electron microscopy and X-ray diffraction confirmed the formation of fibrils. The detailed structure of the fibrils was further investigated by IR and CD spectroscopy, and the results support our prediction.

6. Future plans

The study of SERCA1a mutants was limited by the low expression level. How to increase the yield is of great interest. We will also study several aspects regarding the coupling between transport and catalytic cites, and proton countertransport, for example:

The P domain is connected to the transmembrane helices M 6 and M7 through a loop L6-7. Several mutants on the M, P domain and also the connecting L6-7 loop have shown to be able to affect Ca^{2+} pumping. Effects of these mutants, especially their effect on the active site conformation around the phosphate group will provide a better understanding regarding the communication between the P and the transmembrane domains.

The transport of two Ca^{2+} from the cytosol to SR is accompanied by countertransport of 2-3 H^{+} from the opposite direction. Upon Ca^{2+} release, the protons bind to the empty Ca^{2+} binding sites and are then released to the cytoplasm when the enzyme is dephosphorylated. Sites for protonation can be identified from mutant studies.

Acknowledgements

Finally, I will, hopefully, soon be crowned with the titel “Doctor Li”, and that would not be possible without all the help and joy that I get from the past few years.

First of all I would like to thank you **Andreas** for accepting me as his PhD student. I did not know anything about infrared spectroscopy before coming to the lab (I did mention that in my CV), and you said infrared is your thing, you gonna help me with it. And you did, from the scratch. I really appreciate your patience and enthusiasm though the whole time.

Lena Mäler, my second supervisor, **Pia Ädelroth**, my mentor, and **Stefan Nordlund**, thank you for guiding me through all the PhD check points. And **Elzbieta Glaser**, together with Pia, thank you for the many constructive comments on my pre-dissertation.

My collaborators from France, **Cédric Montigny** and **Marc le Maire**, thank you so much for the help and useful discussions regarding the SERCA1a project; and thank you **Erik Goormaghtigh** for the help on the co-fitting paper; **Magnus Andersson** and **Harsha**, it is a pleasure collaborating with you, and a big thank you to Magnus for taking me to the synchrotron center; **Sabine** and **Ingemar André**, thank you for the collaboration on the peptide project.

To members of the AB group: **Saroj**, although our overlap time in the group was only two months, you showed me a lot things in the lab that makes my future work much easier; **Nadja and Maurizio**, my most long-lasting office and lab mates, it is always fun to work and talk to you, you guys are always so enthusiastic, so active, and always so helpful when I needed, I miss so much that whenever I have problem, I can just shout either of your name from the lab and you will show up the second after, with a facial expression saying “what now”; **Ljubica**, together with Nadja, we make the best fika/foosball/orange company; **Cesare**, **Jan** and **Beatrice**, the newer office mates, it is a pleasure to share the office with you; many more that joined the group in these years, thank you for making the work atmosphere so pleasant and friendly.

To the NMR gang: **Weihua** and **Biao**, thank you for being such good friends all the time, and taking care of my plants every time I take vocation, **Jobst**, thanks for the good times both in the corridor, and during the membrane course; **Potus**, thank you for having the patience always talking Swedish with me; **Jüri** and **Astrid,** thank you for the nice discussions on the peptide and the CD results; **Christian**, **Scarlet, Ida, Axel**, **Johannes,** it is always nice talking to you in the lunch room, you guys make the 3rd floor such a friendly and cozy place to work in.

Friends from RD group: **Dan,** my slightly penetration into Rob´s group starts with you, we did a lot things together, lunch and weekend´s shopping, it was so fun; **Hao**, always so energetic and helpful; **Diogo** and **Cata**: the coolest Portuguese couple, you guys are always so friendly, so nice to me, there is always this positivity and energy coming from you guys that I will definitely miss; **Henrik**, listening to your discussions about politics has made the lunch hour so much interesting; **Joppe**, first time meeting you was having lunch together with Dan, and then is more lunches, now we have dinners together ☺, you are always so supportive and nice to me, you believe in me when I doubt myself, a big ❤ goes to you; and **Rob**, thank you for the nice talks we had, and not kindly asking me to leave when I am in the office too often.

Beata and **Pedro**, thank you for the nice time we had during lunches and fika; **Tomas**, **Zhe**, **Anna**, **Grietje,** always fun talking to you, and Tomas, mein Freund, God I forgot how many times I asked you or Diogo to open the rotor cap for me; **Candan**, for teaching me about lipids and also for the good time we had in the lab; **Rickard**, friend since the spectroscopy course, and that was 2011; **Claudio** and **Kiavash**, enemies in foosball match, friends otherwise; my newer Chinese gang in the K corridor, **Fan**, **Xin**, **Huabing,** the K corridor is so much more energetic and alive with you guys here; **Ingrid**, you help me so much in your lab for the last year; **Johan** and **Markus**, it was good time we had in Bochum; and together with many more, you made my stay here at DBB so pleasant and fun, thank you.

Torbjörn and **Mattew**, thank you for all the fixing so that my lab work can be done smoothly, **Haidi, Alex** and the whole **secretary crew,** thank you for all the administration work.

To **my friends** outside the DBB radar, thank you for making life so interesting and making Stockholm like home.

Finally, I would not be here without the unconditional support and love from **my family** back in China, big thank you to you all.

Sammanfattning på svenska

Infraröd (IR) spektroskopi är en spektroskopisk metod i det infraröda området (0,78-250 μm) av det elektromagnetiska spektrumet, som motsvarar energinivåerna hos molekylära vibrationer. IR-spektroskopi är en enkel, kostnadseffektiv och snabb teknik. Till skillnad från andra spektroskopiska metoder behövs endast små proteinmängder, och storleken på proteinet är inte en begränsning. Mycket information kan dras från ett IR-spektrum, och i denna avhandling presenteras hur man kan använda IR-spektroskopi för att studera proteinstruktur och dynamik.

Artikel I och II visar hur man använder IR spektroskopi för att studera membranproteinet sarko/endoplasmatiska nätverket Ca^{2+}-ATPase (SERCA). SERCA är en kalciumpump som sitter i membranet på sarkoplasmatiska nätverket (SR), och står för 70-80 % av alla proteiner i SR-membranet. SERCA spelar en viktig roll i muskelavslappning genom att pumpa kalcium från insidan av cellen tillbaka till SR. Pumpandet är en energikrävande process som drivs av cellens bränsle ATP. Maskineriet som pumpar kalcium är komplicerat. IR-studier av SERCA-mutanter är av stort intresse för att få en bättre förståelse för hur SERCA fungerar och hur övergången mellan dess olika proteintillstånd går till. För att göra det uttrycktes rekombinant SERCA och renades upp från jäst. Dess kinetik och struktur jämfördes med naturligt SERCA från kanin. Resultaten visade att rekombinant och naturligt SERCA har liknande strukturförändringar under reaktionscykeln, men en liten skillnad kunde ses i deras sekundärstuktur. Ytterligare experiment visade att strukturskillnaden härstammade från proteinaggregering.

IR-spektrumet för ett protein ger upphov till ett brett och ostrukturerat band. För att få en bättre upplösning av spektrumet, och kunna dra information om proteiners sekundärstruktur, är spektrumet ofta kurvanpassat med dess komponentband. Traditionellt kurvanpassas endast absorptionsspektrumet på detta vis. Artikel III presenterar en ny kurvanpassningsmetod där spektrumet kurvanpassas samtidigt både till absorptionsspektrumet samt till sekundära derivatspektrumet. Ett syntetiskt spektrum och spektrum från tre olika proteiner användas för att validera metoden. Resultaten visade att metoden leder till en bättre analys av proteinets sekundärstruktur.

Artikel IV visar hur IR-spektroskopi kan användas för att studera en peptid. Peptiden designades och förväntades självt kunna bilda en fördefinierad struktur. Peptiden syntetiserades, och huruvida den kan bilda den definierade struktur kontrollerades med olika biofysikaliska metoder. Både elektronmikroskopi och röntgenkristallografi visade att peptiden självt kan bilda en struktur som är typisk för amyloid. Detaljerad information av sekundärstruktur drogs från CD och IR spektroskopi, och resultat visade att peptiden har förväntad struktur.

References

1 van Meer, G., Voelker, D. R. & Feigenson, G. W. Membrane lipids: where they are and how they behave. *Nature Reviews. Molecular Cell Biology* **9**, 112-124 (2008).

2 Luckey, M. *Membrane structural biology : with biochemical and biophysical foundations*. (Cambridge University Press, Cambridge 2008).

3 Lehninger, A. L., Nelson, D. L. & Cox, M. M. *Lehninger principles of biochemistry*. 5th edn, (W.H. Freeman, New York 2008).

4 von Heijne, G. Principles of membrane protein assembly and structure. *Progress in Biophysics and Molecular Biology* **66**, 113-139 (1996).

5 von Heijne, G. The membrane protein universe: what's out there and why bother? *Journal of Internal Medicine* **261**, 543-557 (2007).

6 Vinothkumar, K. R. & Henderson, R. Structures of membrane proteins. *Quarterly Reviews of Biophysics* **43**, 65-158 (2010).

7 Seshadri, K., Garemyr, R., Wallin, E., von Heijne, G. & Elofsson, A. Architecture of beta-barrel membrane proteins: analysis of trimeric porins. *Protein Science* **7**, 2026-2032 (1998).

8 Grigorieff, N., Ceska, T. A., Downing, K. H., Baldwin, J. M. & Henderson, R. Electron-crystallographic refinement of the structure of bacteriorhodopsin. *Journal of Molecular Biology* **259**, 393-421 (1996).

9 Pautsch, A. & Schulz, G. E. Structure of the outer membrane protein A transmembrane domain. *Nature Structural Biology* **5**, 1013-1017 (1998).

10 Jidenko, M., Nielsen, R. C., Sørensen, T. L., Møller, J. V., Nissen, P. & Jaxel, C. Crystallization of a mammalian membrane protein overexpressed in *Saccharomyces cerevisiae*. *Proceedings of the National Academy of Sciences of the United States of America* **102**, 11687-11691 (2005).

11 Lenoir, G., Menguy T., Corre, F., Montigny, C., Pedersen, P. A., Thinès, D., le Maire, M. & Falson, P. Overproduction in yeast and rapid and efficient purification of the rabbit SERCA1a Ca^{2+}-ATPase. *Biochimica et Biophysica Acta* **1560**, 67-83 (2002).

12 de Meis, L. & Vianna, A. L. Energy interconversion by the Ca^{2+}-dependent ATPase of the sarcoplasmic reticulum. *Annual Review of Biochemistry* **48**, 275-292 (1979).

13 Skou, J. C. The influence of some cations on an adenosine triphosphatase from peripheral nerves. *Biochimica et Biophysica Acta* **23**, 394-401 (1957).

14 Hasselbach, W. & Makinose, M. The calcium pump of the "relaxing granules" of muscle and its dependence on ATP-splitting. *Biochemische Zeitschrift* **333**, 518-528 (1961).

15 Slayman, C. L., Lu, C. Y. & Shane, L. Correlated changes in membrane potential and ATP concentrations in Neurospora. *Nature* **226**, 274-276 (1970).

16 Axelsen, K. B. & Palmgren, M. G. Evolution of substrate specificities in the P-type ATPase superfamily. *Journal of Molecular Evolution* **46**, 84-101 (1998).

17 Lutsenko, S. & Kaplan, J. H. Organization of P-type ATPases: significance of structural diversity. *Biochemistry* **34**, 15607-15613 (1995).
18 Møller, J. V., Juul, B. & le Maire, M. Structural organization, ion transport, and energy transduction of P-type ATPases. *Biochimica et Biophysica Acta* **1286**, 1-51 (1996).
19 Kühlbrandt, W. Biology, structure and mechanism of P-type ATPases. *Nature reviews. Molecular Cell Biology* **5**, 282-295 (2004).
20 Nissen, P. An Introduction to P-type ATPase Research. *Methods in Molecular Biology* **1377**, 1-2 (2016).
21 Hasselbach, W. The sarcoplasmic calcium pump. A model of energy transduction in biological membranes. *Topics in Current Chemistry* **78**, 1-56 (1979).
22 Periasamy, M. & Kalyanasundaram, A. SERCA pump isoforms: their role in calcium transport and disease. *Muscle and Nerve* **35**, 430-442 (2007).
23 Rossi, A. E. & Dirksen, R. T. Sarcoplasmic reticulum: the dynamic calcium governor of muscle. *Muscle and Nerve* **33**, 715-731 (2006).
24 Wray, S. & Burdyga, T. Sarcoplasmic reticulum function in smooth muscle. *Physiological Reviews* **90**, 113-178 (2010).
25 Brini, M. & Carafoli, E. Calcium pumps in health and disease. *Physiological Reviews* **89**, 1341-1378 (2009).
26 Ebashi, S., Endo, M. & Otsuki, I. Control of muscle contraction. *Quarterly Reviews of Biophysics* **2**, 351-384 (1969).
27 Szent-Györgyi, A. G. Calcium regulation of muscle contraction. *Biophysical Journal* **15**, 707-723 (1975).
28 Brody, I. A. Muscle contracture induced by exercise. A syndrome attributable to decreased relaxing factor. *The New England Journal of Medicine* **281**, 187-192 (1969).
29 Odermatt, A., Taschner, P. E., Khanna, V. K., Busch, H. F., Karpati, G., Jablecki, C. K., Breuning, M. H. & MacLennan, D. H. Mutations in the gene-encoding SERCA1, the fast-twitch skeletal muscle sarcoplasmic reticulum Ca^{2+} ATPase, are associated with Brody disease. *Nature Genetics* **14**, 191-194 (1996).
30 Møller, J. V., Olesen, C., Winther, A. M. & Nissen, P. The sarcoplasmic Ca^{2+}-ATPase: design of a perfect chemi-osmotic pump. *Quarterly Reviews of Biophysics* **43**, 501-566 (2010).
31 Toyoshima, C., Nakasako, M., Nomura, H. & Ogawa, H. Crystal structure of the calcium pump of sarcoplasmic reticulum at 2.6 A resolution. *Nature* **405**, 647-655 (2000).
32 Sørensen, T. L., Møller, J. V. & Nissen, P. Phosphoryl transfer and calcium ion occlusion in the calcium pump. *Science* **304**, 1672-1675 (2004).
33 Toyoshima, C. & Mizutani, T. Crystal structure of the calcium pump with a bound ATP analogue. *Nature* **430**, 529-535 (2004).

34 Olesen, C., Picard, M., Winther, A. M., Gyrup, C., Morth, J. P., Oxvig, C., Møller, J. V. & Nissen, P. The structural basis of calcium transport by the calcium pump. *Nature* **450**, 1036-1042 (2007).

35 Toyoshima, C. & Nomura, H. Structural changes in the calcium pump accompanying the dissociation of calcium. *Nature* **418**, 605-611 (2002).

36 Toyoshima, C., Norimatsu, Y., Iwasawa, S., Tsuda, T. & Ogawa, H. How processing of aspartylphosphate is coupled to lumenal gating of the ion pathway in the calcium pump. *Proceedings of the National Academy of Sciences of the United States of America* **104**, 19831-19836 (2007).

37 Olesen, C., Sørensen, T. L., Nielsen, R. C., Møller, J. V. & Nissen, P. Dephosphorylation of the calcium pump coupled to counterion occlusion. *Science* **306**, 2251-2255 (2004).

38 Jensen, A. M., Sørensen, T. L., Olesen, C., Møller, J. V. & Nissen, P. Modulatory and catalytic modes of ATP binding by the calcium pump. *The EMBO Journal* **25**, 2305-2314 (2006).

39 Bublitz, M., Musgaard, M., Poulsen, H., Thøgersen, L., Olesen, C., Schiøtt, B., Morth, J. P., Møller, J. V. & Nissen, P. Ion pathways in the sarcoplasmic reticulum Ca^{2+}-ATPase. *The Journal of Biological Chemistry* **288**, 10759-10765 (2013).

40 Yamasaki, K., Daiho, T. & Suzuki, H. Remarkable stability of solubilized and delipidated sarcoplasmic reticulum Ca^{2+}-ATPase with tightly bound fluoride and magnesium against detergent-induced denaturation. *The Journal of Biological Chemistry* **277**, 13615-13619 (2002).

41 Obara, K., Miyashita, N., Xu, C., Toyoshima, I., Sugita, Y., Inesi, G. & Toyoshima, C. Structural role of countertransport revealed in Ca^{2+} pump crystal structure in the absence of Ca^{2+}. *Proceedings of the National Academy of Sciences of the United States of America* **102**, 14489-14496 (2005).

42 Toyoshima, C., Nomura, H. & Tsuda, T. Lumenal gating mechanism revealed in calcium pump crystal structures with phosphate analogues. *Nature* **432**, 361-368 (2004).

43 Takahashi, M., Kondou, Y. & Toyoshima, C. Interdomain communication in calcium pump as revealed in the crystal structures with transmembrane inhibitors. *Proceedings of the National Academy of Sciences of the United States of America* **104**, 5800-5805 (2007).

44 Sipe, J. D. & Cohen, A. S. Review: history of the amyloid fibril. *Journal of Structural Biology* **130**, 88-98 (2000).

45 Kyle, R. A. Amyloidosis: a convoluted story. *British Journal of Haematology* **114**, 529-538 (2001).

46 Rambaran, R. N. & Serpell, L. C. Amyloid fibrils: abnormal protein assembly. *Prion* **2**, 112-117 (2008).

47 DePace, A. H., Santoso, A., Hillner, P. & Weissman, J. S. A critical role for amino-terminal glutamine/asparagine repeats in the formation and propagation of a yeast prion. *Cell* **93**, 1241-1252 (1998).

48 Jahn, T. R., Makin, O. S., Morris, K. L., Marshall, K. E., Tian, P., Sikorski, P. & Serpell, L. C. The common architecture of cross-beta amyloid. *Journal of Molecular Biology* **395**, 717-727 (2010).

49 Chiti, F. & Dobson, C. M. Protein misfolding, functional amyloid, and human disease. *Annual Review of Biochemistry* **75**, 333-366 (2006).

50 Pepys, M. B. Pathogenesis, diagnosis and treatment of systemic amyloidosis. *Philosophical Transactions of the Royal Society of London. Series B, Biological sciences* **356**, 203-210 (2001).

51 Sipe, J. D. Amyloidosis. *Critical reviews in Clinical Laboratory Sciences* **31**, 325-354 (1994).

52 Gertz, M. A. & Rajkumar, S. V. *Amyloidosis: Diagnosis and Treatment*. (Humana Press, New York 2010).

53 Howie, A. J., Brewer, D. B., Howell, D. & Jones, A. P. Physical basis of colors seen in Congo red-stained amyloid in polarized light. *Laboratory Investigation* **88**, 232-242 (2008).

54 Howie, A. J. & Brewer, D. B. Optical properties of amyloid stained by Congo red: history and mechanisms. *Micron* **40**, 285-301 (2009).

55 Eanes, E. D. & Glenner, G. G. X-ray diffraction studies on amyloid filaments. *Journal of Histochemsitry and Cytochemistry* **16**, 673-677 (1968).

56 Pauling, L. & Corey, R. B. Configuration of polypeptide chains. *Nature* **168**, 550-551 (1951).

57 Sikorski, P., Atkins, E. D. & Serpell, L. C. Structure and texture of fibrous crystals formed by Alzheimer's abeta(11-25) peptide fragment. *Structure* **11**, 915-926 (2003).

58 Geddes, A. J., Parker, K. D., Atkins, E. D. & Beighton, E. "Cross-beta" conformation in proteins. *Journal of Molecular Biology* **32**, 343-358 (1968).

59 Sunde, M., Serpell, L. C., Bartlam, M., Fraser, P. E., Pepys, M. B. & Blake, C. C. Common core structure of amyloid fibrils by synchrotron X-ray diffraction. *Journal of Molecular Biology* **273**, 729-739 (1997).

60 Goldsbury, C., Frey, P., Olivieri, V., Aebi, U. & Muller, S. A. Multiple assembly pathways underlie amyloid-beta fibril polymorphisms. *Journal of Molecular Biology* **352**, 282-298 (2005).

61 Smith, J. F., Knowles, T. P., Dobson, C. M., Macphee, C. E. & Welland, M. E. Characterization of the nanoscale properties of individual amyloid fibrils. *Proceedings of the National Academy of Sciences of the United States of America* **103**, 15806-15811 (2006).

62 Cherny, I. & Gazit, E. Amyloids: not only pathological agents but also ordered nanomaterials. *Angewandte Chemie International Edition in English* **47**, 4062-4069 (2008).

63 Mankar, S., Anoop, A., Sen, S. & Maji, S. K. Nanomaterials: amyloids reflect their brighter side. *Nano Reviews* **2** (2011).

64 Reches, M. & Gazit, E. Casting metal nanowires within discrete self-assembled peptide nanotubes. *Science* **300**, 625-627 (2003).

65 Scheibel, T., Parthasarathy, R., Sawicki, G., Lin, X. M., Jaeger, H. & Lindquist, S. L. Conducting nanowires built by controlled self-assembly of amyloid fibers and selective metal deposition. *Proceedings of the National Academy of Sciences of the United States of America* **100**, 4527-4532 (2003).

66 Baldwin, A. J., Bader, R., Christodoulou, J., MacPhee, C. E., Dobson, C. M. & Barker, P. D. Cytochrome display on amyloid fibrils. *Journal of the American Chemsitry Society* **128**, 2162-2163 (2006).

67 Blum, M. M. & John, H. Historical perspective and modern applications of Attenuated Total Reflectance-Fourier Transform Infrared Spectroscopy (ATR-FTIR). *Drug Testing and Analysis* **4**, 298-302 (2012).

68 Derrick, M. R., Stulik, D. & Landry, J. M. *Infrared spectroscopy in conservation science*. (Getty Conservation Institute, Los Angeles 1999).

69 Griffiths, P. R. & De Haseth, J. A. *Fourier transform infrared spectrometry*. 2nd edn, (Wiley-Interscience, New York 2007).

70 Rogalski, A. Infrared detectors: an overview. *Infrared Physics and Technology* **43**, 187-210 (2002).

71 Kong, J. & Yu, S. Fourier transform infrared spectroscopic analysis of protein secondary structures. *Acta Biochimica et Biophysica Sinica* **39**, 549-559 (2007).

72 Barth, A. Infrared spectroscopy of proteins. *Biochimica et Biophysica Acta* **1767**, 1073-1101 (2007).

73 Krimm, S. & Bandekar, J. Vibrational spectroscopy and conformation of peptides, polypeptides, and proteins. *Advances in Protein Chemistry* **38**, 181-364 (1986).

74 Bandekar, J. Amide modes and protein conformation. *Biochimica et Biophysica Acta* **1120**, 123-143 (1992).

75 Goormaghtigh, E., Cabiaux, V. & Ruysschaert, J. M. Determination of soluble and membrane protein structure by Fourier transform infrared spectroscopy. I. Assignments and model compounds. *Sub-cellular Biochemistry* **23**, 329-362 (1994).

76 Timasheff, S. N., Susi, H. & Stevens, L. Infrared spectra and protein conformations in aqueous solutions. II. Survey of globular proteins. *The Journal of Biological Chemistry* **242**, 5467-5473 (1967).

77 Seshadri, S., Khurana, R. & Fink, A. L. Fourier transform infrared spectroscopy in analysis of protein deposits. *Methods in Enzymology* **309**, 559-576 (1999).

78 Goormaghtigh, E., Cabiaux, V. & Ruysschaert, J. M. Determination of soluble and membrane protein structure by Fourier transform infrared spectroscopy. III. Secondary structures. *Sub-cellular Biochemistry* **23**, 405-450 (1994).

79 Barth, A. & Zscherp, C. What vibrations tell us about proteins. *Quarterly Reviews of Biophysics* **35**, 369-430 (2002).

80 Haris, P. I. & Severcan, F. FTIR spectroscopic characterization of protein structure in aqueous and non-aqueous media. *Journal of Molecular Catalysis B: Enzymatic* **7**, 207-221 (1999).

81 Jackson, M. & Mantsch, H. H. The use and misuse of FTIR spectroscopy in the determination of protein structure. *Critical Reviews in Biochemistry and Molecular Biology* **30**, 95-120 (1995).

82 Dong, A., Huang, P. & Caughey, W. S. Protein secondary structures in water from second-derivative amide I infrared spectra. *Biochemistry* **29**, 3303-3308 (1990).

83 Mäntele, W. Reaction-induced infrared difference spectroscopy for the study of protein function and reaction mechanisms. *Trends in Biochemical Sciences* **18**, 197-202 (1993).

84 Gerwert, K. Molecular reaction mechanisms of proteins monitored by time-resolved FTIR-spectroscopy. *Biological Chemistry* **380**, 931-935 (1999).

85 Siebert, F. Infrared spectroscopy applied to biochemical and biological problems. *Methods in Enzymology* **246**, 501-526 (1995).

86 Jung, C. Insight into protein structure and protein-ligand recognition by Fourier transform infrared spectroscopy. *Journal of Molecular Recognition* **13**, 325-351 (2000).

87 Vogel, R. & Siebert, F. Vibrational spectroscopy as a tool for probing protein function. *Current Opinion in Chemical Biology* **4**, 518-523 (2000).

88 Wharton, C. W. Infrared spectroscopy of enzyme reaction intermediates. *Natural Product Reports* **17**, 447-453 (2000).

89 Zscherp, C. & Barth, A. Reaction-induced infrared difference spectroscopy for the study of protein reaction mechanisms. *Biochemistry* **40**, 1875-1883 (2001).

90 McCray, J. A. & Trentham, D. R. Properties and uses of photoreactive caged compounds. *Annual Review of Biophysics and Biophysical Chemistry* **18**, 239-270 (1989).

91 Goeldner, M. & Givens, R. *Dynamic studies in biology : phototriggers, photoswitches and caged biomolecules*. (Wiley-VCH, Weinheim 2005).

92 Morad, M., Davies, N. W., Kaplan, J. H. & Lux, H. D. Inactivation and block of calcium channels by photo-released Ca^{2+} in dorsal root ganglion neurons. *Science* **241**, 842-844 (1988).

93 Hartung, K., Grell, E., Hasselbach, W. & Bamberg, E. Electrical pump currents generated by the Ca^{2+}-ATPase of sarcoplasmic reticulum vesicles adsorbed on black lipid membranes. *Biochimica et Biophysica Acta* **900**, 209-220 (1987).

94 Janko, K. & Reichert, J. Proton concentration jumps and generation of transmembrane pH-gradients by photolysis of 4-formyl-6-methoxy-3-nitrophenoxyacetic acid. *Biochimica et Biophysica Acta* **905**, 409-416 (1987).

95 Dantzig, J. A., Walker, J. W., Trentham, D. R. & Goldman, Y. E. Relaxation of muscle fibers with adenosine 5'-[gamma-thio]triphosphate (ATP[gamma S]) and by laser photolysis of caged ATP[gamma S]: evidence for Ca^{2+}-dependent affinity of rapidly detaching zero-force cross-bridges. *Proceedings of the National Academy of Sciences of the United States of America* **85**, 6716-6720 (1988).

96 Kaplan, J. H. & Ellis-Davies, G. C. Photolabile chelators for the rapid photorelease of divalent cations. *Proceedings of the National Academy of Sciences of the United States of America* **85**, 6571-6575 (1988).

97 Goormaghtigh, E., Cabiaux, V. & Ruysschaert, J. M. Secondary structure and dosage of soluble and membrane proteins by attenuated total reflection Fourier-transform infrared spectroscopy on hydrated films. *European Journal of Biochemistry / FEBS* **193**, 409-420 (1990).

98 Byler, D. M. & Susi, H. Examination of the secondary structure of proteins by deconvolved FTIR spectra. *Biopolymers* **25**, 469-487 (1986).

99 Holler, F., Burns, D. H. & Callis, J. B. Direct use of second derivatives in curve-fitting procedures. *Applied Spectroscopy* **43**, 877-882 (1989).

100 Lee, D. C., Hayward, J. A., Restall, C. J. & Chapman, D. Second-derivative infrared spectroscopic studies of the secondary structures of bacteriorhodopsin and Ca^{2+}-ATPase. *Biochemistry* **24**, 4364-4373 (1985).

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

102 Garidel, P. & Schott, H. Fourier-transform midinfrared spectroscopy for analysis and screening of liquid protein formulations. Part 2: Detailed analysis and applications. *BioProcess International* **4**, 48-55 (2006).