Insulin unfolding and aggregation: a multi-disciplinary study

Author: Sergio Mauri

Supervisor: Dr. Heike Arnolds

Surface Science Research Centre

Department of Chemistry

Submitted for the degree of Doctor of Philosophy

August 2014

Abstract

This thesis aims at understanding the interaction between insulin and interfaces with a multi-disciplinary approach.

We investigate three facets of the interaction.

In the first part (chapter 4), we study the interaction of insulin with the air/water interface, for different oligomeric compositions of the solution phase. With the help of Sum Frequency Spectroscopy and calculations of the second order nonlinear susceptibility, we can show that insulin monomers segregate to the hydrophobic air/water interface. Since the insulin monomer is the key species to denature and refold to fibrils, our finding explains for the first time why agitation of insulin solutions and the accompanying increase in air/water interface area accelerates fibril formation.

In the second part (chapter 5), we investigate the interaction of insulin monomers at low pH with model hydrophilic and hydrophobic solid surfaces. We use a combination of spectroscopic methods, like ATR FT-IR, XPS, SFG and QCM-D to characterise the silicon functionalised solid surface, to quantify the amount of adsorbed protein and to determine its secondary structure. We show that, contrary to physiological conditions, where insulin monomers are known to change secondary structure upon adsorption, an acidic environment leads to near-native adsorbed insulin, which is stable for at least a day. We further show that heat is needed to restructure the

adsorbed insulin monolayer and that this restructured monolayer appears to provide the template for further growth.

In the final part (chapter 5), we apply a comparatively simple experimental method, Reflection Anisotropy Spectroscopy for the first time to the formation of amyloid fibrils at interfaces. In a comparison with FT-IR spectroscopy of our model solid surfaces, we show that a drastic change in the peptide backbone arrangement occurs at a hydrophobic sureface, when FT-IR merely detects a thick layer with partial beta-sheet structure. We believe this structural change is the beginning of insulin fibril formation and we use the new tool to explore further changes in the adsorbed layer as it ages over several months.

Acknowledgements

Finally after almost four years of work, and after some crazy weeks writing this thesis, it is now time to spend few words for all the people whom have contributed directly or indirectly to this work. I will do it chronologically, since I don't want to forget to thank anybody. The first person I would like to acknowledge is Heike Arnolds, since without her all this would have not been possible. I would like to thank you for the incredible support you gave me during the course of these 4 amazing years. It has been a pleasure to work for and with you. Thank you for the patience you had especially at the beginning, and for the knowledge transfer and for the motivation you gave me when things seemed not to go very well. I will miss our baby, the laser... well maybe she's more a grumpy old woman... I will definitely remember it like my personal photonic ultrafast auntie. Well, and of course thanks for the amazing chocolate cake you baked for my birthday, and for showing me how to prepare the most incredible *feuerzangenbowle*. Then it comes to the person who transformed me from an engineer to a protein freak: thanks to Martin Volk, who showed me the ropes of the magic world of proteins. I am really glad you did it though! It was very pleasant having such a passionate mentor. At this point I should mention the "Italian mafia", and the gang: thanks to Emiliano, Ilaria, Luigi, Tommaso, Giulia, Diego, Enzo and I'm pretty sure I forgot some. When I moved to Liverpool I could not imagine

that I was going to meet such amazing people. You all made this years happy... Happy as it was always sunny... and we know in Liverpool is not! You really made life here great, each of you in our own way. I am sure that I will not find anybody else like Tommaso, I could speak Chinese with (well, our personal version) and spend half of our salary in "Nero bits". I would also like to acknowledge the special guest Vittorio, whom I met quite randomly. Thanks for the long chats we used to have, and for the crazy nights out too! I will leave Liverpool hopefully with a PhD degree, but most of all with an infinite collection on funny, crazy memories and stories. I would like to write some of them, but I would end up with this section probably as long as the rest of the thesis. I acknowledge Tobias Weidner and his group for welcoming me at the Max Planck Institute during the last months of my Ph.D., and for giving me the possibility to work in such a thrilling environment (shall I mention here the numbers of cakes too?) I want to thank Lucia, Filippo, Samet, Natalia and Vasillis, whom I recently met in Mainz. Let's cut it short: I like the Mediterranean flavour you bring everyday to our Mainzer life. I would then acknowledge Caroline Smith, Peter Weightman and Ben Johnson for contributing to this work. A special thanks goes also to Natalia for the interesting scientific and nonscientific discussion we used to have over lunch breaks and so. Il piú grande e sincero grazie lo rivolgo ai miei genitori, Adelia e Giovanni, che in tutti questi anni hanno saputo indirizzarmi lungo la strada delle personali aspirazioni, senza farmi perdere di vista il valore del sacrificio, del lavoro, della dedizione. Vi ringrazio per il supporto morale ed economico che mi ha permesso di completare questa grandiosa esperienza. Vorrei ringraziare anche la nonna Marisa, che é venuta a mancare recentemente. Grazie per aver creduto ed essere stata orgogliosa di me, senza onestamente sapere cosa facessi. Ho sempre pensato che un

cosí grande atto di fiducia fosse impossibile, ma tu hai saputo dimostrarmi in errore (e non succede spesso!). Thanks to all the other members of the family, especially to my sister Chiara, and to my Brother Alessio and my sister in law Irene.. And of course to your amazing kids Elena and Gabriele! I know I need to mention Elisa, as the adopted sister. I thank you for being a reliable confidant, and for being as mad as a hatter too.. we all need some sort of excitement! I want to thank my auntie Maria, my uncle Rosario, my cousin Davide, because they dare to come over and visit me sometimes. Your company was always a pleasure. I would like to thank my cousin Enrica, for being always present and supportive over the years. It is always nice to wake up in the morning and receive an e-mail from you, just to say: how are you? I thank Andrea and Filippo too... two chaps I've been good friend of... since I can remember! I finally want to thank all the people I have not explicitly mentioned here, mainly because I'm writing this at 4am in the night, and honestly I feel so emotionally overwhelmed, that I really want to conclude.

Contents

List of Figures

Chapter 1

Introduction

Understanding the process of protein denaturation and aggregation is still challenging researchers in various fields of science [1] [2]. Proteins, in general, express their functions thanks to their structure (*native state*). Denaturated or unfolded states are usually undesired, since they make the protein ineffective. Such states might further cause protein aggregation, which is, again, usually not desired. The stated topic sees involvement from many fields of science: biologists, chemists and physicists have been trying to understand the mechanism of protein denaturation for years. Of course, great progress has been made, but generally speaking, these mechanisms are so complex that we are far from fully understanding them. The complexity of protein denaturation and aggregation mechanisms lies in the complexity of proteins themselves. Proteins are in general amphiphilic molecules, they come in different sizes and shapes, and they carry out a numerous set of functions in organisms. As a consequence, each protein is "one of a kind ". Amongst the large number of proteins in nature, the ones that play a role in the human body are particularly interesting. In fact, some proteins in our body undergo denaturation, and sometimes they aggregate in fibril-like structures, called

amyloids. Amyloids formation is associated with diseases known indeed as amyloidosis. Alzheimer's disease and Creutzfeld-Jakob disorder, just to mention two, are caused by amyloid formation in neuronal cells [3] [4]. One of the most popular proteins for studying fibril formation is insulin in its human or animal variants. This popularity is due to two main reasons: on one side, insulin's small size, and mainly α-helical structure make it the perfect model peptide for studying amyloid formation. On the other side, insulin is the most common protein based medication. In pharmacological applications it is extremely important to deliver the medication while keeping the protein in its native structure. Therefore, protein denaturation should be avoided.

1.1 Problem statement

The scientific community has worked intensively to study the phenomenon of amyloid formation. Even if the reasons driving the protein fibrillation process are manifold and still partially obscure, one idea emerges from the large number of studies done: surfaces seem to constitute one trigger element behind the denaturation process [2]. Studying the interaction between proteins and surfaces requires two steps: the binding of the protein to the surface and the folding/unfolding mechanism. Due to its amphiphilic character, a protein may interact with both hydrophobic and hydrophilic surfaces. In the case of insulin, it is commonly believed that hydrophobic surfaces/interfaces enhance the denaturation and aggregation process. The actual reason why this occurs is still under discussion. One complicating factor is that insulin is found in nature in different oligomeric forms: the most common ones are monomers, dimers and hexamers. Only the first species can undergo denaturation, whereas dimers and hexamers are more stable [5]. Brange *et al.* [6]

have discussed how different oligomeric compositions of insulin can affect the time lag before fibrils form. They have shown, by measuring the viscosity of insulin solutions, that if monomers are present in solution fibrils are formed in 1.5 days (while agitating a 0.6 M solution at 37°C) whereas in the presence of dimers, the formation of fibrils occurs only after 8 days. Nielsen *et al.* [7] have studied the aggregation of various insulin mutants, using a fluorescent dye (Thioflavin T), and proved that only monomers participate in the fibril formation process. Up to date, despite the large number of studies published, none of them has focused systematically on studying the interaction of insulin with well-defined surfaces. Claesson *et al.* [8] have investigated insulin adsorption on hydrophobic mica by measuring the forces between the surface and the adsorbed layer, and between different layers of insulin. They show the irreversible adsorption of an insulin layer onto the hydrophobic mica surface. Nilsson *et al.* [9] have demonstrated the adsorption of insulin hexamers on hydrophilic surfaces and of the monomeric form on hydrophobic silica surfaces by means of ellipsometry and they show that the amount of adsorbed insulin on hydrophobic surfaces is largely independent of protein concentration, ionic strength and addition of ions. From their work they conclude that *there is strong evidence that it is the monomeric form of insulin that adsorbs to the the hydrophobic surface and that the driving force is hydrophobic interaction.* More recently Nault et al. [10] [11] have proposed a model to explain the mechanism of insulin fibrillation mediated by surfaces. From a combination of techniques (Infrared Spectroscopy, Surface Plasmon Resonance and Fluorescence microscopy) they propose that the insulin layer adsorbed on a surface is the intermediate along the α to β structural transition that results in the formation of amyloids on these surfaces. The drawback of this study is that different hydrophobic surfaces were used for each technique, therefore

experimental results are difficult to compare.

1.2 Outline of the thesis

This introduction has provided some general information about the importance of studying insulin aggregation and denaturation at surfaces.

Chapter 2 is an introduction specifically to human insulin. The concept of protein primary and secondary structures is presented. Particular attention is bestowed on the differences between insulin oligomers: monomers and dimers especially. The last section of this chapter is dedicated to the importance of insulin in studying fibril formation in general.

Chapter 3 is an experimental chapter, where I explain the various methods used in this work. In particular, I focus on how these techniques have been used to characterise both the surfaces used and the adsorbed protein. Special attention is given to Sum Frequency Spectroscopy (SFS), and how it is used to characterise surfaces and measure protein structure at interfaces.

The experimental results are presented in 3 parts.

In chapter 4, we use SFS to determine that insulin only exists as monomers at the air/water interface regardless of bulk composition. This explains why the air/water interface enhance insulin denaturation and aggregation.

Chapter 5 investigates insulin adsorption at low pH on model hydrophobic and hydrophilic functionalised silicon wafers. Surfaces and insulin adsorption is characterised by SFS, infrared spectroscopy, XPS and QCM-D. Surprisingly insulin retains its native structure on both surfaces for extended times at room temperature, and only temperature increase finally triggers unfolding and denaturation

In chapter 6, we monitor the formation of insulin fibrils on the same

solid surfaces as in chapter 5. By using Reflection Anisotropy Spectroscopy (RAS), we measure the growth of insulin fibrils and compare it to infrared spectra.

1.3 Contributions

The work of this thesis has contributed to the following publications:

"Insulin Adsorption on Hydrophobic and Hydrophilic Silane Self-assembled Monolayers." *S. Mauri, M. Volk, S. Byard, H. Berchtold and H. Arnolds, ready for submission to Langmuir*

"The structure of insulin at the air/water interface: monomers or dimers?" *S. Mauri, T. Weidner and H. Arnolds, Phys. Chem. Chem. Phys., 2014, 16, p. 26722-26724, DOI: 10.1039/C4CP04926H*

Chapter 2

Human insulin

Proteins are large amphiphilic molecules. They are ubiquitous in nature and they are present in every organism. They fulfill many different functions: haemoglobin is responsible for oxygen transport in the blood, keratin is a part of a family of fibrous structural proteins that constitute the main component of hair and nails. Enzymes in cells are proteins that catalyse biochemical reactions. Size and shape vary a lot amongst proteins, but despite their high complexity, proteins are made with a limited set of building blocks and regular/common patterns can often be found [12].

2.1 Protein structure: from the amino acid sequence to tertiary structure

21 amino acids constitute the small set of building blocks proteins are made of. Each amino acid consists of three main parts: an amine group (-NH_2), a carboxyl group (-COOH) and a side chain which is specific to each amino acid. As shown in fig. 2.1 side chains express different chemical properties, like polarity, charge, hydrophobicity and aromaticity.

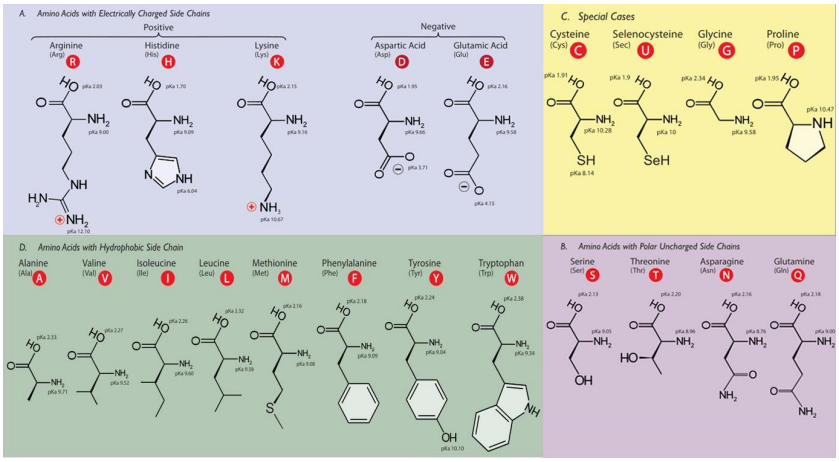

Figure 2.1: *Chart of the 21 amino acids available in nature. A high resolution image can be found in appendix E[from Wikipedia]*

Like letters in an alphabet, amino acids can be assembled together to form an incredibly large vocabulary. The amine group of one amino acid can react with the carboxyl group of another one: this reaction results in a very strong chemical bond, known as peptide bond (see fig. 2.2). Such reaction can repeat many times, giving birth to a *polymer-like* chain composed of amino acids.

The sequence of amino acids uniquely identifies a protein: such sequence is called the *primary structure*. The primary sequence determines also the spatial arrangement of the protein. Sometimes features with particular shapes are found in proteins: helical and sheet-like arrangements are quite common. Hydrogen bonding between a carboxyl oxygen of one amino acid and an amine hydrogen of a different one, along with hydrophobic and electrostatic interactions, causes these local arrangements. These shapes formed are known as secondary structure. Helices can have different periodicity. The number of residues per turn determines the nature of the helix. The so called α-helices have 3.6 residues per turn, whereas the less common 3_{10} helices have

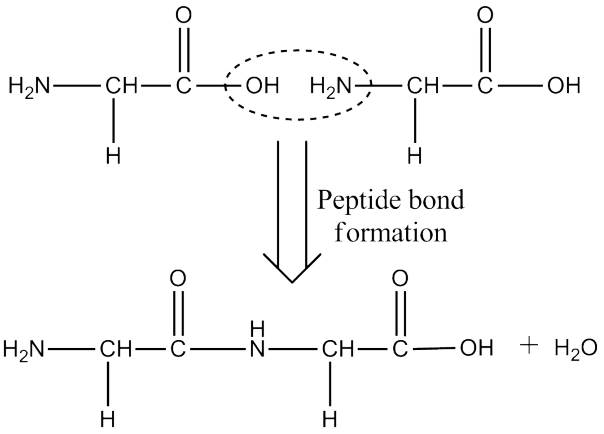

Figure 2.2: *Two single glycine molecules react. A peptide bond is formed, and a water molecule released.*

3 residues per turn. β-sheets can be parallel or anti-parallel, according to the backbone direction of two adjacent strands. A picture of these structures is presented in fig 2.3.

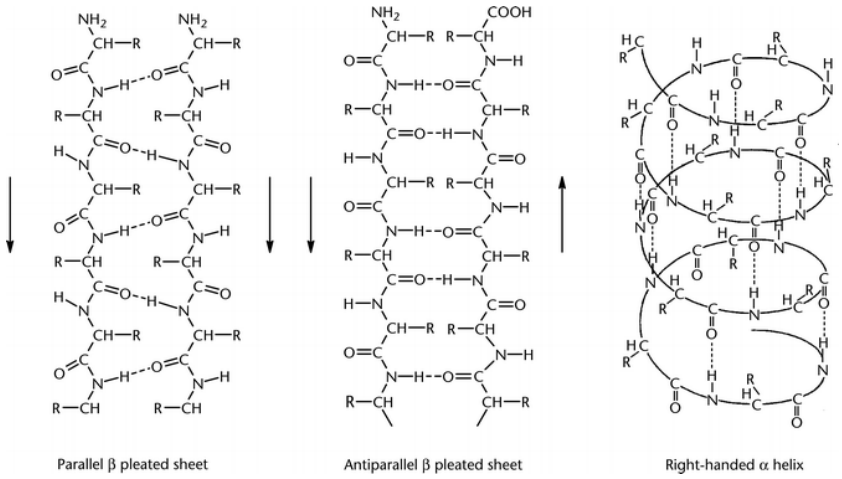

Figure 2.3: *Secondary structures: parallel β-sheet, antiparallel β-sheet, α-helix (left to right). [Encyclopedia of Life Sciences (Nature Publishing Group, London, 2001)].*

Finally, the overall arrangement of the protein in space is known as tertiary structure.

2.2 Insulin structure and self-association: Monomers, dimers and hexamers

Amongst the numerous proteins present in nature, there are some which play a key role in the human body. Insulin is one of them as it controls the sugar level in animals' blood. Insulin deficiency in the blood or cells' high resistance to this hormone are pathological. The disease is well-known as *diabetes*. Correlation between pancreatic deficiency and diabetes was first proposed by Langerhans in 1869 [13]. Insulin was first discovered by Banting and Best in 1922. They extracted a white liquid from calves, and injected it, after a rough purification, into dogs whose pancreas had been removed to make them diabetic. The treatment kept diabetic dogs alive, so they proceeded to test it on human patients. The results were immediate and the patients experienced a quick recovery from a previously incurable disease. In 1923 Banting and Best were awarded the Nobel Prize for medicine, and Hoechst in Germany and Eli Lilly in USA start to produce the drug. A few years later, the isolation of insulin was reported [14] and soon after research focused on insulin medications rapidly expanded and insulin became widespread as a cure for diabetic people, but its structure remained unknown for a long time. The primary structure of insulin was first proposed in 1951 [15], and the 3D structure was only published 50 years after its discovery [16].

Insulin is a small hormone made of 51 amino acids arranged in two separate chains known as A and B chains as shown in fig.2.4.

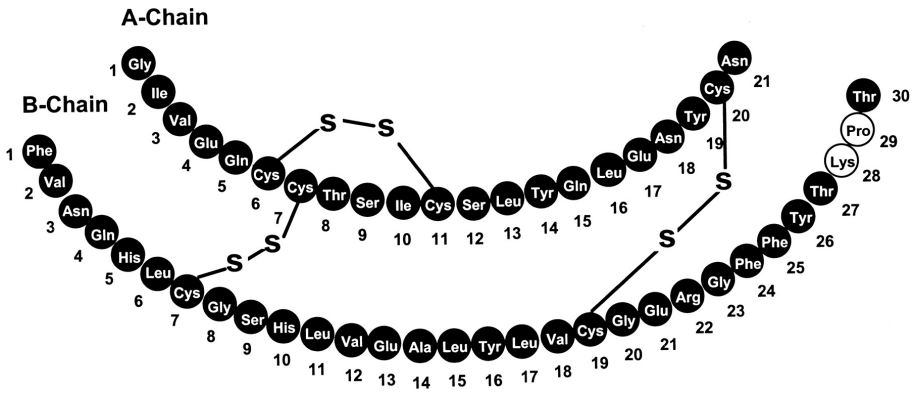

Figure 2.4: *Insulin primary sequence. Disulphide bonds are also shown [17]*

The A chain consists of 21 amino acids. Residues A1 to A8 and A12 to A18 are arranged in α-helices The B chain consists of 30 amino acids, where residues B9 to B19 form an α-helix and the dimer interface is made by residues B24 to B28. The two chains are linked together by two disulphide bonds. An extra disulphide bond links two cysteine residues in the A chain, giving insulin a compact and stable shape.

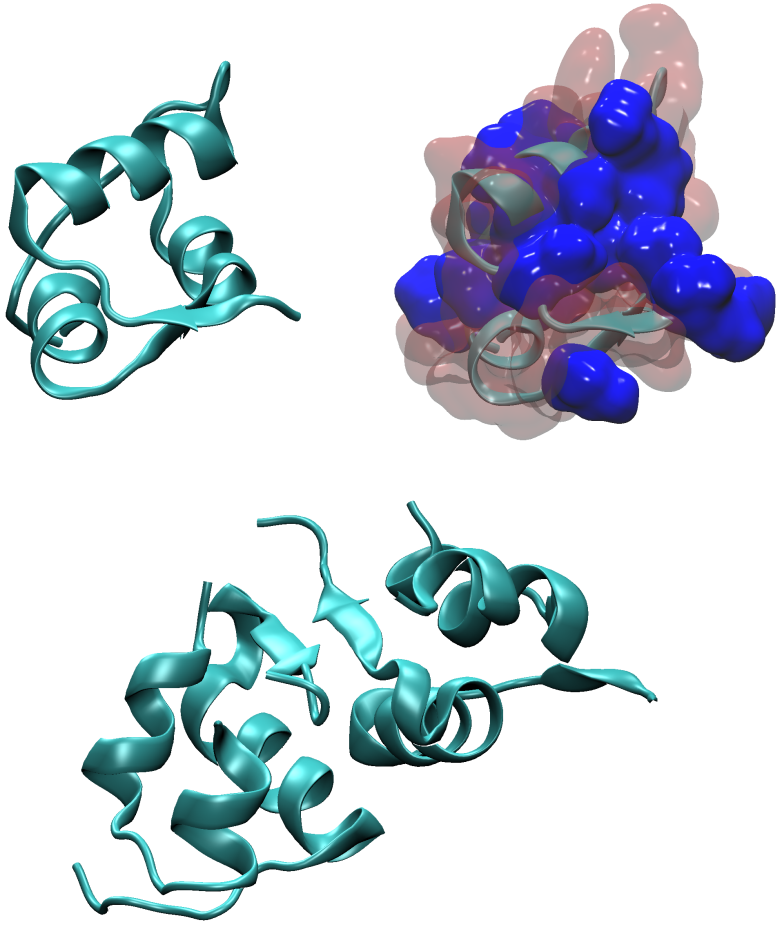

Figure 2.5: *Secondary structure of insulin oligomers. Top: insulin monomer (pdb identifier 2JV1). On the right side, the hydrophobic surface is shown in solid blue and the hydrophilic surface is shown in red (transparent). Bottom: insulin dimers (pdb identifier: 4INS).*

Insulin oligomers are nature's way for storing insulin. Insulin can be found in the form of monomers, dimers and hexamers. The relative concentration of each oligomer can be chemically controlled, by adjusting pH, concentration and by addition of metal ions in solution. Low concentrations and low pH determine insulin to be largely monomeric in solution, whereas at higher concentrations and pH dimers are predominant. At pH 2, monomers domi-

nate below 0.5 g/l [18], while at pH 7, monomers only dominate below 0.1 g/l [19]. Insulin self-association is driven by the presence of divalent metal ions, such as Zn^{2+} and produces its hexameric form. In insulin monomers, a hydrophobic surface is exposed to water molecules in solution. This creates a lower stability of monomeric insulin, in comparison to the other two oligomeric forms. Hexameric insulin, in particular, is very stable and it is in fact used for storage purposes in pharmaceutical applications [20].

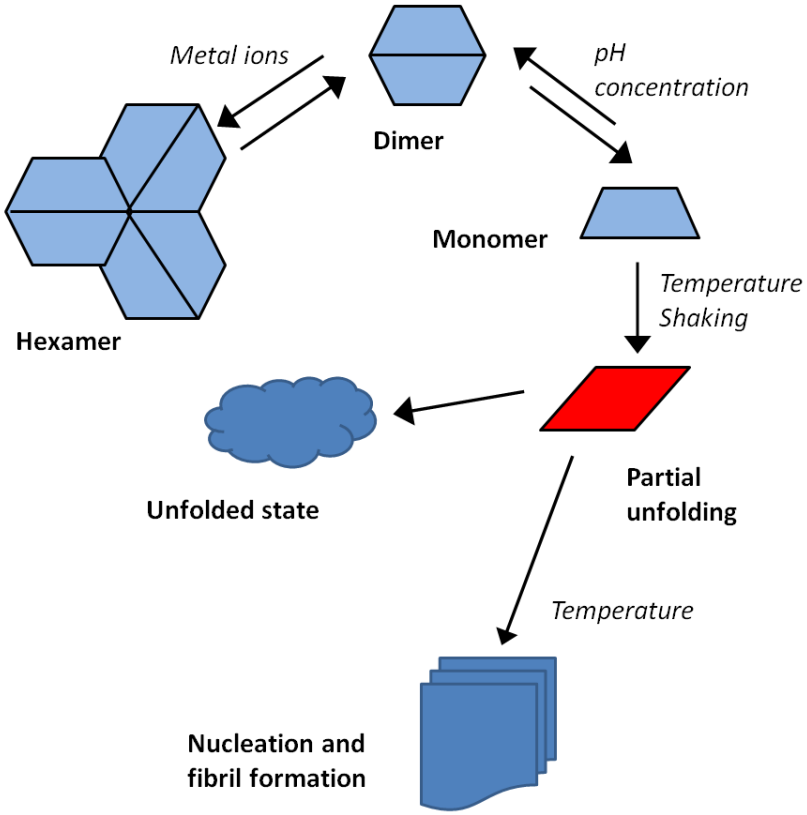

Figure 2.6: *Unfolding/Aggregation pathway. Only insulin monomers can undergo unfolding and further aggregation [7].*

Under certain conditions, insulin undergoes denaturation and further aggregates in fibril-like macro-structures. Effectiveness of insulin medications is

degraded because of fibrillation during storage or injection. Such structures are thus undesired, and very often dangerous for the human body [21] [22]. Insulin fibril formation can be induced by shaking, by adding denaturating agents, or by exposing insulin solutions to high temperatures.

2.3 The importance of human insulin: a model peptide for studying fibril formation

Diseases related to amyloid formation are numerous (like Alzheimer and Creutzfeld-Jakob syndromes) and have been intensively studied. Every amyloidosis is associated with the formation of fibrils of different proteins, nevertheless, the structures of the fibril are always very similar [23]. Ross *et al.* [24] and Petkova *et al.* [25] have extensively discussed details of the fibril structure, describing them as filamentous structures with a width of $\sim 10\mu m$ and a length of $0.1-10\mu m$ characterised by the presence of cross β-structures. Insulin fibrils' properties have been analysed in detail in the literature. Smith *et al.* [26] have studied fibrils' mechanical properties by means of atomic force spectroscopy, and they have proved these fibrils to be extremely stiff and strong. The process of fibril formation though is not clear yet. It has been shown [27] [28] that specific sequences of amino acids can favour fibril formation. These sequences are know as *amyloidogenic sequences*. Insulin has two of these sequences. One consists of 7 amino acids in the B chain: LVEALYL (B11 to B17). The other one is a 6 amino acid sequence in the A chain: NELQYL (A13 to A18). Ivanova *et al.* [27] have discussed the role of the LVEALYL sequence in accelerating and/or retarding insulin fibril formation and shown a concentration dependent behaviour (of extra added LVEALYL in solution). Insulin has become a model peptide for studying fib-

ril formation because it has such a simple structure. Jimenenz *et al.* [29] have proposed a model for insulin fibrils consisting of a cross-β structure. Several cross-β filaments twist around one another to form the fibril (fig. 2.7). This model proposes fibril formation after complete unfolding of the native insulin. A and B chains remain bonded, and stack on top of each other, to further rearrange in a β-like structure (as shown in fig. 2.7b). Those β structures further self-assemble in a fiber-like structure with a typical cross section of 30×40 Å

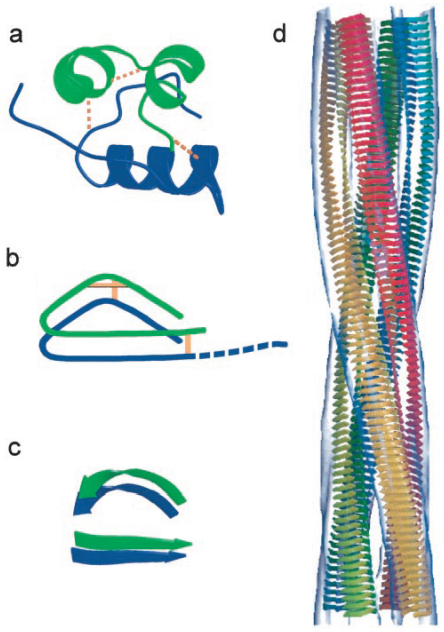

Figure 2.7: *(a) Insulin native structrue. Disulphide bonds are shown in yellow, A chain is depicted in green and B chain is depicted in blue. (b) Unfolded state of a single insulin monomer. Possible model for an amyloid protofilament. (c) β-strand model for a protofilament. Each insulin occupies two layers. (d) β-strand model. From Jimenez et al. [29].*

Jimenez *et al.* proposed moreover the existence of 4 different types of insulin fibrils of different in thicknesses. They measured cryo-EM images, and

their corresponding diffraction patterns and they propose that each insulin fibrils present a common protofilament structure as shown in fig.2.7.

Insulin protofibrils have been recently studied by Amenabar *et al.* [30] by means of nano-FTIR spectroscopy to determine the composition of a single insulin fibril. This work shows that insulin fibrils posseses a core rich in β-sheet , whereas the outer shell is composed of α-helical features. IR

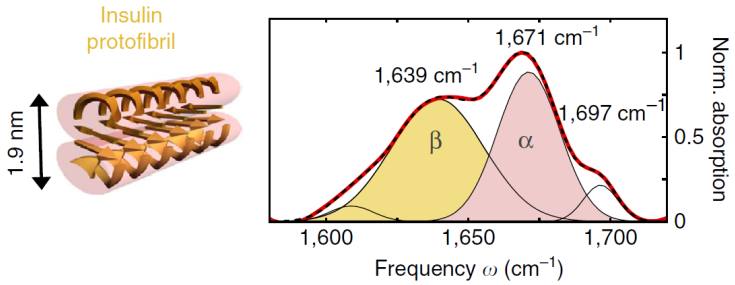

Figure 2.8: *Left: Sketch of an insulin protofibril. Right: nano-FTIR spectrum of insulin fibrils. Form Amenabar at al. [30]*

spectroscopy is well-suited to identify different types of secondary structures, as discussed later in this thesis (see section 3.1). The two peaks at 1639 cm^{-1} and 1671 cm^{-1} shown in fig. 2.8 confirm the presence of both α-helices and β-sheets in the fibril [31].

2.4 Motivation and challenges

Previous studies have shown repeatedly in both solution and at interfaces that the insulin monomer is the main culprit when it comes to fibril formation. None of these studies, though, have answered the simple question: *Why*?

So to address this, we picked the simplest of the hydrophobic interfaces: the air/water interface, and use the nonlinear optical selection rules of sum

frequency spectroscopy to distinguish between monomers and dimers.

We then focus on insulin monomer adsorption on solid surfaces at low pH, an important area for the production of insulin drugs. No detailed structural studies exist under these conditions.

Finally, motivated by the scarcity of experimental methods to determine protein secondary structure at interfaces, we explore the use of reflection anisotropy spectroscopy to monitor insulin fibril formation.

Chapter 3

Studying proteins at surfaces: principles and techniques

Proteins interact with surfaces in a large number of cases. Proteins are constantly interacting with cell membranes for signaling reasons. Proteins interact with implants, and their interaction determines the biocompatibility of such implants and so on. The importance of proteins interacting with surfaces has been briefly explained in the introduction. Investigating protein adsorption on a surface requires the study of multiple aspects. It is necessary to know how much and how quickly proteins adsorb on the surface [32]. At the same time, it is very useful to see if this adsorption is reversible or irreversible, or competitive with other proteins [33]. Last but not least, one also wants to measure the secondary structure of the protein adsorbed. All three aspects present some challenges. A common one is the difficulty to achieve surface sensitivity. In fact, while the scientific panorama is sprinkled with bulk sensitive techniques, the number and variety of surface sensitive techniques suitable for studying protein interaction with surfaces is limited, although in constant evolution. None of the techniques currently available

provides all the necessary features for addressing the manifold questions of the protein/surface interaction problem. Therefore, in case one wants to obtain an exhaustive set of information about the amount and structure of protein adsorbed on a certain surface, a multi-technique approach has to be adopted. In this chapter, I introduce the basics of the experimental methods which I have used throughout my PhD work, explaining the reasons why such experimental technique has been adopted in a specific case. Practical examples will be shown, too. In fig 3.1, all these techniques are grouped according to their quantitative/qualitative properties. As one can see, these

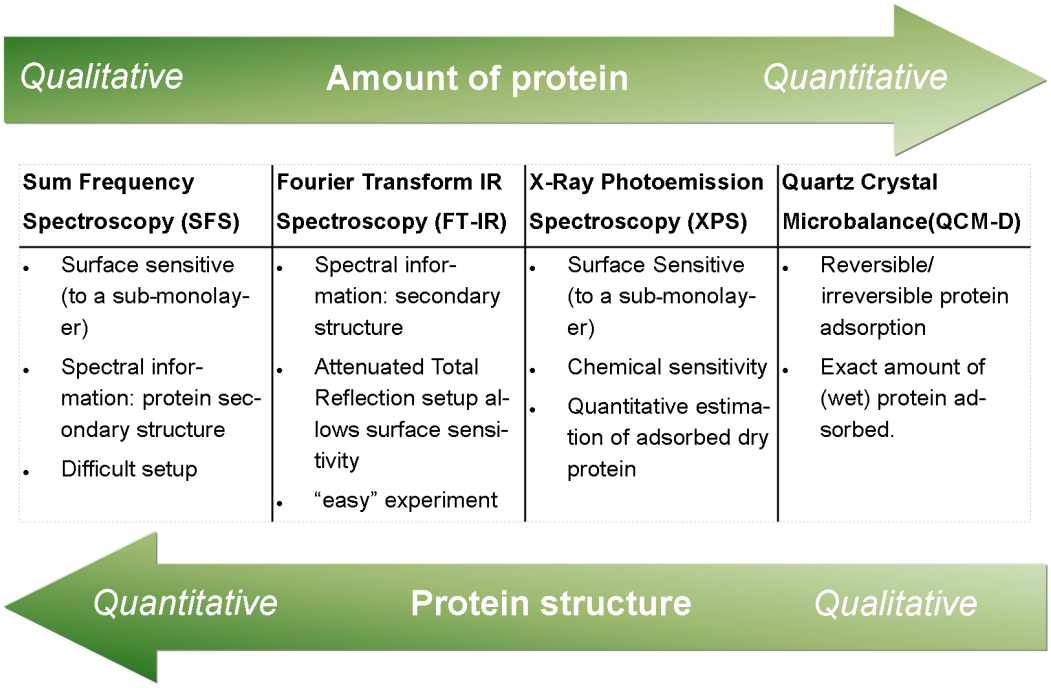

Figure 3.1: *Overview of the experimental techniques used during this PhD work. Pros and cons of each are highlighted.*

four methods are a suitable combination to address the problem of insulin interaction and aggregation with a variety of surfaces.

3.1 Infrared Spectroscopy

InfraRed (IR) spectroscopy is a very well known method which has been used for many years to study the secondary structure of proteins [34] [35] [36]. In general, IR spectroscopy relies on molecular vibrations. A molecular bond vibrates, twists, rotates and so on. These motions are associated with specific frequencies, which occur mostly in the mid-IR spectral region (3500 cm^{-1} to 1000 cm^{-1}). IR light travelling through a sample is partially absorbed by molecular vibrations. An absorption spectrum is then recorded. The measured absorbance is given according to the Lambert-Beer law as follows:

$$A(\bar{\nu}) = \frac{I(L,\bar{\nu})}{I_0(\bar{\nu})} = cL\epsilon(\bar{\nu}) \tag{3.1}$$

where I_0 is the intensity of the IR light, $\bar{\nu}$ is the frequency expressed in wavenumbers, L is the path length through which the light travels, c is the concentration of protein in solution and ϵ is the molar absorption or extinction coefficient. Usually unpolarised light is used, so molecular orientation does not play a role. The wide availability of IR spectrometers, the simplicity of performing the experiment, and the ease of sample preparation have made this a popular and robust method for studying protein secondary structure.

3.1.1 Protein IR absorption: the amide I modes

The interesting frequency region for determining protein secondary structure is the so called *amide I* region [37], as it has contributions of various molecular vibrations within the protein backbone, as shown in fig. 3.2. The amide I mode is mostly due do the C=O stretching vibration of the amide group coupled to the in-phase bending of the N-H bond and the stretching of the C-N bond.

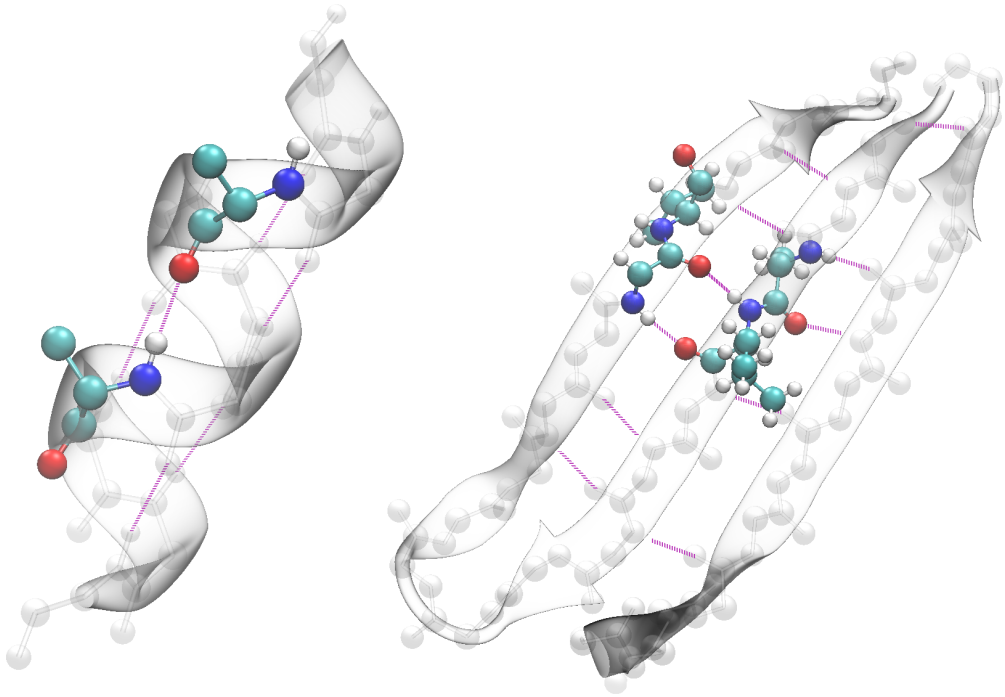

Figure 3.2: *Left: α-helix. Right: β-sheet. The hydrogen bonding found in different secondary structures is depicted by magenta dotted lines. Carbon atoms are shown in light blue, nitrogen in blue, oxigen in red.*

The amide I vibration gives rise to an IR band in the region between 1600 cm^{-1} and 1700 cm^{-1}. Since water absorbs strongly in the same region, most of the experiments must be performed using a different solvent, usually deuterated water. The amide I mode is highly dependent on the spatial conformation of the protein backbone due to its dependence on hydrogen bonding within the backbone itself. The coupling between transition dipole moments in a certain secondary structure pattern determines the amide I frequency of the various secondary structures. In peptides, where only one type of secondary arrangement is present, it is easy to assign an IR peak to their specific secondary structure. Back in 1976, Nevskaya and Chirgadze

have published a series of papers [38] [39] [40] where calculated IR absorption spectra corresponding to α-helices and parallel and anti-parallel β-sheets were reported. In fig. 3.3, one can see that each secondary structure pattern corresponds to a specific vibrational frequency.

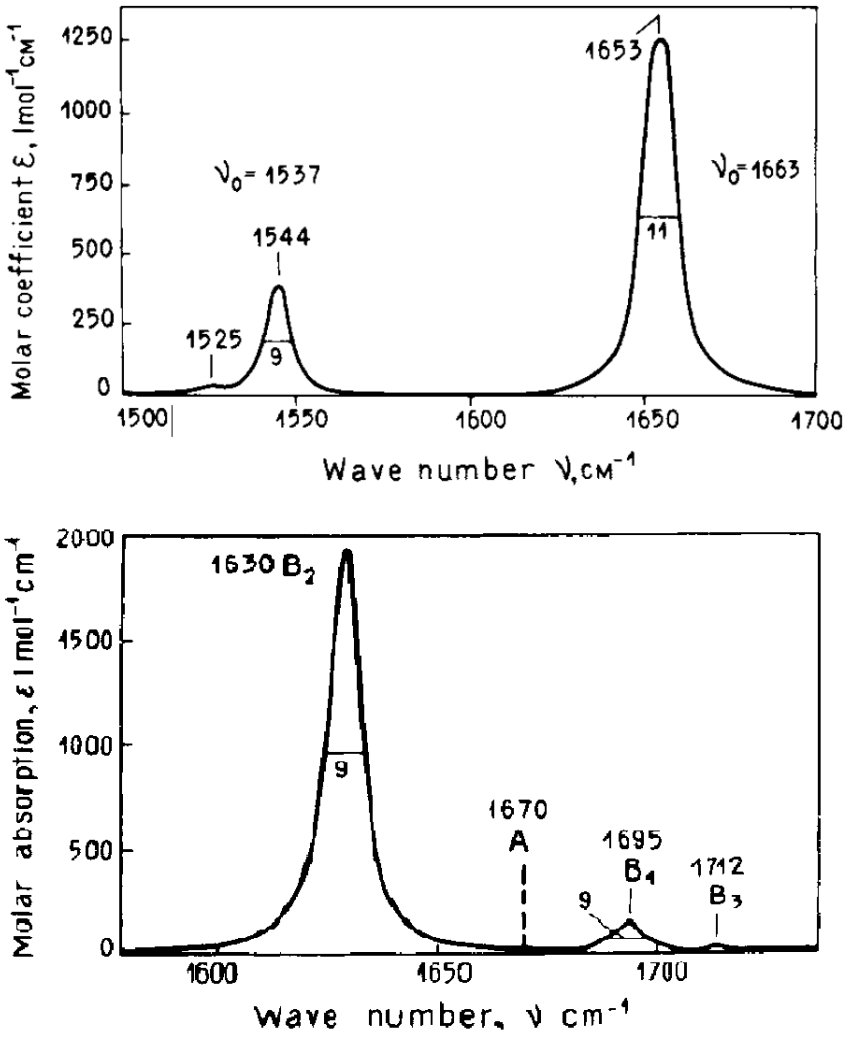

Figure 3.3: *Top: Simulated IR spectrum of an infinitely long α-helix. Bottom: Simulated IR spectrum of an infinitely long β-sheet. Reproduced from et al. [38] [39] [40].*

Theoretically, infinite anti-parallel and parallel β-sheets should absorb at the same frequency, but experiments have shown that the first absorb at a slightly lower frequency. The peak position of the amide I peak, depends also on the solvent surrounding the protein. The amide I frequency shifts due to H-D exchange, thus amide hydrogen atoms in the peptide bonds exchange protons with the solvent. Frequencies corresponding to each secondary structure have been measured [41]and are shown in table 3.1. Proteins

	Band position in H_2O		Band position in D_2O	
Secondary structure	Average	Extreme	Average	Extreme
α-helix	1654	1648-1657	1652	1642-1660
β-sheet	1633	1623-1641	1630	1615-1638
β-sheet	1684	1674-1695	1679	1672-1694
Turn	1672	1662-1686	1671	1653-1691
Random coil	1654	1642-1657	1645	1639-1654

Table 3.1: *Assignment of amide I band positions to secondary structure based on experimental data and assignments of various authors, collected and evaluated by Goormaghtigh et al. [41]*

have usually an uneven composition, though, and therefore band assignment is not always straightforward. The amide I signal of a protein is often the result of multiple overlapped bands. The complexity of the amide I band has been discussed over the years. Susi *et al.* [42] have measured FT-IR spectra of various proteins, and have shown the possibility of interpreting otherwise featureless spectra using self-deconvolution to facilitate peak assignment. Arrondo *et al.* [43] have presented in detail advantages and disadvantages of peak fitting of non-deconvoluted spectra. The possibility of using second derivative spectra to ease peak positioning has been discussed,

too [35], along with smoothing and other strategies to steer FT-IR analysis of proteins to a more quantitative direction. Collected wisdom (or advice) from the cited and many other publications is that great care should be taken while fitting FT-IR spectra, to avoid over-interpretation of experimental results. Human insulin however, has a simple enough structure that FT-IR analysis is comparatively easy.

3.1.2 IR spectroscopy of human insulin

Human insulin in its native structure consists mostly of α-helices (as discussed in chapter 2.2), and therefore the IR spectrum in the amide I region can be relatively easily interpreted. Here we show IR spectra of insulin in its native structure and after refolding and aggregation. In this case, a solution of insulin in D_2O at pD=2.7 is prepared. pD is adjusted using diluted DCl in D_2O. After allowing insulin some minutes to completely dissolve in solution, part of the solution is readily measured, whereas other parts are heated at a temperature of 65°C. H_2O contamination is avoided by storing the vials in bigger vials filled with more D_2O. In fig. 3.4 we report a series of IR spectra taken after heating the solutions at a concentration of 1 mg ml^{-1} for different times. The native structure of insulin is characterised by a peak at 1651 cm^{-1} as shown by previous studies [44] [45]. No insulin denaturation and aggregation occur in the course of a week, if the solution is constantly kept at 65°C. At this low concentration insulin aggregates only if the temperature is raised and shaking applied.

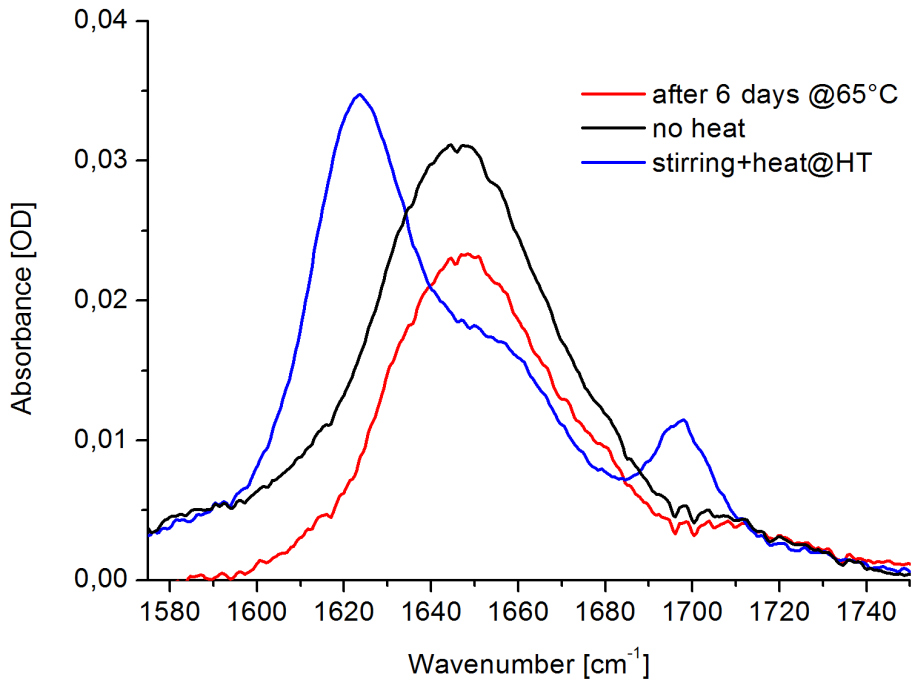

Figure 3.4: *Infrared spectra of human insulin in D_2O before and after heating. Insulin solution is prepared 1mg/ml and pD 2.7*

The presence of β-like structures is the indicated by a sharp peak at 1622 cm^{-1}. Insulin aggregation occurs on a faster timescale when solutions at higher concentrations are heated. Bouchard *et al.* [44] have shown that β-sheets are formed after only 7 hours in 2 mM solutions ($\approx$ 12 $mgml^{-1}$).

3.1.3 Achieving surface sensitivity: Attenuated Total Reflection (ATR) configuration

As mentioned earlier, IR spectroscopy is not intrinsically surface specific. Surface sensitivity can be achieved though, by using a particular optical configuration. The IR light is not transmitted through the sample, but sent into a crystal of a specific shape, usually a trapezoidal silicon or germanium

crystal. If the angle of incidence is greater than the critical angle ($\sin\theta_c = \frac{n(surroundings)}{n(crystal)}$), then the light will be internally reflected as shown in figure 3.5. If we consider a silicon crystal and a layer of protein ($\lambda = 6\mu m$, $n_{Si} = 3.42$ and $n_{protein} \approx 1.55$ [46]) then the critical angle θ_c is $\approx 27°$

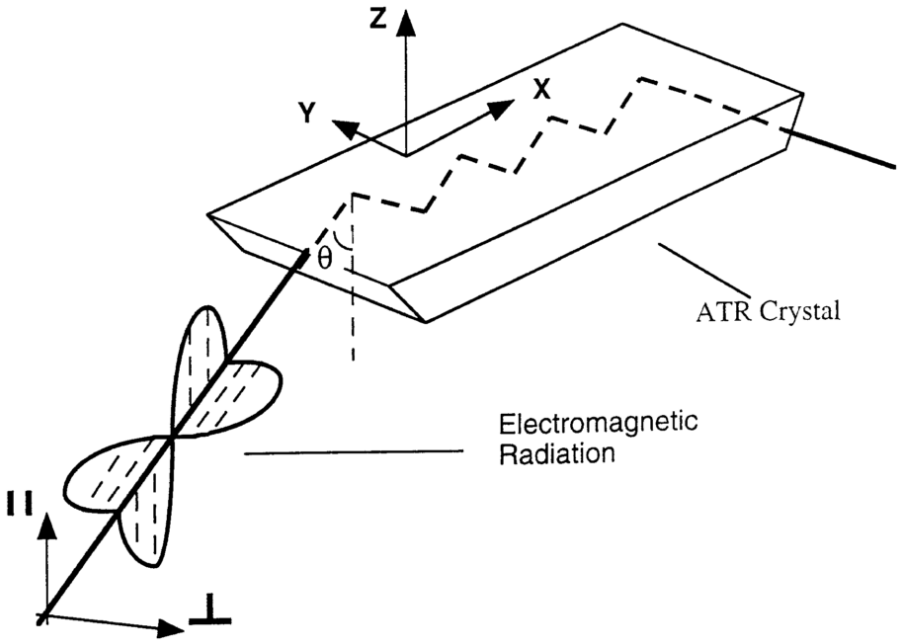

Figure 3.5: *Schematic of an FT-IR ATR experiment [47].*

IR light is assumed to be perpendicularly incident at the input and output faces. The number of reflections depends on the length and thickness of the crystal: $N = \frac{crystal\,length}{crystal\,thickness} \cdot \cot(\theta)$, where θ is the angle of incidence (inside the crystal). In this configuration, called indeed *Attenuated Total Reflection – ATR*, an evanescent electric field is generated at each reflection. The evanescent field is transmitted into the low refractive index medium, and decays exponentially:

$$E(z) = E_0 e^{-\frac{z}{d_p}} \tag{3.2}$$

where z is the direction orthogonal to the crystal surface and d_p is the depth

of penetration into the low refractive index medium. d_p is independent of polarisation [47] and equal to:

$$d_p = \frac{\lambda}{2\pi n_1 \cdot \sqrt{sin^2\theta - \left(\frac{n_2}{n_1}\right)^2}} \tag{3.3}$$

where λ is the wavelength of the IR light, θ is the angle of incidence inside the crystal, n_1 is the refractive index of the ATR crystal and n_2 is the refractive index of the layer (of protein) adsorbed onto the ATR crystal. In fig. 3.6, we show the calculated penetration depth d_p for a silicon crystal in contact with air ($n_2 = 1$) or in contact with a layer with a refractive index of 1.55 (typical value for a protein layer). d_p is in the micrometer range across the IR range.

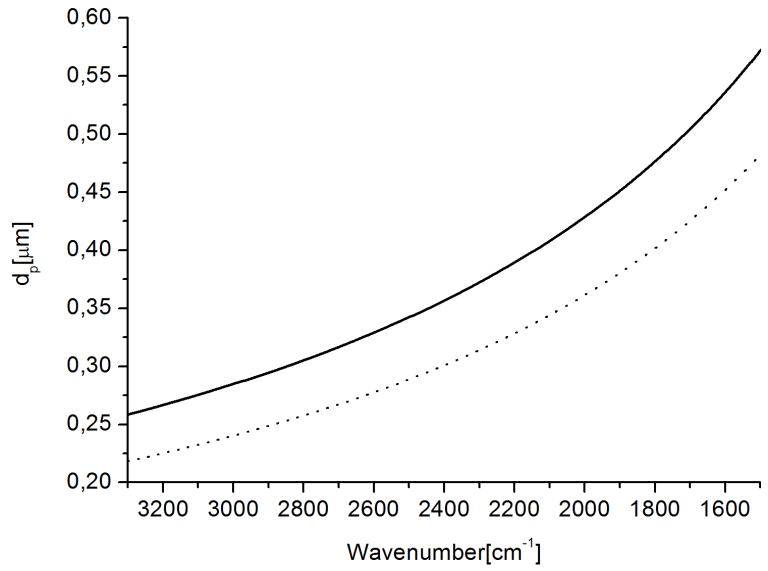

Figure 3.6: *Calculated penetration depth d_p for a Si crystal with 45° internal reflection angle. Dotted line: penetration depth in air. Solid line: penetration depth when the adsorbed layer has a refractive index of 1.55*

To easily compare FT-IR ATR measurements with FT-IR experiments done in transmission, an *effective thickness* d_e can be defined according to

Harrick [48] as follows:

$$d_e = \frac{(\frac{n_2}{n_1}){E_0}^2 d_p}{2cos(\theta)} \tag{3.4}$$

IR absorbance measurements can then be used to quantify the amount of protein adsorbed onto a surface. This calculation requires the molar absorptivity ε of a protein, which can be obtained from a transmission experiment, assuming the validity of the Beer-Lambert law (see eq. 3.1). The absorbance for each reflection can be written as [49] [50] [51]:

$$\frac{A}{N} = \frac{(\frac{n_2}{n_1}){E_0}^2 \varepsilon}{cos(\theta)} \int_0^\infty C(z) exp\left(\frac{-2z}{d_p}\right) dz \tag{3.5}$$

where A is the integrated absorbance, N is the number of reflections, ε is the integrated molar absorptivity (cm^{-2} mol), z is the distance from the surface, C(z) is the concentration as function of distance from the surface and d_e is the effective thickness (as presented in eq. 3.4). If we assume the adsorbed protein layer to have a thickness t and concentration C_i, followed by a bulk solution of concentration C_b for $z > t$, we can split the integral 3.5 into two parts:

$$\frac{A}{N} = \frac{(\frac{n_2}{n_1}){E_0}^2 \varepsilon}{cos(\theta)} \int_0^t C_i exp\left(\frac{-2z}{d_p}\right) dz + \frac{(\frac{n_2}{n_1}){E_0}^2 \varepsilon}{cos(\theta)} \int_t^\infty C_b exp\left(\frac{-2z}{d_p}\right) dz \tag{3.6}$$

t is usually much smaller than d_p, thus $exp\left(\frac{-2t}{d_p}\right) \approx 1 - \frac{2t}{d_p}$. Therefore, the first integral simplifies to $\varepsilon\left(\frac{2d_e}{d_p}\right)(C_i t)$. The second integral is equal to $\varepsilon C_b d_e$ where d_e is the effective thickness as defined in eq. 3.4. The simplified equation for the absorbance per reflection is given by:

$$\frac{A}{N} = \varepsilon\left(\frac{-2d_e}{d_p}\right)(C_i t) + \varepsilon C_b d_e \tag{3.7}$$

This equation shows that the thickness of an adsorbed layer of protein can be determined quantitatively from known or measurable quantities. In fact, C_b

is the concentration of proteins of the bulk solution (and can be determined with accuracy, by using for example UV-Vis spectrscopy) and C_i can be estimated. Similarly, ε can be estimated for a certain protein, or can be obtained from IR spectra of bulk solutions.

3.2 Sum Frequency Spectroscopy – SFS

Sum Frequency Vibrational Spectroscopy (SFS) is based on the second order nonlinear optical process of Sum Frequency Generation. SFS is capable of providing vibrational spectra of sub-monolayers of molecules ordered at an interface. SFS has been used intensively in the last few decades for studying a large variety of systems, including self-assembled monolayers on metal surfaces, proteins at various interfaces and so on [52] [53]. Sum Frequency Generation is here discussed in the dipole approximation. For an extended and thorough description of the SFG process, one can refer to textbooks by Shen and Boyd [54] [55]. Let us consider the interaction of a single molecule with an electromagnetic field. At the molecular level, the electric field $\mathbf{E}$ will induce a dipole moment μ as follows:

$$\mu = \alpha \mathbf{E} \tag{3.8}$$

where α is the molecular polarizability. On the macroscopic scale, eq. 3.8 translates into eq. 3.9:

$$\mathbf{P} = \varepsilon_0 \chi \mathbf{E} \tag{3.9}$$

where ε_0 is the permittivity of vacuum, $\mathbf{P}$ is the polarization vector and χ is the molecular susceptibility. If the electric field interacting with molecules is strong, we need to include higher orders of the electric field in eq. 3.8 and eq. 3.9. The vector polarisation $\mathbf{P}$ becomes nonlinearly dependent on the

electric field:

$$\mathbf{P} = \varepsilon_0 \left(\chi^{(1)} \cdot \mathbf{E} + \chi^{(2)} : \mathbf{EE} + ... \right) \tag{3.10}$$

where $\chi^{(n)}$ are the nonlinear susceptibilities. $\chi^{(n)}$ is related to the *n-th* order hyperpolarizability as explained in appendix A. In SFG, two separate electric fields at frequencies ω_1 and ω_1 are incident onto a sample. As a result, a polarization $\mathbf{P}(\omega_3)$ at a frequency $\omega_3 = \omega_1 + \omega_2$ is generated. This second order nonlinear optical process occurs only in those systems which are not centrosymmetric. Let us consider a polarisation $\mathbf{P}$ oscillating at frequency 2ω in the direction z.

$$\mathbf{P}(2\omega, z) = \chi^{(2)} : \mathbf{E}(\omega, z)\mathbf{E}(\omega, z) \tag{3.11}$$

In a centrosymmetric medium $\chi^{(2)}$ is uniform across the whole medium. If we now reverse the axis z into -z, eq. 3.11 can be rewritten as

$$\begin{aligned} \mathbf{P}(2\omega, -z) &= \chi^{(2)} : \mathbf{E}(\omega, -z)\mathbf{E}(\omega, -z) = \\ &= \chi^{(2)} : \mathbf{E}(\omega, z)\mathbf{E}(\omega, z) = -\mathbf{P}(2\omega, z) \\ \rightarrow \mathbf{P}(2\omega, -z) &= \mathbf{P}(2\omega, z) \end{aligned} \tag{3.12}$$

This can only be fulfilled if $\chi^{(2)} = 0$, meaning that $\mathbf{P}(2\omega)$ is not vanishing only whenever there is a symmetry break, like at a surface, making SFG intrinsically surface sensitive. The simplest way for writing the polarisation vector $\mathbf{P}(\omega_{SF} = \omega_2 + \omega_1)$ is to consider an electric field $\mathbf{E}$ oscillating at frequencies ω_2 and ω_2:

$$\mathbf{E} = \mathbf{E}_1 + \mathbf{E}_2 = \mathbf{E}_1(\omega_1)\cos(\omega_1 t) + \mathbf{E}_2(\omega_2)\cos(\omega_2 t) \tag{3.13}$$

The second order polarisation will then be proportional to the square of the electric field **E**:

$$\begin{aligned}\mathbf{P}^{(2)} &= \varepsilon_0 \chi^{(2)} : \mathbf{E}^2 = \varepsilon_0 \chi^{(2)} \left(\mathbf{E}_1 \cos(\omega_1 t) + \mathbf{E}_2 cos(\omega_2 t)\right)^2 \\ &= \left(\mathbf{E}_1^2 + \mathbf{E}_2^2\right) + \left(\frac{1}{2}\mathbf{E}_1^2 \cos(2\omega_1 t) + \frac{1}{2}\mathbf{E}_2^2 \cos(2\omega_2 t)\right) + \\ &+ \left(\mathbf{E}_1 \mathbf{E}_2 \cos\left[(\omega_1 - \omega_2)t\right]\right) + \left(\mathbf{E}_1 \mathbf{E}_2 \cos\left[(\omega_1 + \omega_2)t\right]\right)\end{aligned} \tag{3.14}$$

The two fields give rise to a DC component. The second term of eq. 3.14 describes the generation of the second harmonic (SHG) of the electric fields $\mathbf{E}_1$ and $\mathbf{E}_2$. The third one defines the generation of a field at the difference frequency (DFG). The last term of eq. 3.14 is the one describing the generation of a photon at the sum frequency $\omega_{SF} = \omega_1 + \omega_2$.

Sum Frequency Spectroscopy is used typically for probing molecular vibrations at an interface, thus the frequency ω_{IR} of one of the two fields occurs in the IR spectral range. Since the detection efficiency of most detectors is high in the visible spectral range, the other photon is tuned to a frequency that produces a SF signal in the visible region. Therefore the other frequency ω_{VIS} is chosen in the visible/near IR spectral region. When ω_{IR} matches one of the vibrational frequencies of the investigated molecular layer, the SF signal is greatly enhanced through resonance. When ω_{IR} is out of resonance, the intensity of the SF signal is small, as illustrated different line thicknesses in fig. 3.7

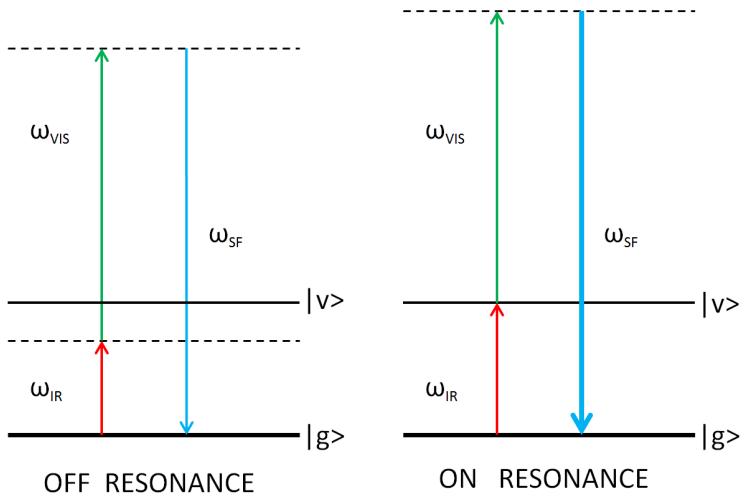

Figure 3.7: *SFG scheme. Left: generation of SF out of resonance. In this case the SF signal will be weak. Right: SF is generated in resonance with a vibrational mode of a molecule at a surface. Because of this resonance, the SF intensity is greatly enhanced.*

3.2.1 Broadband Visible-IR SFG: ultrafast pulses, tunable IR light generation and suppression of non-resonant background

The Sum Frequency Generation process usually has a low yield, being a non-linear effect. SF photons are generated when laser pulses with a sufficiently high peak intensity are mixed at a surface. Mainly two different kinds of setup have been developed over the years. The first type consists of a laser system that generates picosecond pulses of both visible and IR light, which have a narrow bandwidth (few cm^{-1}). In order to record a full SF spectrum, the IR frequency has to be scanned through the spectral range of interest. Another way of performing SFS requires the use of femtosecond laser pulses. We refer to these systems as *broadband SF spectrometers*. Femtosec-

ond pulses of sufficient intensity can be generated using a combination of a laser oscillator (commonly a Ti:Sapphire laser), followed by a regenerative amplifier. A Ti:Sapphire oscillator usually produces pulses with an energy in the nJ range, insufficient for performing SFS. Therefore these pulses are amplified in a regenerative amplifier, with the working principle explained in fig. 3.8, called *Chirped Pulse Amplification* [56]. A femtosecond ($\sim$ nJ) pulse is chirped using a grating. The stretched pulse has the same energy of the original one, but a longer duration. The stretched pulse is then amplified in a laser crystal. At this point, the pulse has a higher energy than the incoming one, but is not short enough in time. After the amplification, the pulse is therefore compressed again using a combination of mirrors and grating. The final output will be a femtosecond pulse with energy of few mJ.

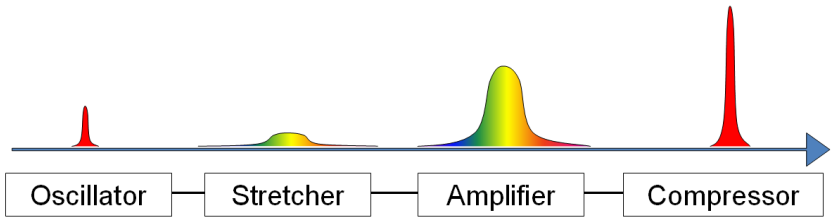

Figure 3.8: *Chirped Pulse Amplification. A nJ femtosecond pulse is succesively sent trough a stretcher, amplifier and compressor to obtain a mJ femtosecond pulse*

The femtosecond mJ train of pulses is used to produce both visible and IR light. IR light is generated in an Optical Parametric Amplifier (OPA). OPAs are commercially available, and allow to generate tunable broadband IR pulses. Since the IR pulses are broadband, a single pulse can excite IR vibrations within $\sim$100 cm^{-1}. Therefore with a single shot, it is possible to record a full spectrum. The spectral response of an SFG spectrograph is determined by the lineshape of both infrared and visible laser pulses, and in particular correspond to the convolution of IR and visible lineshape. A

purely nonresonant SF spectrum can then be recorded (e.g. off a gold surface) and used for normalising spectra. In order to achieve spectral resolution, a narrow-band visible pulse is necessary. This can be achieved in two ways. One can use either a pulse shaper or an etalon. The use of a pulse shaper allows to vary the spectral resolution from experiment to experiment, but the optical setup can sometimes be complicated. The use of an etalon fixes the spectral resolution to a certain number ($\sim$10 cm^{-1}), but offers other advantages. First of all, the optical setup is simpler, and requires only minimal alignment. Secondly, an etalon produces pulses with an exponential temporal decay, which allows suppression of the nonresonant background. This method was introduced by Latguchev *et al.* [57]. The infrared polarisation generated on the surface has two sources, a nonresonant one and a resonant one. The first decays on the timescale of the IR pulse, the second one decays with the dephasing time of the molecular polarisation (see fig. 3.9). Therefore, if the visible pulse is delayed for a sufficient time, it will only upconvert the resonant contribution, simplifying spectral analysis. Fig. 3.9 illustrates this with spectra (measured, not normalised) of ODT (Octadecanethiol)/Au for zero and 1 ps delay. In fig.3.9 (b), the vibrational modes of the methyl endgroup of the ODT self-assembled monolayers, results as dips instead of peaks, because of the relative phase between the nonresonant signal (arising from the gold substrate) and the resonant signal (arising from the monolayer).

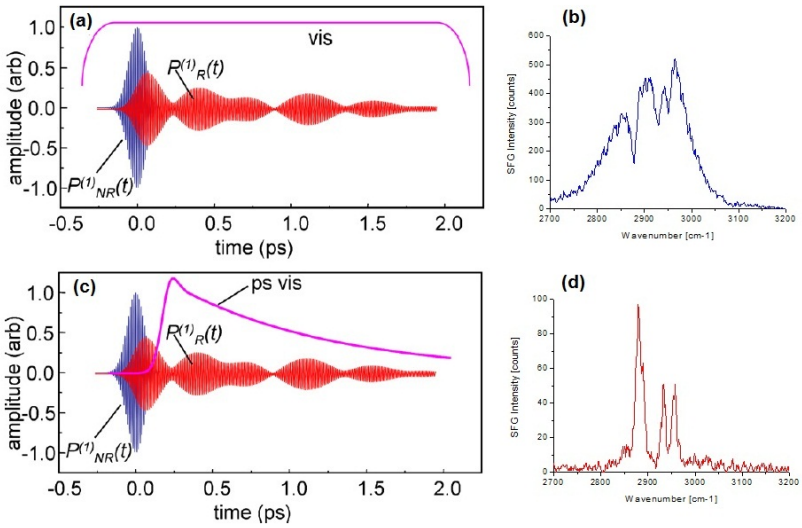

Figure 3.9: *Fast-decaying nonresonant (blue) and slow-decaying resonant (red) polarization are sketched. (a) Time-symmetric visible pulse (pink line): nonresonant and resonant signal can't be discriminated. (b) Spectrum of ODT SAM on Au(111). Etalon at zero delay time: nonresonant background is not suppressed. (c) Time-asymmetric visible pulse (pink line): a purely resonant SFG signal can be obtained. (d) Nonresonant background free spectrum of ODT SAM on Au(111). Etalon at 1 ps delay time [57]*

3.2.2 The interaction of light with a surface: *s* and *p* polarised light, Fresnel coefficients and the SF equation

When an electromagnetic field is reflected off a surface, it is possible to resolve its components into components which are parallel and perpendicular to the plane of incidence of the electromagnetic field. These components are in general identified as p and s components, as shown in fig. 3.10.

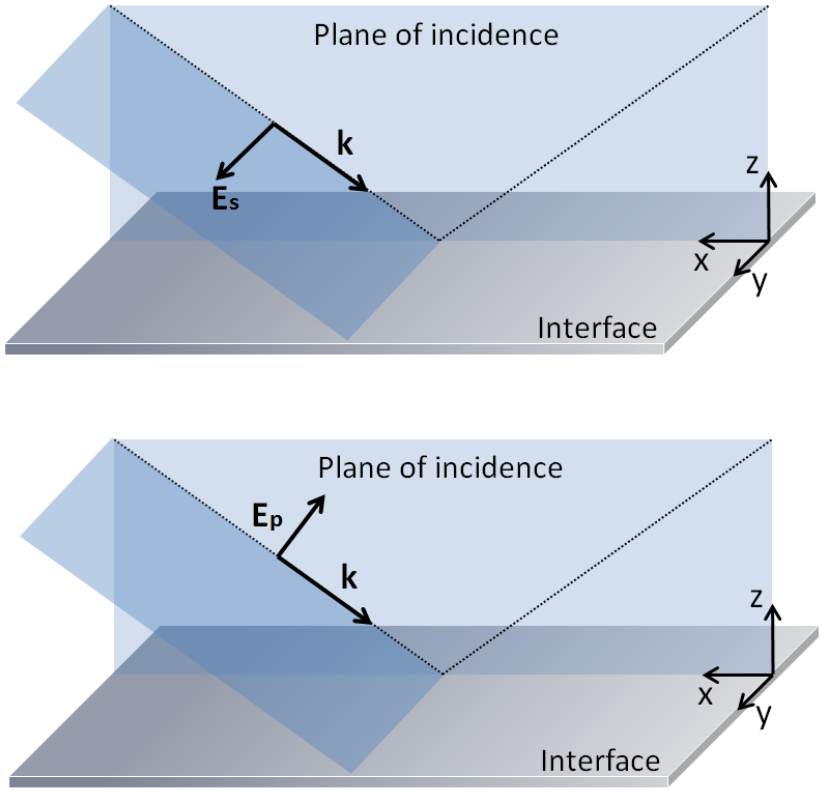

Figure 3.10: *Top: s-polarised light.* ***s*** *comes from German 'senkrecht', meaning perpendicular to the plane of incidence. Bottom: p-polarised light.* ***p*** *comes from German parallel, so parallel to the plane of incidence. Figure adapted from [58].*

According to the Cartesian coordinate system shown in fig. 3.10, the electric fields in the x, y and z directions can be written as follows:

$$\begin{aligned} \mathbf{E}_x^I &= E_x^I \hat{\mathbf{x}} \rightarrow E_x^I = \pm E_p^I \cos(\theta_i) \\ \mathbf{E}_y^I &= E_y^I \hat{\mathbf{y}} \rightarrow E_y^I = E_s^I \\ \mathbf{E}_z^I &= E_z^I \hat{\mathbf{z}} \rightarrow E_z^I = E_p^I \sin(\theta_i) \end{aligned} \tag{3.15}$$

where $\hat{\mathbf{i}}$ are the unit vectors for the cartesian axes x,y and z. The reflected and transmitted electromagnetic fields, $\mathbf{E}^R$ and $\mathbf{E}^T$ respectively, can be also calculated. The extent to which an electromagnetic field is transmitted or

reflected depends on polarisation, incident, reflected and transmitted angles (θ_i, θ_r and θ_t), and refractive indices of the two media (n_1 and n_2) (see fig. 3.11).

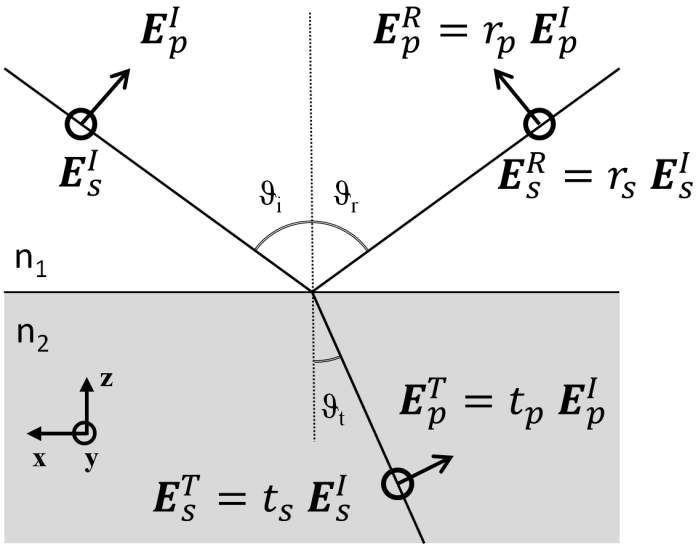

Figure 3.11: *s and p polarised incident E fields are transmitted and reflected at an interface between two media. Figure adapted from [58].*

The Fresnel amplitude coefficients of transmission t_p, t_s and reflection r_p, r_s can be easily derived:

$$\begin{aligned} t_p &\equiv \left(\frac{E_p^T}{E_p^I}\right) = \frac{2n_1\cos\theta_i}{n_1\cos\theta_t + n_2\cos\theta_i} \\ t_s &\equiv \left(\frac{E_s^T}{E_s^I}\right) = \frac{2n_1\cos\theta_i}{n_1\cos\theta_i + n_2\cos\theta_t} \\ r_p &\equiv \left(\frac{E_p^R}{E_p^I}\right) = \frac{n_2\cos\theta_i - n_1\cos\theta_t}{n_1\cos\theta_t + n_2\cos\theta_i} \\ r_s &\equiv \left(\frac{E_s^R}{E_s^I}\right) = \frac{n_1\cos\theta_i - n_2\cos\theta_t}{n_1\cos\theta_i + n_2\cos\theta_t} \end{aligned} \tag{3.16}$$

The total electric field at the surface is the sum of incident and reflected

beams. We can write each component as follows:

$$\begin{aligned}
E_x &= E_x^I + E_x^R = E_x^I + r_p E_x^I = \mp \left(E_p^I \cos\theta_i - r_p E_p^I \cos\theta_i \right) = \\
&= \mp E_p^I \cos\theta_i (1 - r_p) \\
E_y &= E_y^I + E_y^R = E_y^I + r_s E_y^I = E_s^I + r_s E_s^I = E_s^I \left(1 + r_s\right) \\
E_z &= E_z^I + E_z^R = E_z^I + r_p E_z^I = E_p^I \sin\theta_i + r_p E_p^I \sin\theta_i = \\
&= E_p^I \sin\theta_i \left(1 + r_p\right)
\end{aligned} \tag{3.17}$$

The electric field in the x direction can be positive or negative, according to the direction of propagation of one of the two fields. When $\mathbf{E}_1$ and $\mathbf{E}_2$ propagate in the same direction, then E_x in eq. 3.17 has a negative sign. The sign of E_x becomes positive in the opposite case. From eq. 3.14, it follows that a polarisation $\mathbf{P}_{SF}$ is generated at a frequency ω_{SF}:

$$\begin{aligned}
\omega_{SF} &= \omega_{VIS} + \omega_{IR} \\
\mathbf{P}_{SF}^{(2)} &= \varepsilon_0 \chi^{(2)} \mathbf{E}_{VIS} \mathbf{E}_{IR}
\end{aligned} \tag{3.18}$$

The SF signal is generated coherently at the interface once temporal and spatial overlap of the two fields is achieved. Because of its coherent nature, the direction of the emitted SF field is determined by the conservation of momentum:

$$\mathbf{k}_{SF} = \mathbf{k}_{\parallel,VIS} + \mathbf{k}_{\parallel,IR} \tag{3.19}$$

Eq. 3.19 is known as *phase matching* condition. The SF angle can be easily derived:

$$\theta_{SF} = \arcsin\left(\frac{n_{VIS} k_{VIS} \sin(\theta_{VIS}) + n_{IR} k_{IR} \sin(\theta_{IR})}{n_{SF} k_{SF}} \right) \tag{3.20}$$

where n_{SF},n_{VIS},n_{IR} are the refractive indices of the medium through which the corresponding electric field is propagating, θ are the angles to the surface normal, and $k = \frac{\omega}{c}$ (where c is the speed of light). Eq. 3.18 is valid for

all coordinate systems. To make it specific to the Cartesian one defined in fig. 3.11, the third rank tensor $\chi^{(2)}$ is defined using the same axes, where in general:

$$\mathbf{P}^{(2)}_{i,SF} = \varepsilon_0 \chi^{(2)}_{ijk} \mathbf{E}_{j,VIS} \mathbf{E}_{k,IR} \tag{3.21}$$

The indices *i,j,k* take the values *x,y,z*. The total second order polarization $\mathbf{P}^{(2)}_{SF}$ can then be expressed as shown in eq. 3.22:

$$\mathbf{P}^{(2)}_{SF} = \sum_{i}^{x,y,z} \mathbf{P}^{(2)}_{i,SF} = \varepsilon_0 \sum_{i}^{x,y,z} \sum_{j}^{x,y,z} \sum_{k}^{x,y,z} \chi^{(2)}_{ijk} \mathbf{E}_{j,VIS} \mathbf{E}_{k,IR} \tag{3.22}$$

3.2.3 K and L factors

Eq. 3.21 refers strictly to the fields at the surface, and not the propagating fields in space. Using eq. 3.16 and eq. 3.17, the electric fields at the surface can be written in terms of the s and p components of the incident beams. $\mathbf{E}_x$, $\mathbf{E}_y$ and $\mathbf{E}_z$ can be written as follows:

$$\begin{aligned}
\mathbf{E}_x &= E_x \hat{\mathbf{x}} = \mp E^I_p (1 - r_p) \cos\theta_i \hat{\mathbf{x}} = K_x E^I_p \hat{\mathbf{x}} \\
\mathbf{E}_y &= E_y \hat{\mathbf{y}} = E^I_s (1 + r_s) \hat{\mathbf{y}} = K_y E^I_s \hat{\mathbf{y}} \\
\mathbf{E}_z &= E_z \hat{\mathbf{z}} = E^I_p (1 + r_p) \sin\theta_i \hat{\mathbf{z}} = K_z E^I_p \hat{\mathbf{z}}
\end{aligned} \tag{3.23}$$

where K_x, K_y and K_z are defined (according to the eq. set 3.16) as:

$$\begin{aligned}
K_x = \mp(1 - r_p)\cos\theta_i = \mp\cos\theta_i \left(1 - \frac{n_2 \cos\theta_i - n_1 \cos\theta_t}{n_1 \cos\theta_t + n_2 \cos\theta_i}\right) = \\
= \mp \frac{2 n_1 \cos\theta_i \cos\theta_t}{n_1 \cos\theta t + n_t \cos\theta_i} \\
K_y = (1 + r_s) = \frac{2 n_1 \cos\theta_i}{n_1 \cos\theta_i + n_2 \cos\theta_t} \\
K_z = (1 + r_p)\sin\theta_i \hat{\mathbf{z}} = \frac{2 n_1 \sin\theta_i \cos\theta_i}{n_1 \cos\theta_t + n_t \cos\theta_i}
\end{aligned} \tag{3.24}$$

Also, the induced polarisation at the interface $\mathbf{P}^{(2)}_{i,SF}$ needs to be related to the emitted electric field $\mathbf{E}_{i,SF}$. These two vectors are linked to each other

via to the so-called nonlinear *Fresnel Factors* or L Factors. The L factors can be derived according to the continuity condition of the electric field across the interface and the phase matching condition for the SF emission [58].

$$\begin{aligned}
L_x^R &= -\frac{i\omega_{SF}}{c\varepsilon_0}\frac{\cos\theta_t^{SF}}{n_2\cos\theta_i^{SF}+n_1\cos\theta_t^{SF}} \\
L_y^R &= \frac{i\omega_{SF}}{c\varepsilon_0}\frac{1}{n_1\cos\theta_i^{SF}+n_2\cos\theta_t^{SF}} \\
L_z^R &= \frac{i\omega_{SF}}{c\varepsilon_0}\frac{\left(\frac{n_2}{n_{Layer}}\right)^2\sin\theta_t^{SF}}{n_1\cos\theta_t^{SF}+n_2\cos\theta_i^{SF}} \\
L_x^T &= -\frac{i\omega_{SF}}{c\varepsilon_0}\frac{\cos\theta_i^{SF}}{n_2\cos\theta_i^{SF}+n_1\cos\theta_t^{SF}} \\
L_y^T &= \frac{i\omega_{SF}}{c\varepsilon_0}\frac{1}{n_1\cos\theta_i^{SF}+n_2\cos\theta_t^{SF}} \\
L_z^T &= \frac{i\omega_{SF}}{c\varepsilon_0}\frac{\left(\frac{n_1}{n_{Layer}}\right)^2\sin\theta_i^{SF}}{n_1\cos\theta_t^{SF}+n_2\cos\theta_i^{SF}}
\end{aligned} \tag{3.25}$$

Knowing K and L factors, both p and s SF intensities can be written:

$$\begin{aligned}
I_{p,SF} &\propto \left|\mathbf{E}_{x,SF}+\mathbf{E}_{z,SF}\right|^2 \\
&\propto \left|L_x\mathbf{P}_{x,SF}^{(2)}+L_z\mathbf{P}_{z,SF}^{(2)}\right|^2 \\
&\propto \left|L_x\varepsilon_0\sum_j^{x,y,z}\sum_k^{x,y,z}\chi_{xjk}^{(2)}K_j\mathbf{E}_{j,VIS}^I K_k\mathbf{E}_{k,IR}^I + L_z\varepsilon_0\sum_j^{x,y,z}\sum_k^{x,y,z}\chi_{zjk}^{(2)}K_j\mathbf{E}_{j,VIS}^I K_k\mathbf{E}_{k,IR}^I\right|^2
\end{aligned} \tag{3.26}$$

$$\begin{aligned}
I_{s,SF} &\propto \left|\mathbf{E}_{y,SF}\right|^2 \\
&\propto \left|L_y\mathbf{P}_{y,SF}^{(2)}\right|^2 \\
&\propto \left|L_y\varepsilon_0\sum_j^{x,y,z}\sum_k^{x,y,z}\chi_{yjk}^{(2)}K_j\mathbf{E}_{j,VIS}^I K_k\mathbf{E}_{k,IR}^I\right|^2
\end{aligned} \tag{3.27}$$

Since the intensity of the SF signal depends on L and K factors (as shown in eq. 3.26 and eq. 3.27), A calculation of these helps to find the experimental

conditions to maximise the SF signal. In fig. 3.12 we show the intensity of the sum frequency signal in ppp polarisation for the air/gold interface in the -CH stretching spectral region (IR wavelength $\sim$3300cm^{-1}) and for the air/water interface in the amide I region in ssp polarisation combination.

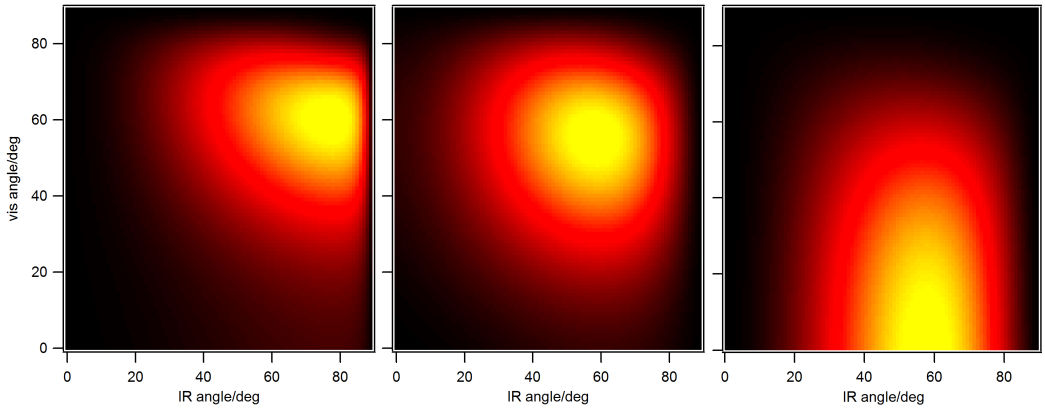

Figure 3.12: *Left: air/gold interface, ppp polarisation combination,* $\lambda_{vis} = 800nm$,$\lambda_{IR} = 3300nm$, I_{max}*=0.355. Centre: air/silicon interface, ppp polarisation combination,* $\lambda_{vis} = 800nm$,$\lambda_{IR} = 3300nm$, I_{max}*=0.089 . Right: air/water interface, ssp polarisation combination,* $\lambda_{vis} = 800nm$,$\lambda_{IR} = 6000nm$, I_{max}*=0.085*

3.2.4 Determining molecular orientation

We have shown that SFS is a surface specific spectroscopic method. Eq. 3.26 and eq. 3.27 show the dependence of SF intensity on the various components of $\chi^{(2)}$, which represents the macroscopic average of the molecular hyperpolarizability β.

β and therefore $\chi^{(2)}$ are strongly dependent on the IR (ω_{IR}) frequency because they account for the molecular vibrations. An accurate description on how to relate β and $\chi^{(2)}$ is discussed in appendix A. The hyperpolarisability

β for a single vibrational resonance can be expressed as:

$$\beta \propto \frac{1}{\omega_{IR} - \omega_0 + i\Gamma} \cdot \frac{\partial \alpha}{\partial Q} \cdot \frac{\partial \mu}{\partial Q} \tag{3.28}$$

where ω_{IR} is the IR frequency, ω_0 and Γ are, respectively, the frequency and the width of the resonance, α is the Raman polarisability, μ is the dipole moment and Q are the vibrational coordinates. Briefly, at the molecular level we can define an axis system within the molecule *(a,b,c)*, in contrast to the Cartesian system defined at the surface *(x,y,z)* as shown in fig. 3.13. For a molecule tilted by an angle θ, the relationship between the E fields in the (x,y,z)-coordinate system and the (a,b,c)-coordinate system can be described simply by a rotation matrix with a single polar angle θ.

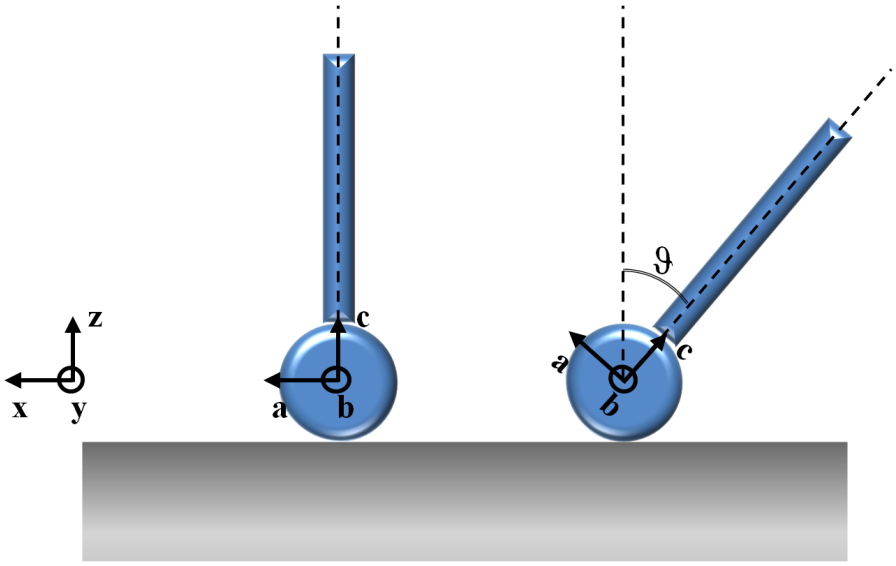

Figure 3.13: *Molecule with C_∞ symmetry shown on a surface. Left: the molecule stands upright at the surface, the molecular axis system and the surface cartesian axis system coincide. Right: the molecule is tilted by an angle θ. Figure adapted from [58].*

Due to symmetry, some of the $\beta_{\alpha\beta\gamma}$ vanish. For example a C_∞ has only four independent components ($\beta_{aac} = \beta_{bbc}$, $\beta_{aca} = \beta_{bcb}$, $\beta_{caa} = \beta_{cbb}$ and

β_{ccc}). The non-vanishing elements result in a certain SF intensity at a certain frequency. If the molecular structure is known (as it is in most cases), then the molecular orientation can be determined observing two or more vibrations within the same molecule. As an example, we discuss here the behaviour of octadecanethiol (fig. 3.14) self assembled monolayers (ODT SAMs) on gold surfaces in contact with solvents of a different nature, using the symmetric and asymmetric stretches of the methyl group.

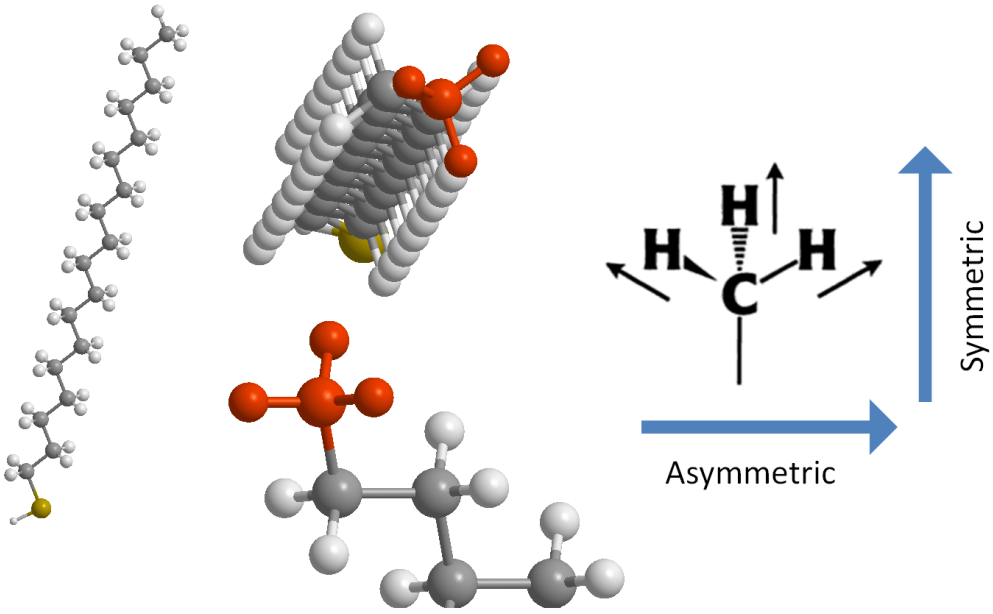

Figure 3.14: *Structure of octadecanethiol. The methyl -CH_3 head group is shown in red. Right: black arrows show the direction of the dipole moment corresponding to the asymmetric and symmetric vibration of the methyl head group.*

ODT is a long alkane chain, terminated on one side with a thiol group (-SH) and on the opposite side with a methyl group (-CH_3) (highlighted in red in fig. 3.14). The preparation of an ODT SAM on a gold surface is well established, and discussed in [59] [60] and appendix B. Freshly prepared

SAMs are measured through a prism, and a thin layer of solution is placed in between the SAM surface and the prism (as shown in fig. 3.15–top). We record the vibrational spectrum in the -CH stretching region (around 2900 cm^{-1}). In SFG, only the methyl end group is "visible" because (if the chain is well ordered) most of the contributions from the methylene groups cancel out due to an inversion centre between adjacent groups and unpaired -CH_2's tend to only give a weak signal [61].

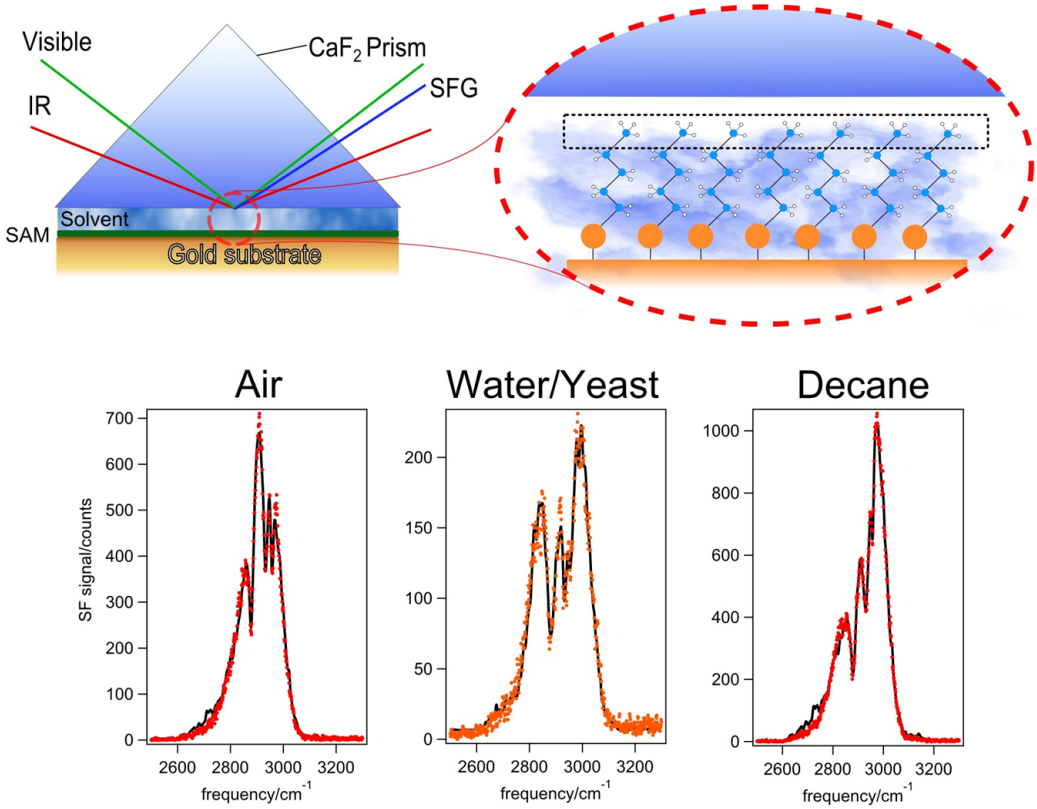

Figure 3.15: *Top: Schematic of the experimental layout: ODT SAMs are prepared on a gold surface. A layer of liquid is placed in between the gold surface and a CaF_2 prism. Bottom: SF spectra obtained with 3 different liquid interlayers. From left to right: air, water/yeast and decane respectively.*

Different spectra are recorded, wetting the SAM with water, decane and

yeast (fig. 3.15–bottom, from left to right). These spectra show three clear features at frequencies of 2877 cm^{-1}, 2932 cm^{-1} and 2958 cm^{-1} [62]. These three spectral features correspond to the symmetric stretching vibration, the Fermi resonance and the asymmetric stretching of the methyl head group. These three vibrations of the methyl group in the ODT molecule on Au result in a dip, due to phase interference between the resonances and the nonresonant signal generated from the gold surface [57]. The relative intensity of the symmetric and asymmetric spectral features can be used to determine the orientation of the methyl head group [61] from the factor R [63]

$$R = \frac{I_{symmetric} + I_{FermiResonance}}{I_{asymmetric}} \tag{3.29}$$

The relationship between R and the tilt angle can be derived from the molecular hyperpolarisabilities $\beta_{\alpha\beta\gamma}$ and a coordinate transformation.

Here we obtain R=2.7 for air, 1.9 for decane and 14.1 for the yeast/sugar mixture. Bulard *et al.* [63] determined R=3.7 for air, 5.8 for water and 8.8 for hydrophilic bacteria, corresponding to -CH_3 tilt angles of 55°, 53° and 50° respectively. We obtain the same trend, larger$R$ for more hydrophilic environment, meaning a more upright -CH_3 group. Detailed values differ as we did not fit the -CH_2 stretches here.

3.3 Quartz Crystal Microbalance with Dissipation – QCM-D

In the previous sections, we discussed two different spectroscopic methods to determine protein conformation and molecular structure at interfaces. Even though FT-IR has quantitative capability, it is not straightforward to determine the amount of adsorbed protein from an IR spectrum. This issue

has been thoroughly discussed in section 3.1 and appendix C. To achieve better accuracy regarding the amount of adsorbed protein on a surface, it is necessary to use different techniques. One of the possible solutions is to use a Quartz Crystal Microbalance with Dissipation (QCM-D) [64]. A QCM-D device functions by measuring the resonance of a quartz crystal. A quartz crystal can be electrically driven at its intrinsic resonance frequency. In QCM-D devices this is most commonly the *thickness shear deformation* (fig. 3.16).

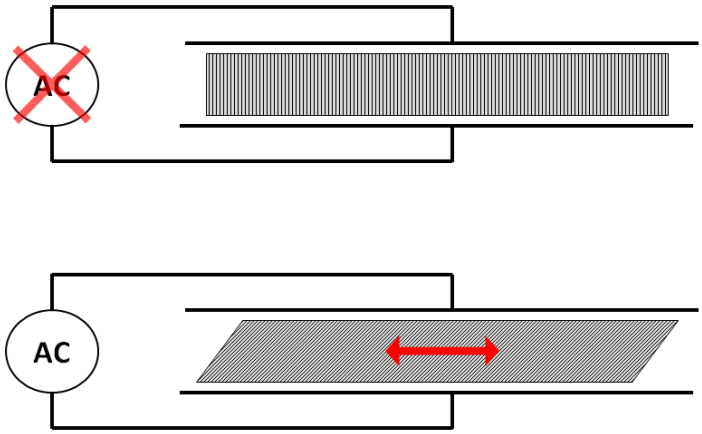

Figure 3.16: *Top: without external potential the quartz crystal does not show any deformation. Bottom: the quartz crystal deforms when an AC electric signal is applied to the electrodes.*

The external AC field induces the crystal to oscillate. When the thickness of the crystal t_q is twice the acoustical wavelength, a standing wave can be established, where the inverse of the frequency of the applied potential is of the period of the standing wave. This frequency is called the resonant

frequency f_0 and is given by the equation

$$f_0 = \frac{\sqrt{\frac{\mu_q}{\rho_q}}}{2t_q} \tag{3.30}$$

where μ_q is the shear module, ρ_q is the density of the crystal and t_q is the crystal thickness. Crystals' resonance frequencies can vary from few MHz to hundreds of MHz, dependent on the shape, size and quartz crystal orientation [65]. Commercial available crystals have resonance frequencies of few MHz (typical 5 MHz) and a variety of coatings is available (Gold, Silicon, Cellulose and so on) [66] [67]. The resonance frequency depends not only on the geometrical properties of the crystal, but also on its mass. Therefore, if any external mass is added onto the crystal, the resonance frequency will shift. This frequency shift can be measured with extreme accuracy, and the deposited mass can be indirectly determined. Despite the simplicity of the functioning principle of a QCM, this method is extremely powerful. Masses of the order of $\sim$ng/cm^2 can be determined.

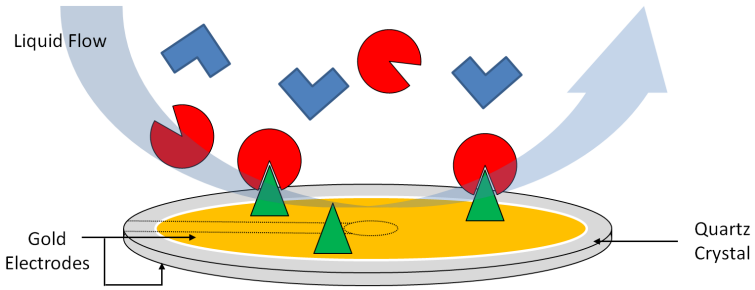

Figure 3.17: *QCM-D can be used to selectively identify specific molecules adsorbing on a surface. Reversible and irreversible adsorption can be monitored.*

The simplest model for describing the frequency shift as a function of mass dates back to 1959, and was proposed by Sauerbrey [68]

$$\Delta f = n\frac{-2f_0^2}{A\sqrt{\mu_q \rho_q}}\Delta m \tag{3.31}$$

where f_0 is the resonance frequency of the crystal, μ_q and ρ_q are respectively the shear module and the density of the quartz crystal, A is the piezoelectrically active area, Δm is the mass variation and n is the n-th overtone. The Sauerbrey equation (Eq. 3.31) is valid when Δf is small, and when the film adsorbed onto the crystal is rigid. In the case of proteins in liquid environments, other models should be used, since the rigid properties of the adsorbed layer are not known 'a priori' [69] [70] [71]. When measuring the frequency shift with a QCM, one can measure various overtones. In fig.3.18 we show a measured set of data with the various measured overtones normalised by n. The reader can appreciate that Δf does not scale with n, and therefore, the Sauerbrey model is not applicable. This is typical for an adsorbed viscoelastic film which acts as a coupled oscillator rather than an additional mass. By measuring both frequency and dissipation of the various overtones, Höök *et al.* [72] developed the so called Voight model, which takes into account the viscoelastic properties of the film. This is now implemented as a fitting routine on commerical instruments.

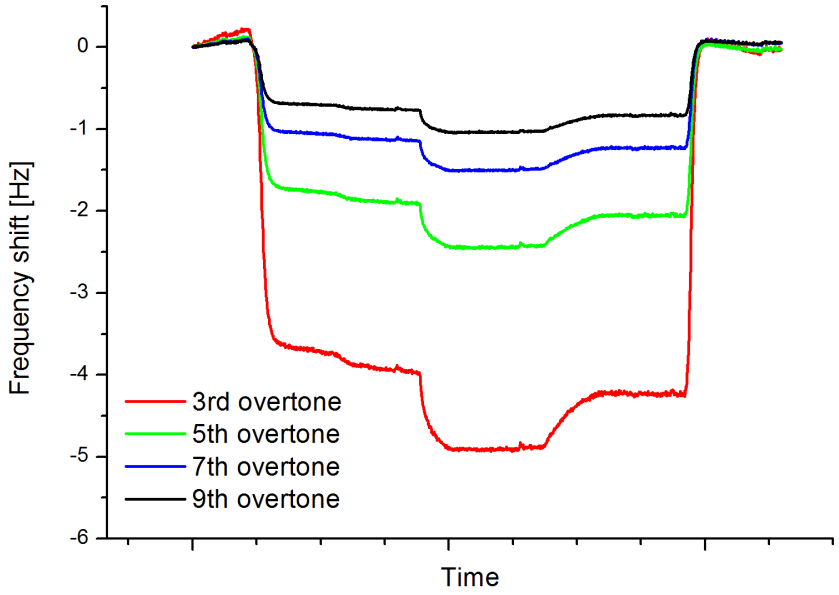

Figure 3.18: *Insulin adsorption on an APTMS functionalised silicon sensor. In the graph are reported the normalised 3rd, 5th, 7th and 9th measured overtones. Since the protein layer is not rigid, the Sauerbrey model is not valid anymore.*

QCM-D is the perfect instrument to measure with precision the amount of adsorbed proteins. When comparing it to other techniques though, care has to be taken regarding the interpretation of the derived mass. QCM measures the mass of protein and its coupled water layer, while eg. FT-IR measures the amount of protein present within the penetration depth of the evanescent field.

3.4 X-ray Photoemission Spectroscopy – XPS

In the previous section, we have discussed how to quantify a layer of adsorbed layer of protein using QCM-D. Two issues still persist: as mentioned the quartz crystal microbalance measures the wet adsorbed mass and there is no chemical sensitivity. To overcome these problems, X-ray Photoemission

Spectroscopy (XPS) can be used. XPS is a rather simple technique which combines surface specificity, chemical sensitivity and versatility to be used with a large variety of solid samples. The principle of XPS consists of irradiating a surface with X-rays [73]. Because of the photoelectric effect, some of the X-rays will cause the emission of electrons. The emitted electrons are collected, and their energy is measured, as schematically sketched in fig. 3.19

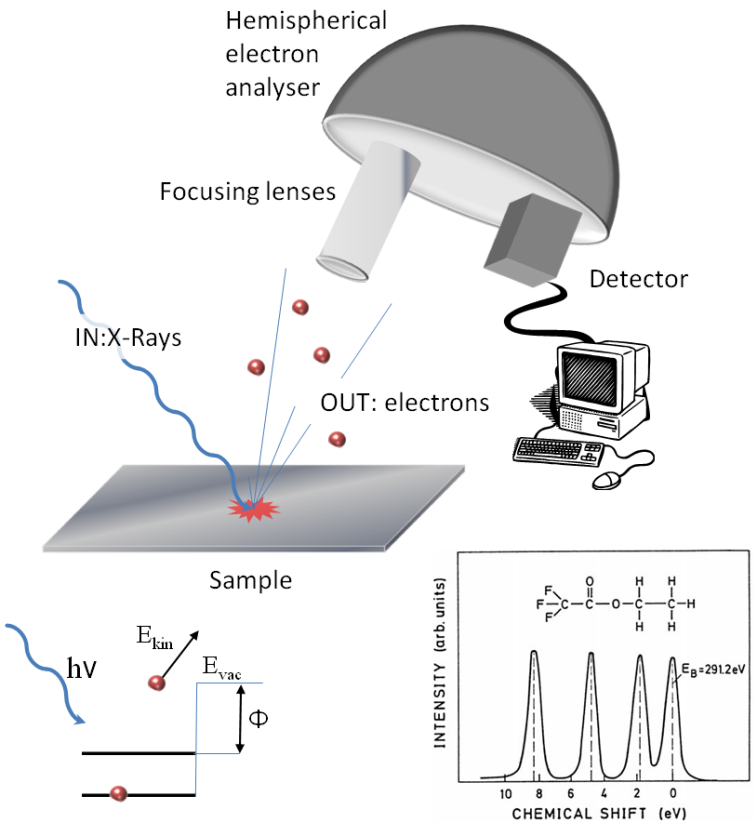

Figure 3.19: *A surface is irradiated by X-rays and then emits electrons. The emitted electrons are focused and collected thanks to a hemispherical electron analyser. The signal is detected, and a XP spectrum recorded. Bottom-Left: energy scheme for the photoemission process. Bottom-Right: XP spectrum (carbon 1s) of a complex molecule with different species of carbon. The C1s binding energy shift according to the different chemical surroundings [74]*

The energy of the emitted electrons is univocally related to the atoms present in the sample. In general, the kinetic energy of the emitted electrons E_{kin} can be written as function of the work function of the sample Φ, the binding energy of the electrons E_B and the X-ray photon energy $h\nu$.

$$E_{kin} = h\nu - \Phi - E_B \tag{3.32}$$

XPS' surface sensitivity comes from the limited electron escape depth when the energy of the electrons is between 10-1000 eV. Almost independently on the material, any photoemitted electron will not escape from a distance more than a few nm at maximum [75] [76] [77]. XPS can be used to estimate film thicknesses on surfaces as the intensity of substrate electrons decays exponentially with increasing overlayer thickness. In chapter 5 we have used the empirical method proposed by Cumpson [78] to derive film thicknesses from XPS peak ratios. While XPS is often used to identify qualitatively the chemical composition of an adsorbed layer of protein, Ray and Shard [79] have demonstrated the possibility of using XPS to quantify the thickness of protein layer based on its nitrogen content. The layer thickness obtained thus can be compared with thicknesses measured with other methods (for example QCM). In accordance with the literature, we found that XPS systematically measures lower protein adsorbed mass than QCM, which is due to the missing hydration layer of the protein.

3.5 Reflection Anisotropy Spectroscopy

Reflection Anisotropy Spectroscopy (RAS) allows the detection of adsorbates which induce an in-plane asymmetry of the surface. Its functioning principle relies on the fact that the reflection of light off a surface depends on the polarisation of the incident light. RAS measures the difference in reflectivity

of light polarised along two orthogonal axes within a surface plane, as shown in eq. 3.33:

$$\frac{\Delta_r}{r} = \frac{2(r_x - r_y)}{r_x + r_y} \tag{3.33}$$

RAS is a linear optical technique, which does not have intrinsic surface sensitivity, because the light incident on the sample usually penetrates for few atomic layers into it, and therefore the signal obtained comes from the bulk material, too. This arguments is valid in general, but not when the substrate is chosen specifically so that there is no variation in reflectivity along the two defined axes of polarisation arising from the bulk material. In this case, any measured $\frac{\Delta r}{r}$ will arise eventually only from the surface. As shown in fig. 3.20, the RAS setup consists of a white light source followed by a polariser. The polarised light is reflected off a sample. The reflected light is sent through a photoelastic modulator (PEM) and a polariser (analyser). The signal is sent into a monochromator and detected by using a photomultiplier, and a RA spectrum is recorded.

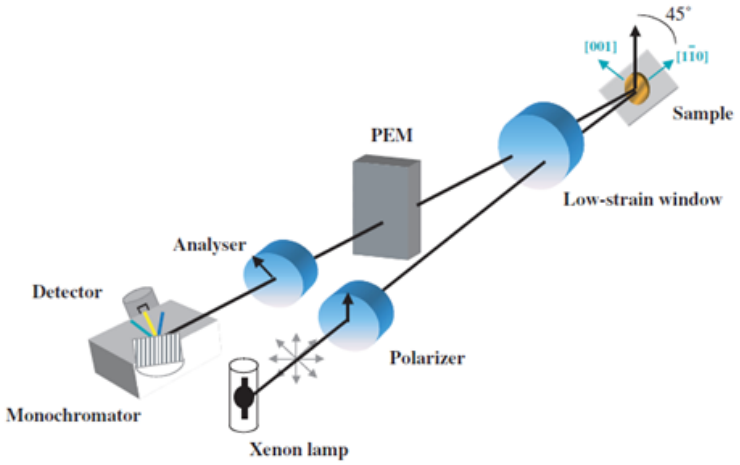

Figure 3.20: *A typical RAS setup is shown. Picture adapted from Weightman et al. [80].*

The PEM introduces a sinusoidal modulation in one of the components of the reflected beam. Each optical device is set at a certain angle to one of the sample's axes. In a standard RAS configuration, the angles $\theta_{Polariser}$,θ_{PEM} and $\theta_{Analyser}$ are set respectively to 0°, 45° and 0°. The complete theory has been described elsewhere by Lane *et al.* [81]. The measured intensity is a superposition of the fundamental I_0 and the second harmonic $I_{2\omega}$ of the PEM. The ratio between these two quantities can be written as:

$$\frac{I_{2\omega}}{I_0} = -2J_2 \frac{Re[r_{ss}^* r_{sp}]}{|r_{ss}|^2 + |r_{sp}|^2/4} \approx -2J_2 Re\left[\frac{r_{sp}}{r_{ss}}\right] \tag{3.34}$$

where J_2 is the second order Bessel function (of the first kind) and r_{ij} are the elements of the reflection matrix R^{sp}:

$$R^{sp} = \begin{pmatrix} r_{ss} & r_{ps} \\ r_{sp} & r_{pp} \end{pmatrix} \tag{3.35}$$

If the sample is set at an angle $\theta_s = \pi/4$, the matrix R^{sp} can be re-written as

$$R^{sp} = \begin{pmatrix} r & \frac{\Delta_{ps}}{2} \\ -\frac{\Delta_{ps}}{2} & r \end{pmatrix} \tag{3.36}$$

and eq.3.34 simplifies to:

$$\frac{I_{2\omega}}{I_0} = 2J_2 \frac{Re[r^* \Delta_r]}{|r|^2 + |\Delta_r|^2/4} \approx 2J_2 Re\left[\frac{\Delta_r}{r}\right] \tag{3.37}$$

hence the signal measured is directly proportional to the quantity $\frac{\Delta_r}{r}$, making RAS surface specific. The RAS signal posseses also an azimuthal dependence as shown by MacDonald *et al.* [82]. When one of the overlayer's axes is aligned with the surface normal, eq. 3.37 can be rewritten as:

$$\frac{I_{2\omega}}{I_0} = 2J_2 \left[\Re \sin 2\theta_s + \frac{\Im^2 - \Re^2}{4} \sin 4\theta_s\right] \tag{3.38}$$

where $\Re$ and $\Im$ are respectively the real and imaginary part of $\frac{\Delta_r}{r}$. Neglecting higher order effects, the RAS signal varies with $\sin 2\theta_s$. This fact is often used to identify the unknown axis for sample alignment, as we show in fig. 3.21.

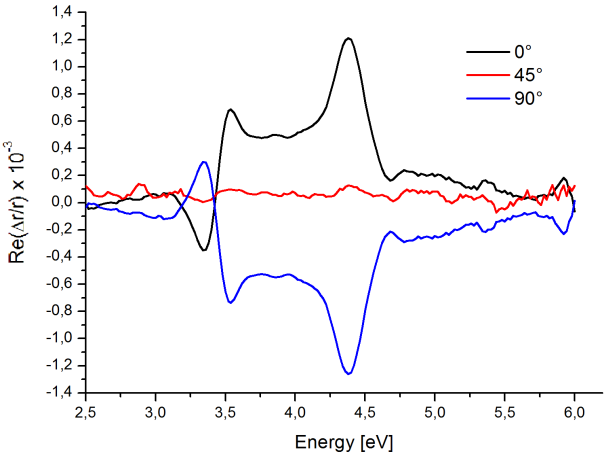

Figure 3.21: *Azimuthal dependence of RAS spectra of a bare Si(111)-vicinal subrate with native oxide layer.*

Fig.3.21 shows two peaks at 3.5 eV and 4.5 eV. These two peaks arise from the anisotropic stress caused by steps of the Si(111)-vicinal surface [83] [84] [85].

3.6 Determining the thickness of adsorbed protein layers

Some of the methods presented in this chapter, can be adopted to quantitatively determine the amount of protein adsorbed on solid surfaces. In particular, XPS, QCM-D and ATR FT-IR together represent a powerful toolbox for determining the amount of dry and wet protein on a surface, and its secondary structure too. The experimental approach presented here, is used in chapter 5 and also discussed in appendix C. A QCM-D is first used to quantify the amount of adsorbed protein, including its hydration layer. A QCM-D measurement provides (after analysing the frequency curves with the Voight model [72]) a area mass density, which can be further converted in film thick-

ness by considering the protein density. As a second step, the same surfaces used in the QCM-D experiments can be measured with XPS. In particular, for quantifying the dry amount of protein adsorbed, we used the method proposed by Ray and Shard [79]. Detail can be found in their manuscript, but in conclusion the simplified model used relies on using eq.3.39 during the integration of the nitrogen peak of measured XPS spectra.

$$d_{protein} = -L_{N_{1s}} \cos(\theta) \left(1 - \frac{[N] - [N]_0}{[N]_\infty - [N]_0}\right) \tag{3.39}$$

Where $L_{N_{1s}}$ is the electron attenuation lenght[1], $[N]$ is the measured fraction of nitrogen, $[N]_0$ is the fraction of nitrogen in the surface (SAM) $[N]_\infty$ is the fraction of nitrogen in the pure protein ($\approx 15\%$) and $\cos\theta$ is the relative angle between the sample and the electron analyser. The successive step is to use ATR FT-IR to determine the protein thickness both in presence or absence of the hydration layer (dry vs. wet). To do so, we refer to the equations presented in section 3.1.3. For dry protein layer, we can derive the thickness Γ after rearranging equation 3.7:

$$\Gamma = \frac{A}{N \cos\gamma} \cdot \frac{1}{\varepsilon} \cdot \frac{n_1}{4n_2} \cdot \sqrt{1 - \frac{n_3^2}{n_1^2}} \tag{3.40}$$

where n_1, n_2 and n_3 are the refractive indeces of the silicon ATR crystal, the protein layer and the air (in the case of dry) or water/solution (wet) protein. For the wet spectra one has to take in account for the adsorption of insulin in solution. This could be easily calculated from insulin solution spectra. In fact we can calculate the absorbance due to insulin solution within the penetration depth d_p, and then subtract this value from the whole absorbance. An example of this calculation is given in appendix C.

[1]L can be calculated as proposed by Green *et al.* [86] as $L = 0.00837 \cdot E^{0.842}$, where E is the binding energy in eV, and L is given in nm.

Chapter 4

The structure of insulin at the air/water interface: monomers or dimers?

These series of experiments has been carried out at the Max Planck Institute for Polymer Research, Mainz, Germany in the Molecular Spectroscopy Group of Prof. M. Bonn, under the supervision of Dr. T. Weidner
This collaboration has been funded through a 4-month DAAD scholarship.
This section is adapted from a publication:

"The structure of insulin at the air/water interface: monomers or dimers?"
S. Mauri, T. Weidner and H. Arnolds
Phys. Chem. Chem. Phys., 2014,16, p26722-26724
DOI: 10.1039/C4CP04926H

Here we study the oligomeric composition of insulin at the air/water interface, the simplest hydrophobic interface, to clarify the effect of hydropho-

bicity on the oligomeric composition of insulin at interfaces.

In section 2.2, we discussed the existence of various insulin oligomers. In solution, insulin is present as monomers, dimers and hexamers [6], but only monomers undergo denaturation and aggregation. Nilsson *et al.* [9] demonstrated that monomeric insulin adsorbs irreversibly on hydrophobic silica surfaces and concluded that the hydrophobic interaction is likely to be the driving force behind the irreversible protein adsorption. Sluzky and coworkers [87] studied the behaviour of insulin on teflon surfaces by means of light scattering (QELS) and found an increase in insulin stability at higher concentrations due to the presence of dimers/hexamers at the surface, which then reduce the available space for monomers to unfold. The insulin monomer therefore is the key species in aggregation and fibril formation at interfaces. However its presence is difficult to ascertain. Linear vibrational spectroscopy is not capable of distinguish between monomers and dimers. The only real structural difference between monomer and dimer would be the interfacial β-sheet that forms upon dimerisation, but its corresponding signal is expected to be very small compared to the rest of the protein, since only 8 amino acids in the B chain of insulin (B23-B30) [6] are involved, and amide I intensity of antiparallel β-sheets is smaller than for α-helices [88]. More sophisticated methods, eg. 2D-IR, are needed [89] to detect the interfacial β-sheet. For bulk solutions, a detailed analysis of circular dichroism spectra can be used to derive the oligomeric composition [90] but this method is not sensitive enough for interfaces.

Johnson *et al.* [91] studied the structure of insulin in the bulk and at the air/water interface using CD, FT-IR and IRRAS spectroscopy. They concluded that insulin at interfaces is mostly composed of α-helices. Data presented in their work are hard to interpret though, since IR surface spectra

present features only slightly above the noise level and the water sub-phase was kept at insulin's isoelectric point, which causes precipitation. Here we use Sum Frequency Spectroscopy to identify which insulin oligomeric species are are present at the surface.

The relative amount of each oligomer in solution can be chemically controlled by varying the concentration of insulin, changing the pH or adding divalent metal ions [92]. At high concentrations, dimers are prevalent in solution, whereas at low concentrations only monomers are observed, as has been shown by Hvidt by means of light scattering [93]. Ganim *et al.* [89] have measured the self-association of insulin dimers by means of 2D-IR. The presence of dimers is negligible at concentrations lower than 10^{-4} M ($\sim$0.6 mg/ml). At higher concentrations (10^{-2} M - 60 mg/ml) the fraction of dimers increases significantly, up to 50%.

Specifically, we use insulin solutions at various concentrations, from 0.1 mg/ml to 50 mg/ml to have predominantly monomers (low concentration) or an equal amount of monomers and dimers (high concentration). By comparing SF spectra with calculated second-order susceptibility values for monomers and dimers, we show that only monomeric insulin is present at hydrophobic interfaces.

4.1 Experimental details

Solution preparation: Human insulin, with 0.3% Zn by weight content, has been provided by Sanofi-Aventis. The composition of the protein powder has been tested with mass spectroscopy, that shows a peak at 5808 Da, indicative of native insulin monomers. Insulin solutions are prepared in D_2O/HCl at pD2.7, at concentrations of 0.1, 1, 3, 10 and 50 mg/ml. After insulin is added,

the pD is carefully adjusted to a value of 2.7 by adding small volumes (few $\sim \mu$l) of concentrated HCl. A teflon trough (10cm diameter) is sonicated in acetone for 10 minutes and then ethanol for 10 minutes, and at last carefully rinsed with ultra pure water ($\sim 18.2\Omega m^{-1}$). Finally, it is dried and stored under nitrogen before the experiment.

SFS setup: A femtosecond pulse train at 800 nm wavelength ($\sim$ 45 fs, 1kHz repetition rate) is generated by a Ti:Sapphire amplifier (SpectraPhysics Spitfire ACE), seeded by a Ti:Sapphire oscillator (SpectraPhysics MaiTai). One part is sent to an etalon, to obtain a picosecond visible beam with a spectral resolution of $\sim$ 15 cm^{-1}. Another part is used to pump an OPA (TOPAS Prime, LightConversion) to generate broadband tunable IR light. The average power of the visible beam is 5 mW (at the specified resolution) and 2 mW for the IR beam (at $6\mu m$). The visible beam is focused loosely with a 20 cm focal length lens at the air/water interface. The IR beam is tightly focused with a 5 cm focal length CaF_2 lens. The two beams are temporally and spatially overlapped at the interface. The SF is then detected with a spectrograph and an intensified CCD camera. SFG intensity is proportional to $\left|\chi^{(2)}\right|^2$ (see eq. 3.26 and eq. 3.27). The polarisation of all three beams can be controlled. The experiment is done in a nitrogen atmosphere. The laser setup is shown in fig. 4.1. All spectra are collected using *ssp* polarization combination. The spectrograph is calibrated by measuring a lipid monolayer (DOPC) at the air/D_2O interface (see fig.4.2). Such lipid monolayer is characterised by the presence of a single peak at 1730 cm^{-1} (shown in fig. 4.2). The cleanness of the trough is checked by measuring milli-Q water in the amide region. The trough is considered clean only if no lipids spectrum is measured, but only the water bending mode.

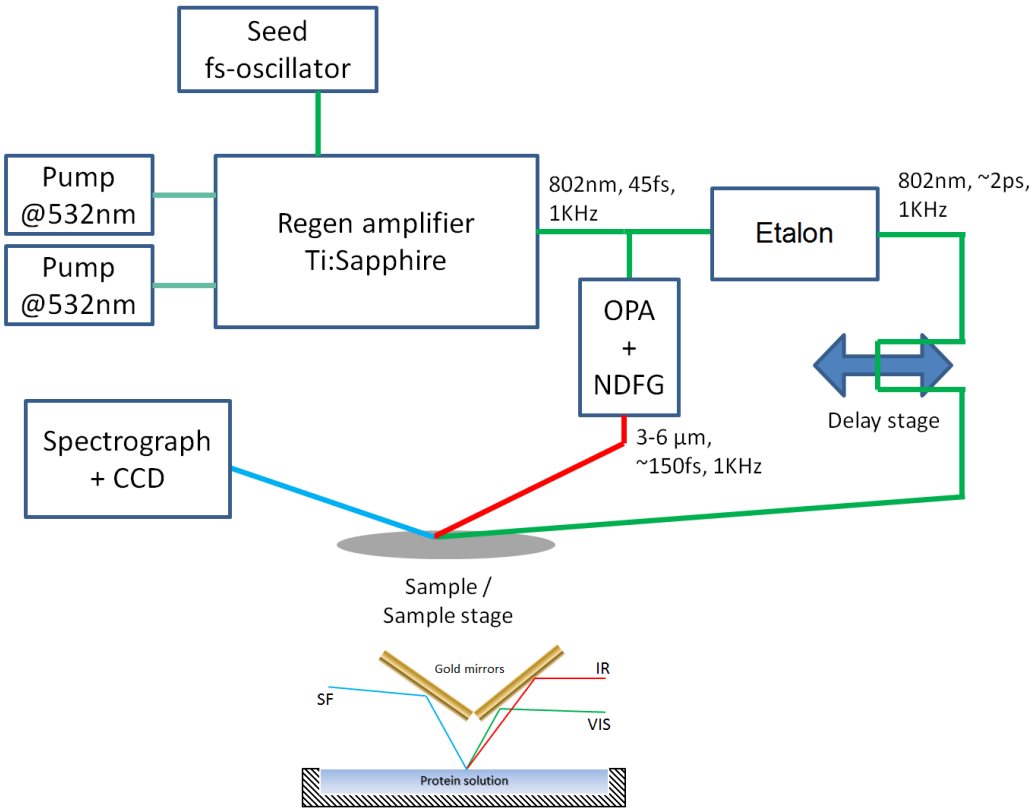

Figure 4.1: *Static SFG experimental setup.*

Data acquisition and processing: Each spectrum is the average of two spectra, acquired for 10 minutes each, from which we subtract a background (measured in the same conditions as the SF experiments, but blocking the IR beam)[1]. Spectral intensities have been normalised according to a SF nonresonant spectrum of a gold surface (the nonresonant spectrum of gold gives the spectral response of the SFG setup used.). All spectra have been measured with the same IR and visible power (2 mW and 5 mW respectively). The IR and visible power are stable within 1%, and the spectra can be therefore quantitatively compared.

[1]The lipid spectrum shown in fig.4.2 has been acquired for 5 minutes. The lipid layer is prepared at a surface pressure of ($\approx 15mN/m$)

4.2 Results and discussion

We measure insulin solution spectra at the air/water interface, at different protein concentrations. All solutions have been prepared in deuterated water, so that the amide I band does not overlap with the -OH bending mode of water. Small H_2O contamination are negligible, since the -OH bending mode is broad, and it overlaps with the amide I signal from insulin does not affect spectral interpretation. The spectra have been measured using *ssp* polarisation combination. In fig. 4.2 we show the recorded spectra which reveal that the SF intensity does not change significantly with concentration of bulk solutions (from 50 mg/ml to 0.1 mg/ml).

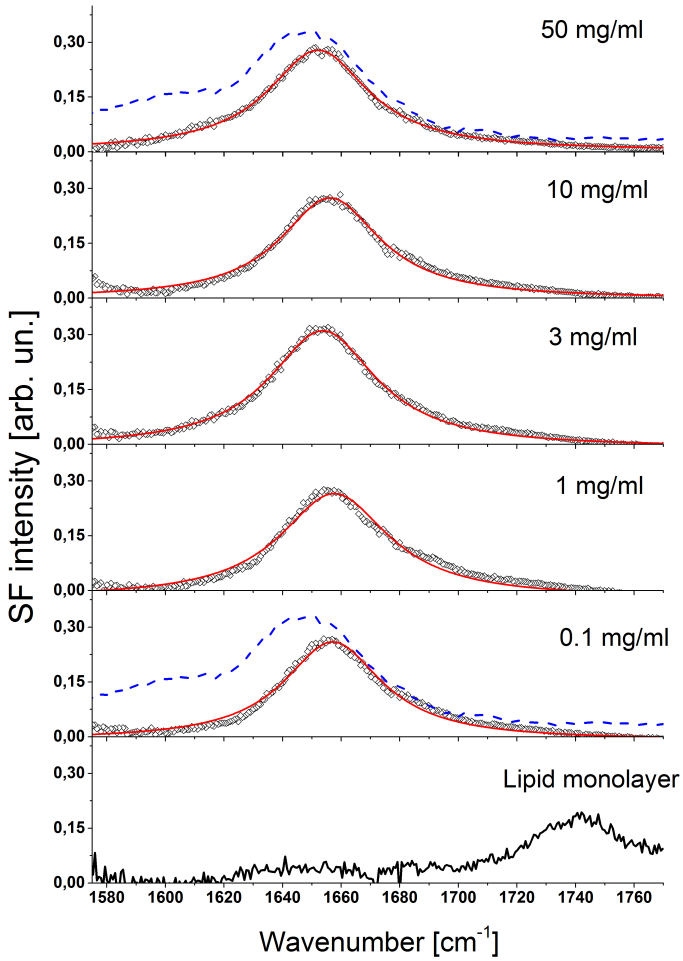

Figure 4.2: *SF spectra in the amide I region. Experimental data (squares) and fitting curve (red solid line) are shown. Spectra at concentrations of 50, 10, 3, 1 and 0.1 mg/ml are shown. Bottom spectrum: SF spectrum of a lipid monolayer at the D_2O/air interface. Blue dashed line: IR insulin spectrum in D_2O at 1mg/ml concentration.*

The amide I band lies in all cases at the same frequency, confirming the mostly helical structure of HI in its native state.

The peak intensities do not vary significantly, despite the bulk concentration of inulin spanning from 50 to 0.1 mg/ml (and the monomer/dimer

ratio being significantly larger in the lowest concentrated solution)[23]. This suggests that only one species of insulin saturates the interface already at low bulk concentration and this species is most likely the monomer.

The peak positions and widths have been fitted with a lorenzian single peak, and the fitting values are reported in the following table:

Conc. [mg/ml]	Peak position [cm^{-1}]	FWHM [cm^{-1}]	Peak intensity
50	1652±1	42±1	18.1±0.1
10	1655±1	42±1	18.0±0.2
3	1653±1	44±1	22.0±0.2
1	1656±1	45±1	20.0±0.3
0.1	1656±1	41±1	17.0±0.3

Table 4.1: *Fitting values for the SF spectra shown in 4.2.*

The symmetry inherent in the insulin dimer should render it almost invisible in a SF experiment. A first intuitive explanation relies on the fact that the two α-helices in the A chain have equally strong dipole moments aligned along the helix, but opposite in sign. Therefore these two contributions cancel each other out (see fig. 4.3). Only the α-helix in the B chain is "measurable" with SFS. In the case of a dimer, also the contributions from the helices in the B chain cancel, as they run in opposite direction . The only feature that could be SF active is the antiparallel β-sheet that forms at the

[2]Preliminary DLS measurement confirm the presence of two different species in solution at 50mg/ml and 1mg/ml. Further DLS measurement are in progress to confirm the size of the species in solutions.

[3]Private communication from Dr. Harald Berchtold: at a concentration of 50mg/ml at pH 2.7 the concentration of dimers is significantly higher than the one of monomers. The presence of hexamers is also possible.

hydrophobic interface between monomers in a single dimer. Therefore the total $\chi^{(2)}$ for an insulin dimer is expected to be quite small. The intensity of the various $\chi^{(2)}$ components can be calculated. Simpson's group at Purdue University [94] [95] has developed a plug-in to be used with the Chimera software [96] to compute the $\chi^{(2)}$ tensor elements of a protein in the lab coordinate frame given its .pdb file. We can thus compare the SF signal expected from dimers and monomers, assuming identical orientation for both oligomeric species. In fig. 4.3 we show the tensor elements calculated for both monomers and dimers. The calculation of these tensor elements is performed in the amide I region, assuming a achiral surface with C_∞, for which only 7 $\chi^{(2)}_{ijk}$ are non vanishing [97].The calculation of the tensor elements is done considering insulin monomers and dimers with a specific orientation. In particular we have set the orientation of monomers so that the hydrophobic patch it is exposed to the air water interface. This is the most likely way for monomers to orient, since it is the most energetic favourable one. For monomer we show how orientational changes do not affect significantly the value of $\chi^{(2)}_{ssp,dimer}$.

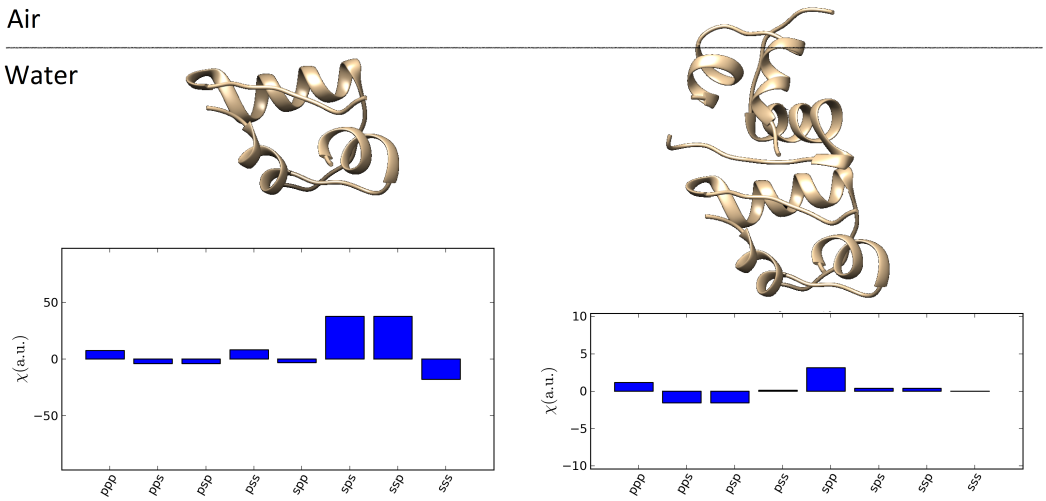

Figure 4.3: *Left: Top- insulin monomer shown as oriented at the surface. Bottom- $\chi^{(2)}$ tensor elements in the lab coordinate frame calculated for the monomer.*

Right: Top- insulin dimer shown with the same orientation as the monomer. Bottom- $\chi^{(2)}$ tensor elements in the lab coordinate frame calculated for the dimer.

The susceptibility component active in our polarisation combination, $\chi^{(2)}_{ssp,monomer}$ is 20 times higher for the monomer than for dimer. This equates to a 400 times lower SF intensity for a dimer compared to a monomer. Since a single dimer does not present any particular hydrophobic surface, we would not expect any specific orientation of the protein at the air/water interface [98]. So we also calculated $\chi^{(2)}_{ssp,dimer}$ for various other orientations of the dimeric species, as shown in fig. 4.4, and again $\chi^{(2)}_{ssp}$ remains at least an order of magnitude lower than the one for monomers.

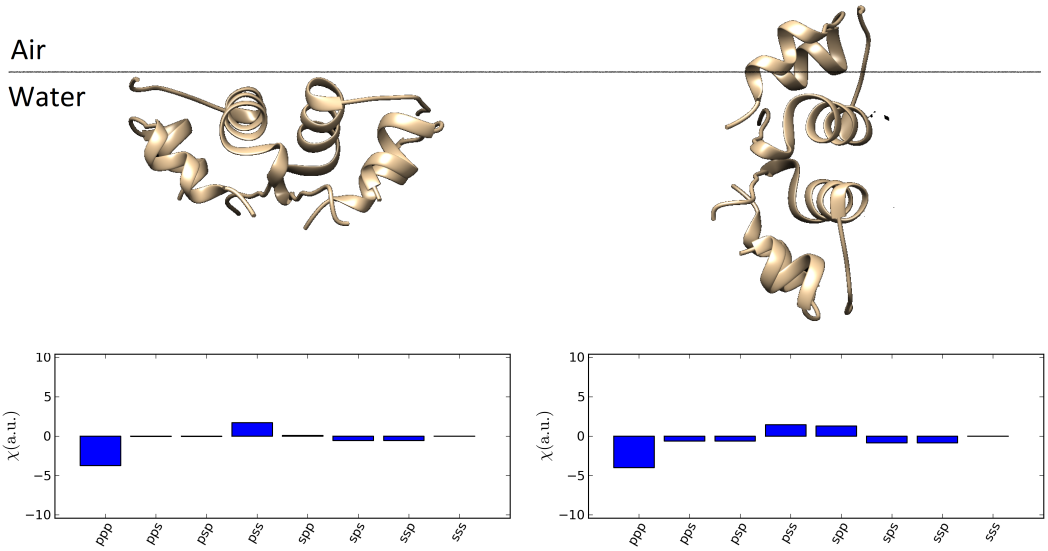

Figure 4.4: $\chi^{(2)}$ *tensor elements calculated for insulin dimers with different orientations.*

We can therefore confidently attribute the peak at 1653 cm^{-1} to insulin monomers at the air/water interface. We propose that insulin monomers segregate to the surface as sketched in fig. 4.5.

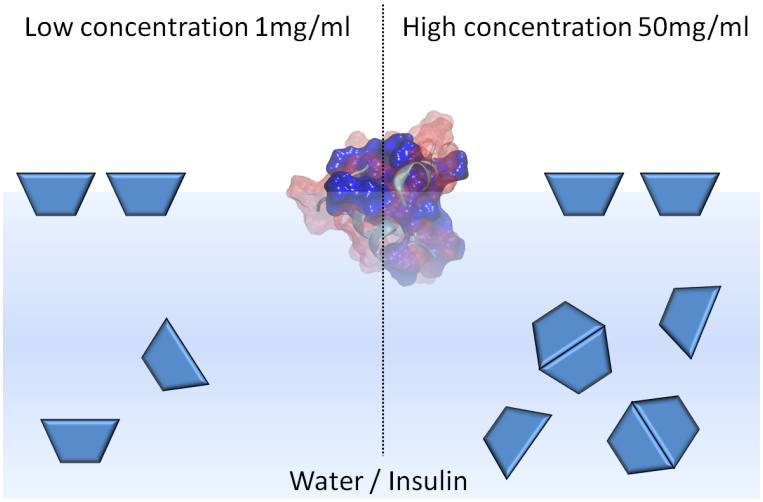

Figure 4.5: *Regardless of the bulk concentration, only monomers segregate to the surface.*

The segregation is most likely driven by the entropy gain of exposing the hydrophobic area of the protein surface to air [99]. Since monomers are known to be the least stable insulin species with regards to denaturation and aggregation, this surface segregation of the monomer is probably the deeper reason why agitation causes aggregation. The formation of a large air/water surface area on agitation will simply accelerate fibril formation [7]. In conclusion, we used SFG to determine that the insulin monomer segregates to the air/water interface by a quantitative comparison of spectra in the amide I region. This finding solves a long-standing puzzle why insulin has a propensity to form fibrils at the air/water interface. This interface segregation might also explain why other hydrophobic interfaces also enhance insulin fibrillation.

4.3 Future work

The results presented in this chapter support the idea that only monomers segregates at the air/water interface. To further proof the presented argument some other work can be done. Other polarisation combinations (e.g. ppp and sps) should be used, in order to determine an accurate orientation of the protein at the surface. We could then compare the whole set of measured spectra with the tensor elements, to verify that the orientation of insulin monomer is as we propose. On the other side, we can also chemically force hexamerisation of insulin in solution by adding Zn^{2+} ions. By doing that the signal we measure should disappear completely. This would support the hint that what we have measured is a truly insulin monomers SFG signal.

Chapter 5

Insulin adsorption on model hydrophilic and hydrophobic surfaces

The experiments presented in this chapter have been carried out at the University of Liverpool, in collaboration with Dr. Caroline Smith - Dept. of Physics (QCM-D) and with Dr. Benjamin Johnson - University of Leeds, Dept. of Physics. This section is adapted from paper ready for submission to Langmuir:

"Insulin Adsorption on Hydrophobic and Hydrophilic Silane Self-assembled Monolayers." *S. Mauri, M. Volk, S. Byard, H. Berchtold and H. Arnolds, submitted to Langmuir*

In the previous chapter we have shown how the hydrophobic nature of the air/water interface causes insulin monomers to segregate to it. The acceleration of amyloid formation by interfaces is generally attributed to stabi-

lization of partially folded intermediates, which go on to form different fibril morphologies from the bulk [100] [101] [102] Most of these prior works have been carried out at physiological pH=7.4, from which the following picture of insulin adsorption emerges.

Insulin monomers generally form a close-packed monolayer at hydrophobic interfaces, where they loose α-helical structure and gain random coil or β-sheet structure [9] [103] [104] [100] [98]. Conformational changes in adsorbed insulin appear to allow further binding from solution and pre-adsorbed peptides with a β-sheet structure were found to accelerate surface-induced insulin fibrillation [10] [11]. In solution, fibrils become more hydrophobic with time [105] and recent nano-FTIR and TERS suggest the presence of a β-sheet core with an α-helical shell [106] [107]. However, here we concentrate on acidic environments encountered during the synthesis and purification of insulin, where it exists only as monomers [108]. Working under these conditions, allows to focus our attention specifically on the interaction of monomers with surfaces.

Reichart *et al.* [109] proposed to use poly(acrylic acid) (PAA) brushes for adsorption of insulin as a protein-friendly substrate, since no fibril formation is observed at pD=2 even under high temperature conditions. PAA brushes are not easy to implement in the practical application of protein synthesis and storage, though. In general, the little attention bestowed upon surface characterization in other low pH studies and the scarcity of structural information for sub-monolayer to low multilayer insulin coverages, leaves open questions - what is the structure of adsorbed insulin monomers at low pH, an what is the nature of intermediates on the path to fibril formation? Surprisingly, here we find that insulin monomers keep their native structure upon adsorption at room temperature. Only after incubation at high tempera-

ture insulin monolayer form with a significant β-sheet content. These then provide the template for subsequent multilayer growth.

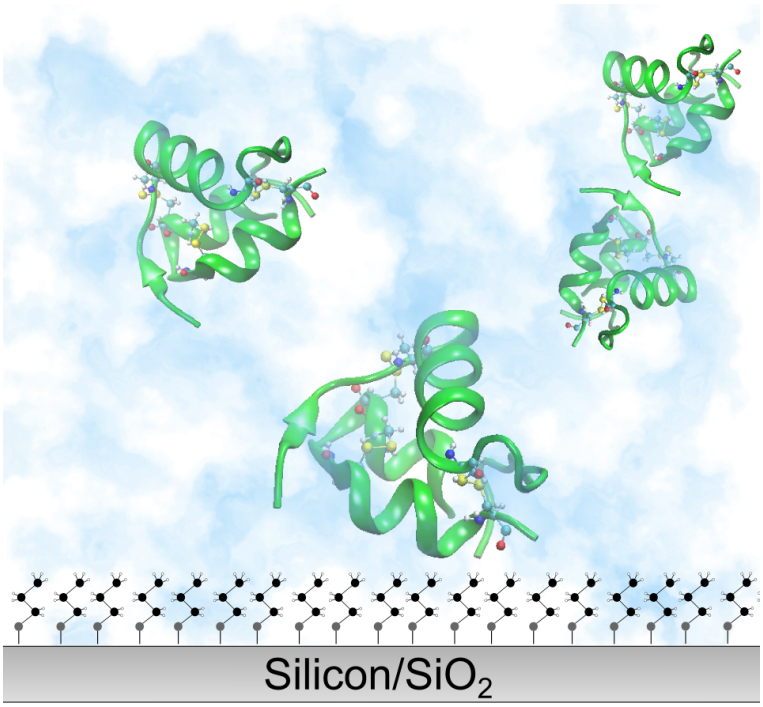

Figure 5.1: *Sketch of our surface layer (a methyl or amine-terminated propyl chain attached to SiO_2/Si) and two different orientations of insulin showing the B chain (long helix with 4 turns) connected via disulfide bonds (yellow) to the A chain (two helices).*

In this study, we use well-characterised hydrophobic and charged hydrophilic silane self-assembled monolayers (SAMs) formed on silicon surfaces and employ a combination of electronic and vibrational spectroscopies to elucidate the nature of the surface-insulin interaction. We compare adsorption of insulin at a hydrophobic surface to adsorption at a charged hydrophilic surface, which is known to keep porcine insulin physiologically active [110]. To understand insulin adsorption, we first quantify the adsorbed amount by comparing results from XPS, QCM-D and ATR FT-IR, we analyse the silane

SAM and insulin side-chain orientation by SFG and finally use ATR FT-IR to investigate the secondary structure of monolayer insulin at room temperature and later unfolding and multi-layer growth at elevated temperature.

5.1 Experimental details

5.1.1 Sample preparation

Si wafers and ATR crystal were cleaned using Piranha solution (70% H_2SO_4, 30% H_2O_2 solution (35% in water)) for 10 minutes at room temperature in a sonication bath. QCM crystals were cleaned using UV-ozone and only briefly dipped into cold Piranha solution (3 minutes at room temperature). After cleaning, all samples are thoroughly rinsed with copious amounts of ultrapure water (18.2 MΩ) and dried under Argon flow. All glassware is cleaned in Piranha solution at 70°C for 1 hour, rinsed several times with ultrapure water and ethanol, dried under a stream of dry air and finally rinsed with toluene. Trimethoxypropylsilane (TMPS) and 3-aminopropyltrimethoxysilane (APTMS) solutions were prepared in toluene at a concentration of 5 mM and 0.5 mM respectively. Low concentration of APTMS in solution is used to avoid polymerisation and/or multilayer formation. Silicon wafers are left in solution for 24 hours at room temperature, sonicated in toluene to remove potential multilayers and rinsed with copious amounts of ultrapure water. Finally, substrates are dried under an argon flow. Insulin solutions are prepared in Citrate-Phosphate buffer at pH 2.7, and at a concetration of 0.44 mg/ml. The buffer is prepared as follows: 44.6 ml of citric acid solution (0.1 M in D_2O) and 5.4 ml of dibasic sodium phosphate solution (0.2 M in D_2O) are mixed. The final volume is adjusted up to the final volume of 100 ml and the final pD is carefully adjusted by addition of small amounts of either

concentrated citric acid or dibasic sodium phosphate, by using a sensitive pH meter. Human insulin with 0.3% Zn by weight was kindly provided by Sanofi-Aventis Deutschland GmbH.

5.1.2 Spectroscopies

To monitor the amount of adsorbed protein under flow, we used quartz-crystal microbalance with dissipation (QCM-D) on a Q-sense E4 (Biolin Scientific) with silicon coated quartz sensors. XPS spectra of the sensors were recorded on a PSP Vacuum Technology analyser at 0.5 eV resolution for survey spectra and 0.025 eV resolution for analysis of binding energy shifts. The $2p_{1/2}$ level of argon embedded in the silicon crystal was used as internal energy standard with an energy of 241.2 eV [111]. FT-IR measurements were performed using a Bruker IFS 66v/S spectrometer. Each spectrum is averaged over 500 scans at 1 cm^{-1} resolution. We used an ATR setup from Harrick equipped with a silicon crystal (80 x 10 x 4 mm 45°, 11 reflections). FTIR-ATR spectra recorded in D_2O show a broad peak around 1647 cm^{-1}, and no amide II band around 1550 cm^{-1}, indicating complete H/D exchange [112]. Sum frequency spectroscopy (SFS) is a second-order nonlinear optical technique which intrinsically provides surface sensitivity and high spectral resolution and represents a coherent combination of an IR absorption and a Raman anti-Stokes transition. SFS was done on Si(100) wafers with a 1 kHz amplified Ti:Sapphire laser system from Coherent Inc. The laser output is used to both generate a 7 cm^{-1} narrowband visible beam with an etalon and to pump a commercial OPA (OPerA Solo from Coherent Inc.) for tunable broadband IR light generation. Spectra are collected by a spectrograph equipped with an iCCD camera from Andor with an 800 fs time-delayed etalon pulse to suppress the nonresonant background of the silicon substrate.

The intensity in SF spectra can then be described as $I \sim |\chi_{resonant}|^2$ where $|\chi_{resonant}| = \sum_i \frac{A_i}{\omega_i - \omega + \imath\Gamma_i}$ where A_i, ω_i and Γ_i are the amplitude, the frequency and the width of the i-th resonance. In the analysis of the methyl group orientation, we use the Lorentzian peak-height $(A/\Gamma)^2$

5.2 Results and discussion

5.2.1 Surface quantification

Our standard surface preparation leads to highly hydrophobic propyl-terminated surfaces with a static water contact angle of 90° and hydrophilic aminopropyl-terminated surfaces with a static water contact angle below 50° with approximately monolayer silane surface coverage. The oxide thickness as derived from the Si:SiO_2 peak ratio in XPS increases during silanisation, and the silane film thickness varies between 0.4 and 1 nm as derived from the C1s signals (table 5.1 below [78]). By using anhydrous solvents and low concentrations, we avoid the well-known polymerization reaction and multilayer formation of APTMS [113]. APTMS appears to be less reactive than TMPS in the silanisation reaction, as the APTMS layer has a lower coverage than the TMPS layer and around 10% of the C1s signal can be attributed to unreacted methoxy groups at a binding energy of 286.8 eV. This is confirmed by N1s spectra, which are composed of an NH_3^+ peak at 401.9 eV binding energy [114], and a peak at 400.2 eV typically assigned to NH_2. Protonation is thought to occur by hydrogen bonding to surface silanol groups [115] and the 4:3 ratio of NH_3^+ to NH_2 shows that about half of the APTMS chains are lying flat on the surface. A highly tilted APTMS chain is supported by SFG (see below).

Substrate	Oxide [nm]	Silane [nm]
Clean	1.2	—
TMPS	1.6	0.4
APTMS	2.3	1

Table 5.1: *Quantification of adsorbate layer thickness on silicon-coated QCM sensors by XPS [78]*

5.2.2 Protein quantification

We study insulin adsorption at acidic pH of 2.7 for two reasons. Firstly, acidic environments and high temperatures are relevant for pharmaceutical production of insulin, but also lead most easily to protein unfolding and fibril formation. Secondly, insulin exists predominantly as monomers at acidic pH and the monomer has been shown a number of times to be the only species that adsorbs to surfaces and subsequently denatures to form amyloid fibrils [9] [100]. Insulin self-associates in solution to form dimers and hexamers, depending on presence of zinc ions, pH, concentration, ionic strength and type of buffer [100] [92] [116]. Low concentration, low pH and absence of Zn^{2+} ions generally favour the monomer. At our relatively low concentration of insulin (< 0.5 mg/ml) and neutral pH, Jorgensen *et al.* showed that human insulin with zinc consists of 90% monomers and 10% dimers [100]; with decreasing pH, this ratio will shift even further towards the monomer. We used a citrate-phosphate buffer with a low ionic strength of 50 mM, corresponding to a Debye length of 1.4 nm. This long Debye length means comparatively strong repulsive interactions between insulin molecules, which carry a positive charge of z=4.5 at pH 2.7 [108]. This disfavors rapid multilayer growth and allows us to look at conformational changes in the insulin

monolayer. We employ three experimental methods to determine the amount of adsorbed protein. QCM-D measures the mass of adsorbed insulin plus its hydration layer, FTIR-ATR absorbance measures the insulin surface density, while XPS detects the thickness of a dry insulin layer.

Using the approach presented in details in section 3.6[1], we derive the thicknesses of the adsorbed protein layers from XPS, ATR FT-IR and QCM-D spectra as presented in table 5.2

Silane / pH	Oxide thickness [nm]	C1s thickness (before HI adsorption)	Protein thickness [nm] (from XPS N1s peak)	Protein thickness [nm] (from QCMD)	Protein thickness [nm] (from IR ATR)
APTMS 2.7	1.60	0.40	0.07	0.07	0.42
TMPS 2.7	2.30	1.00	0.50	0.78	0.66

Table 5.2: *Oxide, silane and protein layer thickness for the two QCM sensors*

The summary results are shown in Figure 5.2. It can be seen that the hydrophobic surface has a coverage that is close to the theoretical insulin monolayer coverage of 1.1 mg/m^2 based on a rectangularly close-packed layer as derived by Mollmann *et al.* [103]. While there is a systematic difference between FTIR on one hand and XPS and QCM-D on the other hand for the hydrophilic surface, which stems from the nature of the experiment and will be discussed below, all three methods confirm that no more than a monolayer of insulin adsorbs at room temperature and pH 2.7.

[1]More details are discussed also in appendix C

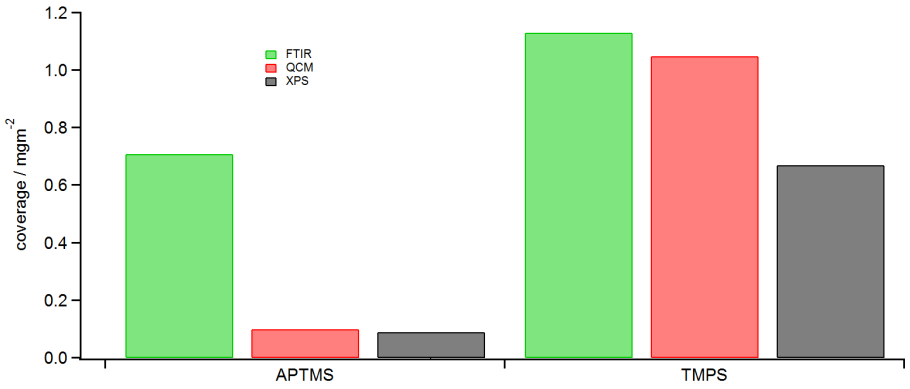

Figure 5.2: *Comparison of three experimental methods (XPS, QCM-D, FTIR) to determine insulin coverage on APTMS and TMPS functionalized surfaces at pH 2.7.*

Figure 5.3 shows in detail insulin adsorption as measured by QCM-D through resonance frequency and dissipation changes. Even though insulin has been classified as a "soft" protein, its globular, which is confirmed by the relatively small change in dissipation, shape and small size are unlikely to change the stiffness of the interface much. After equilibrating the quartz crystals in ultrapure water, buffer solution is introduced and the crystal is allowed to equilibrate for about half an hour. It can be seen that equilibration with the buffer is comparatively fast for the hydrophilic APTMS surface, while the hydrophobic TMPS layer shows a strong baseline shift. Despite degassing all solutions, this is very likely due to the initial formation of microscopic air bubbles at the hydrophobic interface, which reduces the near-surface viscosity [117]. Since insulin adsorption turns the TMPS interfaces hydrophilic, there is no further baseline drift. After insulin adsorption for 30 minutes, we switch back to flowing pure buffer and finally rinse with ultrapure water. The frequency change is significantly larger for the hydrophobic surface, while it is near the detection limit for the hydrophilic surface. We

fit the frequency and dissipation traces using the Voigt model, which takes into account the viscoelastic properties of the protein hydration layer [72], and determine adsorbed layer masses, after rinsing with milli-Q water, of 1.1 mg/m^2 for TMPS and only around 0.1 mg/m^2 for APTMS.

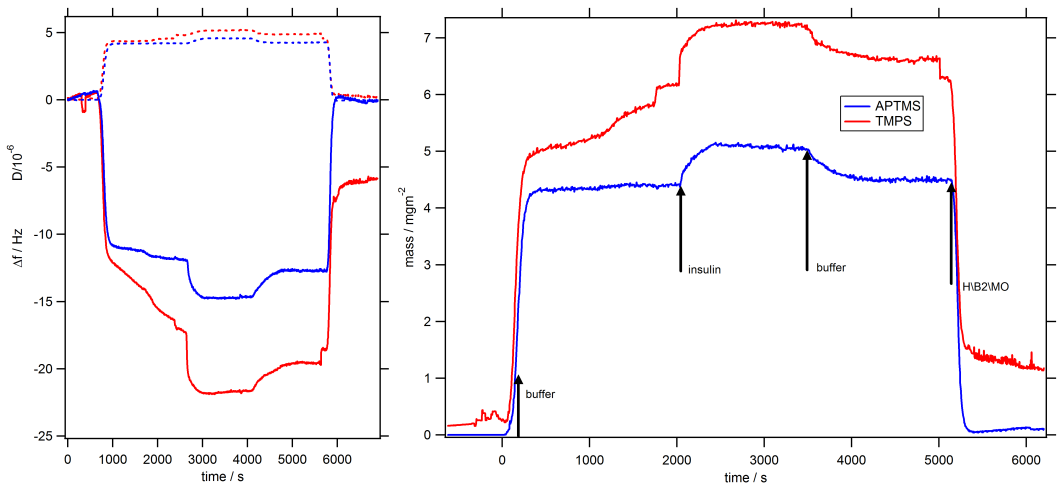

Figure 5.3: *Left: QCM frequency (solid lines) and dissipation (dotted lines) changes upon insulin adsorption on hydrophobic TMPS and hydrophilic APTMS in citrate-phosphate buffer at pH 2.7. Changes between solutions are indicated. Right: Result of the Voigt fit.*

The crystals used in the QCM-D experiments, are then dried under nitrogen and immediately used for XPS measurements. The corresponding XPS spectra of the N1s region are shown in Fig. 5.4, where the insulin-related peak has a binding energy of around 400.7 eV on the hydrophobic TMPS, while it is lower at 400.3 eV on the hydrophilic APTMS. Both binding energies are within the range typical for nitrogen in a protein amide bond [118]. APTMS also shows C1s binding energies that are around 0.4 eV lower than on TMPS and O1s binding energies that are almost 1 eV higher.

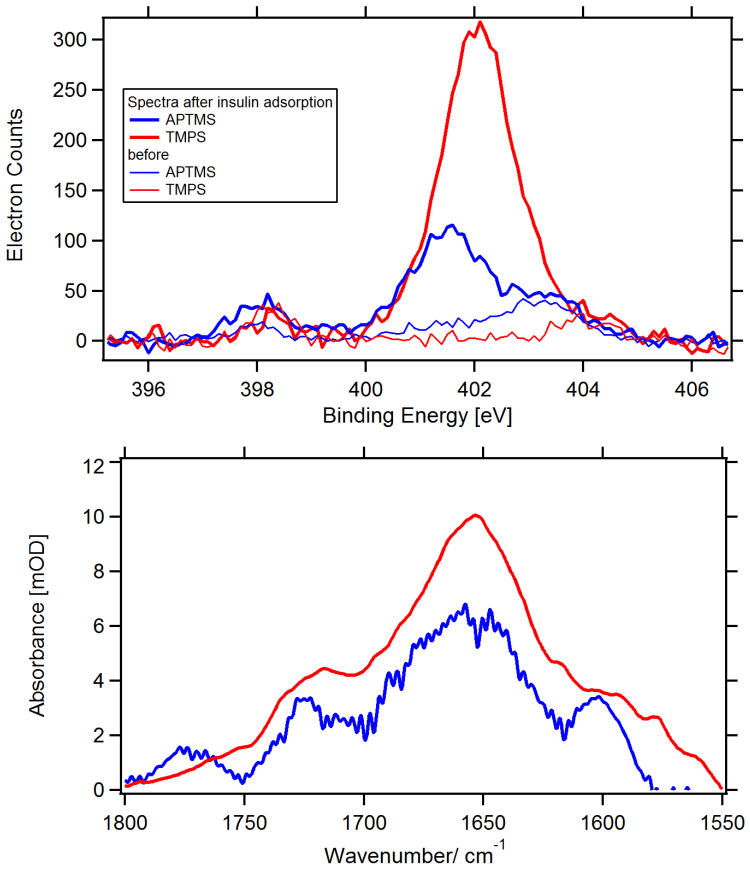

Figure 5.4: *Top: XPS N1s spectra after insulin adsorption on APTMS and TMPS at pH 2.7. Bottom: ATR FTIR spectra from dried insulin films for quantification*

The binding energy difference between the two surfaces could be due to a different orientation or tertiary structure of insulin on the two surfaces, but could also simply be due to different degree of residual hydration of the protein layer. Using the simplified model for protein film thickness estimation developed by Ray and Shard [79], we derive an insulin film thickness, from the N1s peak, of less than 1 nm, with comparable values gained from Si 2p attenuation. Using a density of 1.324 mg/m^3 for insulin [119], a film thickness of 1 nm would then correspond to an adsorbed (dry) mass of 1.324 mg/m^2. The

QCM-D-derived mass is expected to be higher than the XPS-derived one, which is explained by the hydration layer around the protein. For TMPS at pH 2.7, the volume fraction of insulin is 64%. This compares well to volume fractions around 60% found for fibrinogen, IgG and BSA [79]. The lower insulin coverage on APTMS is presumably due to repulsion between insulin and the surface, which are both positively charged at pH 2.7 [114]. Quantification of the layer thickness using ATR-FTIR requires two intermediate steps. Firstly, we used a molar extinction coefficient of 5530 $M^{-1}cm^{-1}$ at 277.5 nm to determine the concentration of the insulin solution [120] and then recorded FTIR transmission spectra of the amide I band in a transmission cell with a 100 μm spacer. For a concentration of 75 μM, we obtain a peak absorbance of 0.025 at 1649.4 cm^{-1}, which translates into a peak extinction coefficient of ε=33,333 $M^{-1}cm^{-1}$. A peak absorbance of 0.01 then roughly corresponds to a coverage of 0.7 mg/m^2. To better account for possible changes in peak shape, we however use the integrated extinction coefficient ψ (appendix C), which is 2.07×10^9 cm mol^{-1}. We then use this value to calculate adsorbed densities from the peaks in the ATR spectra shown in fig. 5.7. Their peak absorbance is shifted up by 5 cm^{-1} to 1654.5 cm^{-1}, due to the reduced hydration shell on the protein. Quantitative analysis of FTIR-ATR spectra is however simpler for dry insulin layers. As fig. 5.2 shows, the FTIR-derived coverage is much higher for APTMS than the coverage determined by QCM-D or XPS. The main reason for the discrepancy is that ATR crystals were rinsed directly with water rather than buffer after immersion in insulin solution. The sudden change back to neutral pH means less repulsion between HI and APTMS and therefore a higher residual coverage. Our insulin coverage compares well with literature values for adsorption at physiological pH 7.4 at hydrophobic surfaces such as teflon [103] [104], polystyrene [98], and

methyl terminated alkanes [9] [10], which are all in the 1-3 mg/m^2 range. There are no comparable quantitative data for these rigid surfaces at acidic pH. Experiments at polyacrylic brushes [121] and charged lipid layers [101] at pH of 2 or lower detected much higher adsorbed masses between 4 and 20 mg/m^2.

5.2.3 SAM and protein side chain orientation

Sum-frequency spectra of siloxane films grown on Si(100) wafers and the corresponding FTIR spectra on a silicon ATR are shown in fig.5.5. The TMPS film possesses a well-defined sum frequency spectrum, which consists entirely of three CH_3-related vibrational modes - the symmetric stretch at 2872 cm^{-1}, the asymmetric stretch at 2959 cm^{-1} and the Fermi resonance between the symmetric stretch and the overtone of the CH_3 bend at 2929 cm^{-1}. These frequencies are close to those found for C_{18} siloxanes on fused silica or glass [122] [123]. The approximate center of inversion between the two CH_2 groups in an all-trans TMPS chain prevents the generation of a sum frequency signal from these groups [124] [61] [62]. The lack of any CH_2-related signal in SFG thus shows the formation of a well-ordered film. The APTMS films produce a similar looking SFG spectrum, which we propose is due to CH_3 vibrations from the unreacted methoxy groups detected by XPS, similar to the assignment by Kim et al. [113] rather than the alternative assignment to symmetric stretches of the individual CH_2 groups by Tong *et al.* [125]. The phase of the sum frequency light emitted by the CH_3 resonances with respect to the silicon nonresonant background is the same for TMPS and APTMS, which implies that both methyl groups have the same general orientation, that is, pointing up from the silicon surface. An aminopropylsilane with an amino group hydrogen bonded to a surface silanol as deduced from XPS is

highly compatible with residual methoxy groups pointing upwards.

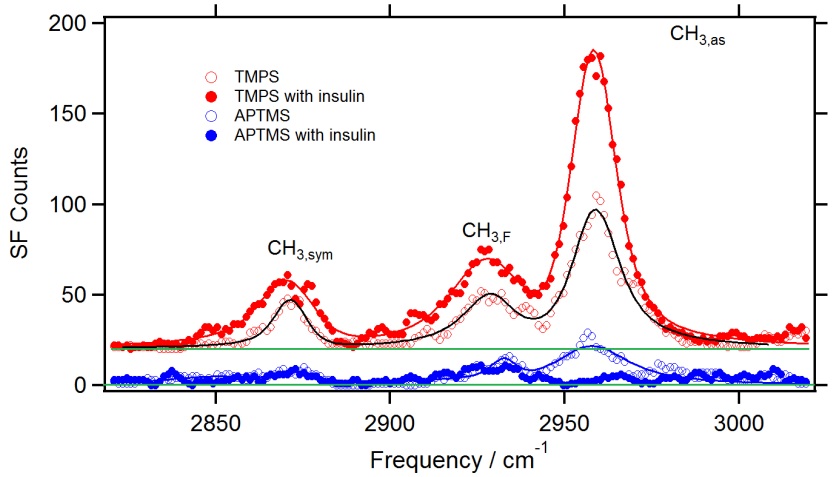

Figure 5.5: *Sum frequency spectra (ppp polarisation combination) of dried hydrophobic and hydrophilic silane layers grown on Si(100) before (open symbols) and after (closed symbols) insulin adsorption at pH 2.7.*

Similar to analysis described by Bourguignon's group [124] [61], and by Wang [126], we can deduce the angle of the CH_3 symmetry axis with respect to the surface normal from the intensity ratio between symmetric and asymmetric peaks [61]

$$R = \frac{I_F + I_{sym}}{I_{asym}} \tag{5.1}$$

To relate this ratio to a methyl tilt angle, we calculate hyperpolarisability values for methyl and methylene groups using the bond additivity model. We account for reduction of the symmetric SFG signal by intensity splitting between the symmetric CH_3 stretch and the Fermi resonance and use the recorded FTIR spectra to deduce the degree of intensity sharing [124]. Details of the calculation can be found in appendix A. We obtain R= 1.05 from the experimental TMPS/Si(100) spectrum, which corresponds to tilt angle of 33° for the methyl group, respectively an angle of $\sim 8°$ for the propyl chain.

This is slightly lower than the backbone tilt angle of around 15° deduced by SFG for a C_{18} silane chain on glass [127]. This average tilt angle probably arises from the strain in the surface generated by fitting Si-O-Si polymer chains onto the Si(100) lattice. Naik *et al.* [128] deduced the presence of so-called *snow moguls*, where not all siloxanes are anchored to the substrate to allow more space for the alkane chain. For the hydrophilic APTMS/Si(100), we obtain a very similar tilt angle of 32° for the CH_3 of the methoxy group. Insulin adsorption leads to clear changes in the sum frequency spectrum of both layers. On the hydrophobic surface, the asymmetric CH_3 peak grows in relationship to the two symmetric peaks, such that the symmetric to asymmetric ratio is reduced from 1.05 to 0.28. The width of the two symmetric peaks increases significantly, while that of the asymmetric peak remains the same, suggesting that this spectrum could have small contributions from nearby CH_2 groups, although they are not clear enough to be fit reliably. There are two possible origins of the change. It can be either due to a hydrophobic interaction between insulin and TMPS and a thereby increased tilt angle of the surface methyl groups or it can be due to a net orientation of insulin side chains on the hydrophobic surface, as observed for amino acids [129] and polypeptides and proteins [126] [130]. In the first case, the decrease in the symmetric/asymmetric peak ratio to 0.28 would correspond to a 22° increase in the tilt angle of the methyl group. This is much larger than any other hydrophobic interaction measured by SFG. For comparison, the CH_3 tilt angle of octadecanethiol on Au(111) increased by 3.3°, when the solvent was changed from distilled water to a hydrophobic bacterium [131]. A similar reorientation of methyl groups has also been observed for protein adsorption on deuterated polymers [132]. We therefore consider a pure hydrophobic effect unlikely to be the sole cause of the observed change. Instead,

it appears that insulin adsorbs in an ordered fashion on the hydrophobic surface with the main effect of increasing the intensity of the asymmetric CH_3 stretch over the symmetric one.

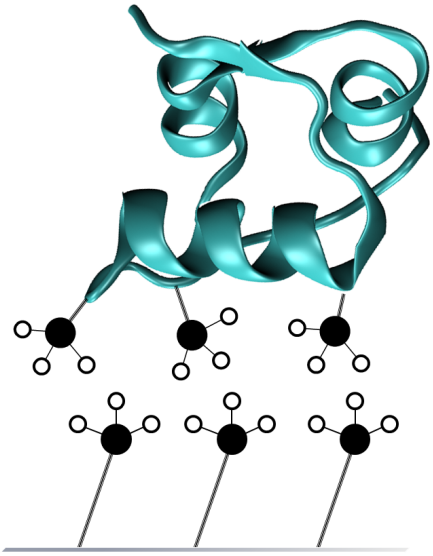

Figure 5.6: *Sketch of insulin adsorption on a TMPS functionalised silicon surface.*

An insulin monomer has a footprint of about 2×3 nm when it adsorbs with its hydrophobic dimer interface to the surface (the most likely orientation). In this orientation, it covers roughly 30 unit cells of an unreconstructed Si(100) surface and its 28 methyl groups are oriented in such a way to make a net contribution of 10 methyl groups to a sum frequency spectrum (appendix C). The TMPS functionalized substrate will at most have one alkane chain per unit cell and we would therefore expect around 30 methyl groups from the surface layer within the insulin footprint. The observed spectral changes are therefore highly compatible with an oriented adsorption of insulin monomers, in particular as the net CH_2 contribution can be low for a wide range of tilt angles (appendix C). Any further analysis is however impossible, since the

total spectrum contains mixed contributions from both insulin and TMPS, which make analysis very difficult. On APTMS, on the other hand, insulin adsorption leads to an almost complete disappearance of the asymmetric CH_3 stretch, so either the residual methoxy groups tilt upwards or insulin adsorbs in such a way as to cancel this signal. The low signal also means there is no net orientation of side chains.

5.2.4 Protein structure

The dried spectra in fig. 5.4 already showed that surface termination and pH only influence the peak position and width of the amide I spectrum by a small amount, which is unexpected given published results. Circular dichroism spectroscopy of insulin adsorbed to teflon and polystyrene beads at pH 7.4 has shown a decrease in α-helical content and an increase in β-sheet or random coil [103] [100] [98], while an FTIR-ATR study of insulin layers on a phenyldimethyl-functionalised silicon crystal interpreted a peak at 1708 cm^{-1} as the appearance of intermolecular β-sheets [10]. In fig. 5.7, we show wet insulin spectra for APTMS and TMPS at pH 2.7 recorded after 30 min, 90 min and 180 min adsorption. Over the course of three hours at room temperature, the shape of the spectrum does not change and, as the overlaid solution spectra show, the spectral shape is extremely similar to native insulin in solution. Wet layers have a roughly twice as high coverage as dry layers which shows that under wet conditions, insulin probably consists of a chemisorbed layer and a physisorbed layer. A small shift towards higher frequencies can be seen for the TMPS spectra, which could indicate a small increase in random coil content, while APTMS has a slightly broader spectrum which could also indicate a small amount of disorder. Overall though, under these conditions, the secondary structure of insulin is hardly affected

by adsorption and TMPS and APTMS functionalized surfaces are as protein-friendly as polyelectrolyte brushes [109].

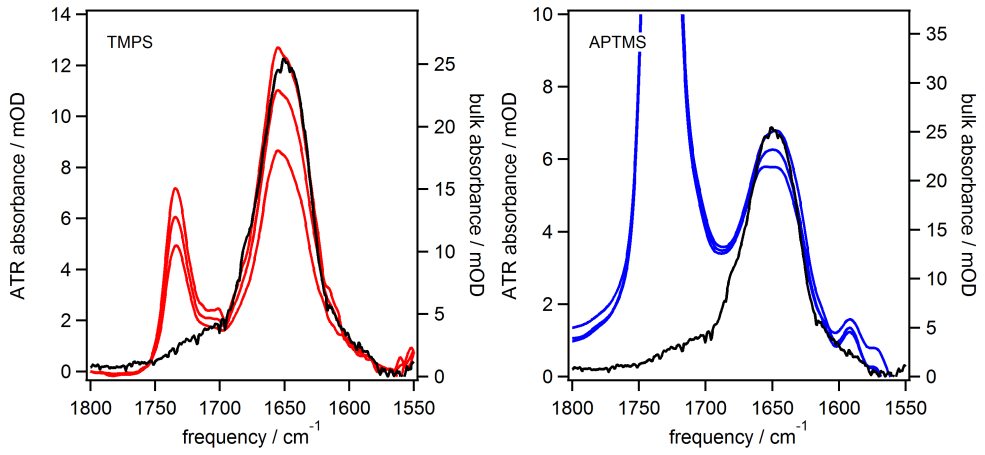

Figure 5.7: *Comparison of wet insulin layer spectra at pH 2.7 for TMPS (red) and APTMS (blue) with bulk insulin spectra (black, right axis). Spectra were recorded after 30, 90 and 180 min in insulin+buffer and the pure buffer spectra were subtracted.*

The narrow peak at 1734 cm^{-1} is due to the C=O stretch in the carboxylate groups of citric acid. At the low pH used, we would normally expect this group to be deprotonated, and therefore detect asymmetric O-C-O stretching peaks around 1572 and 1597 cm^{-1} in a 1:2 ratio [133]. Citrate is a hydrogen bond acceptor, which could bind to several hydrogen bond donors in insulin, specifically -NH_2 or -OH groups. The much more intense C=O peak on the APTMS functionalized surface is due to citrate binding to the surface amine groups, which are also hydrogen donors. We next investigate the influence of elevated temperature on the coverage and structure of adsorbed insulin by heating the solution and crystal to 65°C. As we could not heat the ATR assembly in-situ or keep layers wet for longer than 3 hours without significant depletion of the insulin concentration, we prepared ATR crystals ex-situ and

recorded spectra of dry insulin layers.

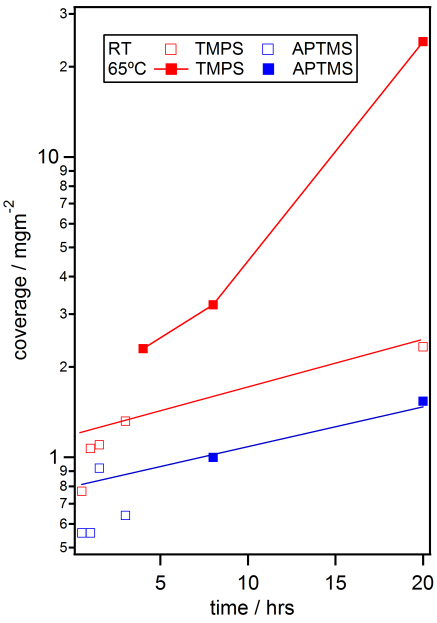

Figure 5.8: *Insulin adsorption at room and elevated temperature at pH 2.7*

Rinsing the crystal does not always remove all traces of citrate ions, so we ignore the changes in the citrate peak found at 1744 cm^{-1} for dry layers. Having established for the wet layers shown in fig. 5.7 that the structure of adsorbed insulin is near-native, we can use the spectra in fig. 5.4 as a reference for native, adsorbed insulin without its hydration shell. The maximum adsorption time was limited to 20 hours by gradual H/D exchange with water vapor in the air. Fig.5.8 shows that extending adsorption at room temperature from 1 hour to 20 hours changes the amount of adsorbed insulin by a factor 2 (corresponding to a mass change from 1.1 mgm^{-2} to 2.3 mgm^{-2}) and slightly shifts the peak to higher frequencies. This hydrophobic surface is therefore "insulin-friendly" for an extended time under acidic conditions at room temperature. Heating the crystal and solution to 65°C leads to a distinct β-sheet spectral signature after 4 hours at a coverage of 2.3 mgm^{-2}.

This is significantly higher than the coverage of 1.4 mgm^{-2} achieved after 3 hour adsorption at room temperature and shows that layer growth goes hand in hand with β-sheet formation. As the scaled spectra in Fig.5.9 show, the overall shape of the amide I band only changes marginally during multilayer growth.

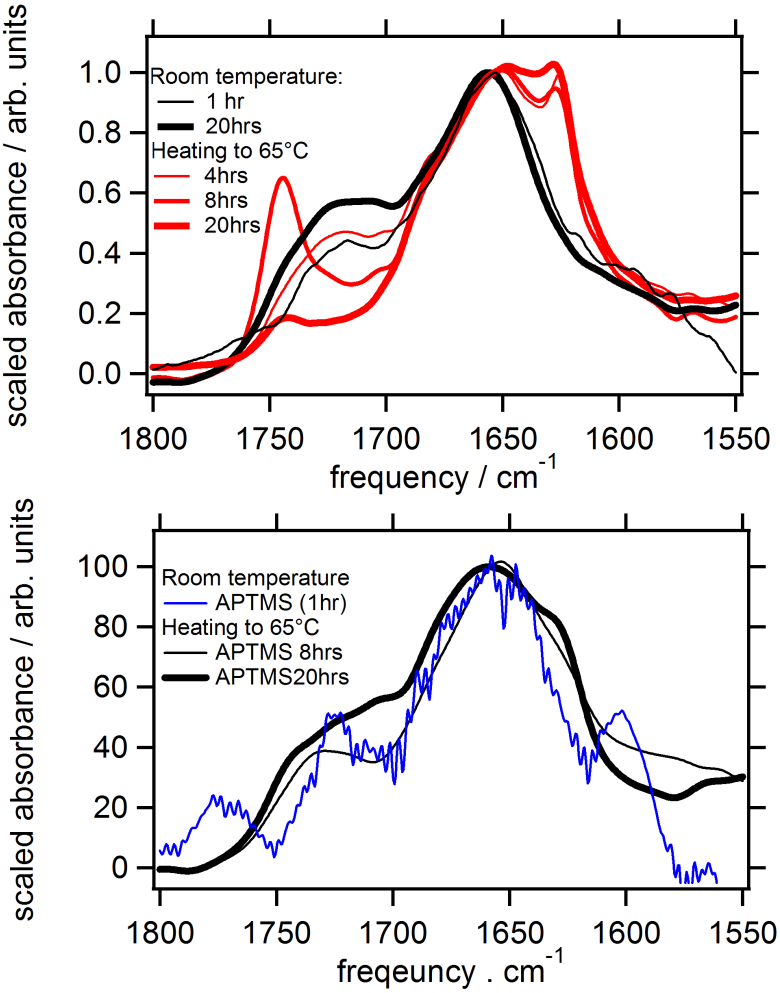

Figure 5.9: *Scaled ATR spectra to show the change in amide I band shape for TMPS (top) and APTMS (bottom). Due to the noise on the cold APTMS spectrum, the TMPS spectrum is repeated in the figure. IR spectra of native insulin and insulin β-sheets are shown in appendix C- fig. C.2*

The APTMS terminated surface stabilizes the adsorbed insulin conformation for at least 8 hours at heat and β-sheets only appear after heating for 20 hours. This supports the literature finding that insulin bound to APTMS-covered silica beads is still physiologically active [110].

5.2.5 Discussion

Surprisingly, the data presented in this work show that insulin monomers maintain their native structure when adsorbed on both hydrophobic and hydrophilic surfaces at low pH. This is contrary to detailed structural studies at physiological pH, which find evidence for restructuring [100] [10]. At both low and neutral pH, the orientation of adsorbed monomers is surmised to be with the hydrophobic surface domain of the monomer attached to the hydrophobic surface [9]. This causes an entropy gain from water repulsion and is similar to the hydrophobic interaction to form the dimer, although the latter is aided by hydrogen-bond formation. In our case, the increased hydrophilicity of the surface after adsorption, as seen in QCM-D and by eye, indirectly confirms that insulin adsorbs with the hydrophobic side down. The SF spectra of TMPS with adsorbed insulin confirm that there is a net order in the side chains, which would be expected form an oriented adsorption. The hydrophobic attraction between insulin and surface is obviously strong enough to overcome the repulsion between the highly charged insulin monomers in a low ionic strength solution. This was proven in a study of the adsorption of various modified insulin variants on polystyrene and Teflon beads [113] [122]. Acylation for example increased insulin's negative net charge and its hydrophobicity. The adsorbed amount on polystyrene was higher compared to unmodified insulin, therefore hydrophobic interactions were more important than electrostatic repulsion between monomers.

Similarly, by attaching a fluorescent label to the N-terminal of the B-chain, Mollmann *et al.* [104] concluded that this end was not attached to the surface. Instead, an accompanying shift in tryptophan fluorescence indicated that it is likely the C-terminus of the B-chain (which forms part of the dimer interface) that is involved in adsorption. Recent molecular dynamics simulations by Nikolic *et al.* [134] on small peptide segments show that beta-sheet formation at a hydrophobic surface can be partly explained by the hydrophobic effect and partly by reduced conformational entropy of the polypeptide chain. They found a strongly reduced number of water molecules bound to side chains and around the peptide backbone. The dehydration increases the entropy and also reduces the conformational flexibility of the polypeptide and encourages formation of new peptide-peptide hydrogen bonds. While an almost complete dehydration of adsorbed insulin leads to a 5 cm^{-1} peak blue shift, the comparison between wet adsorbed insulin and insulin solution spectra in fig. 5.7 shows only a minimal difference. On the hydrophilic surface, a possible binding mode which overcomes the overall repulsion between positively charged insulin and APTMS, is a localized attraction between positively charged -NH_3^+ and negatively side chains. The C termini of the A and B chains (Asparagine and Threonine respectively) of insulin could electrostatically attach to/react with the -NH_3^+ terminated surface. Such a localised binding could explain why the APTMS chains reorient (as shown indirectly by the SF spectrum of their residual methoxy groups) but why there is no contribution from insulin side chains evident in the spectra. A possible reason for why we observe a native structure for adsorbed insulin is that low pH is known to stabilise the native state of the monomer in solution [135], despite increasing overall fibrillation rates by increasing the monomer fraction. Certainly a contributing factor (although this is independent of pH) is the

fact that insulin adsorption renders a hydrophobic surface hydrophilic and similarly charged to insulin in solution. This will reduce further adsorption - in our case, the coverage after 20 hours is only twice as high as after 1 hour. Denaturation occurs only if we heat surfaces and solutions at a temperature of 65°C. Jorgensen *et al.* [100] proposed that the fibrillation of insulin occurs at the surface, but showed that the structure is different from fibrils formed in solution. The dependence of fibril morphology on surface termination has been shown repeatedly in the literature [136]. We equally find that adsorbed and solution spectra (fig. 3.4) differ strongly in their beta-sheet content. In addition, we find that the change from native to restructured adsorbed insulin occurs within the first insulin monolayer (4 hour adsorption at heat) and that only then the multilayer begins to grow, while maintaining a similarly shaped infrared spectrum. The intermediate on the path to fibril formation can thus be identified as the insulin monolayer, which appears to provide the template for further growth. In conclusion, our measurements demonstrate that adsorption of insulin at acidic pH on solid, hydrophobic surfaces can lead to a stable, near-native structure of the protein. Hydrophobic surfaces are therefore not per se detrimental to the secondary structure of insulin, as we find that significant thermal activation is needed before restructuring occurs. Observation of insulin secondary structure at heat showed in addition that the insulin monolayer should be considered the intermediate species on the way to fibrillation.

Chapter 6

Insulin fibril formation on solid surfaces

The experiments presented in this chapter have been carried out at the University of Liverpool, in collaboration with the group of Prof. P. Weightman - Dept. of Physics

In chapter 5 we have shown by means of ATR FT-IR that on hydrophobic surfaces, insulin adsorbs and aggregates, forming a thick layer of β-like structures at high temperature. FT-IR can distinguish native β-sheet proteins from amyloid fibrils, as the former have a characteristic amide I peak in the region of 1630 cm^{-1} to 1646 cm^{-1}, while the latter possess peaks in the range 1611 to 1630 cm-1 [137]. The same study however showed that FT-IR cannot distinguish between amyloids and nonfibrillar aggregate structures. Currently, the formation of protein fibrils is studied using scanning probe microscopies. These methods do not provide spectroscopic information, so even if the morphology of the fibrils can be studied, their composition still remains unknown. Recently, two studies have combined vibrational spec-

troscopy with scanning probe microscopy to provide more structural detail of amyloid fibrils. Amenabar *et al.* [30] used tip-enhanced IR spectroscopy to detect the composition of insulin fibrils and concluded that they consist of an outer shell with dominant α-helical structure. Only vibrations perpendicular to the substrate were found to couple strongly to the near field of the tip, so that this spectroscopy has the potential to not just determine structure but also orientation. Kurouski *et al.* [106] used tip-enhanced Raman scattering and concluded that around 30% of their insulin fibril surface consisted of β-sheet structures, based on amide I band shape and the presence of marker bands for some specific amino acids. While these new approaches are effective, they are experimentally highly challenging and not largely available. As an alternative to vibrations of the peptide backbone, optical transitions in the 180 nm – 240 nm UV range can be used for structural analysis. Circular Dichroism (CD) is commonly used to evaluate the secondary structure content of proteins, and has often been employed to study the transition from α-helices to β-sheets [138]. CD however presents two drawbacks for studying thin protein films at surfaces: it is not sensitive enough in transmission and it cannot work in reflection, as that reverses the direction of polarisation rotation. Alternatively, linear dichroism, defined as the difference between absorption of light polarized parallel and polarized perpendicular to an orientation axis, has been used in transmission to study proteins, which can be aligned in a flow cell or which self-align in a membrane [138] [139], but has not been applied to protein adsorption at solid interfaces.

Here we present a novel approach for studying the formation of insulin fibrils on solid surfaces using Reflection Anisotropy Spectroscopy. RAS is a form of linear dichroism detected in reflection. The experiments are easy to perform, and can be performed in-situ through a liquid layer. We compare

RAS spectra with ATR FT-IR spectra, and propose a model for the formation of fibrils at surfaces.

6.1 Results and discussion

6.1.1 Sample Preparation

Si wafers [Si(111)-vicinal][1] were cleaned using Piranha solution (70% H_2SO_4, 30% H_2O_2 solution (35% in water)) for 10 minutes at room temperature in a sonication bath. After cleaning, all samples were thoroughly rinsed with copious amounts of ultra-pure water (18.2 MΩ) and dried under Argon flow. All glassware was cleaned in Piranha solution at 70°C for 1 hour, rinsed several times with ultra-pure water and ethanol, dried under a stream of dry air and finally rinsed with toluene. Trimethoxypropylsilane (TMPS) and 3-aminopropyltrimethoxysilane (APTMS) solutions were prepared in toluene at a concentration of 5 mM and 0.5 mM respectively. Low concentration of APTMS in solution is used to avoid polymerisation and/or multi-layer formation. Silicon wafers are left in solution for 24 hours at room temperature, sonicated in toluene to remove potential multi-layers and rinsed with copious amounts of ultrapure water. Finally substrates are dried under an argon flow. Insulin solutions are prepared in Citrate-Phosphate buffer at pD 2.7, and at a concentration of 0.44 mg/ml. The buffer is prepared as follows: 44.6ml of citric acid solution (0.1 M in D_2O) and 5.4 ml of dibasic sodium phosphate solution (0.2 M in D_2O) are mixed. The final volume is adjusted up to the final volume of 100 ml and the final pD is carefully adjusted by

[1] A Si(111)-vicinal wafer is used since it is a stepped surface. These steps give rise to a RAS signal which helps to identify the orientation of the surface while doing the experiment.

addition of small amounts of either concentrated citric acid or dibasic sodium phosphate, by using a sensitive pH meter. Human insulin with 0.3% Zn by weight was kindly provided by Sanofi-Aventis Deutschland GmbH.

6.1.2 RAS spectra

RAS spectra have been measured using a RA spectrometer of the type described in the previous section (3.20). Measured spectra are shown in fig. 6.1, and are compared with corresponding ATR FT-IR spectra.

Functionalisation with APTMS and TMPS does not induce any visible changes in the RA spectra. The silanes therefore do not introduce any new anisotropy or modify the stepped surface significantly.

APTMS functionalised substrates are known to be insulin-friendly for extended periods (chapter 5). The FTIR spectra in fig. 6.1 show a main peak at 1654 cm^{-1}, characteristic for an α-helix. After 20 hours at 65°C heat, a small shoulder develops at about 1620 cm^{-1}, which shows that the layer is beginning to restructure into a β-sheet. We previously derived (see fig.5.8 and fig.fig. 5.9) that 8 and 20 nm adsorb on APTMS, which correspond to insulin coverages of $\sim$ 1 $mgcm^{-2}$ and 1.5 $mgcm^{-2}$, that is the insulin films are roughly a monolayer thick. Just like the FTIR spectra, the RA spectra do equally not show any significant changes. Small changes in the reflectance anisotropy are typical for thin adsorbed layers [84] [140].

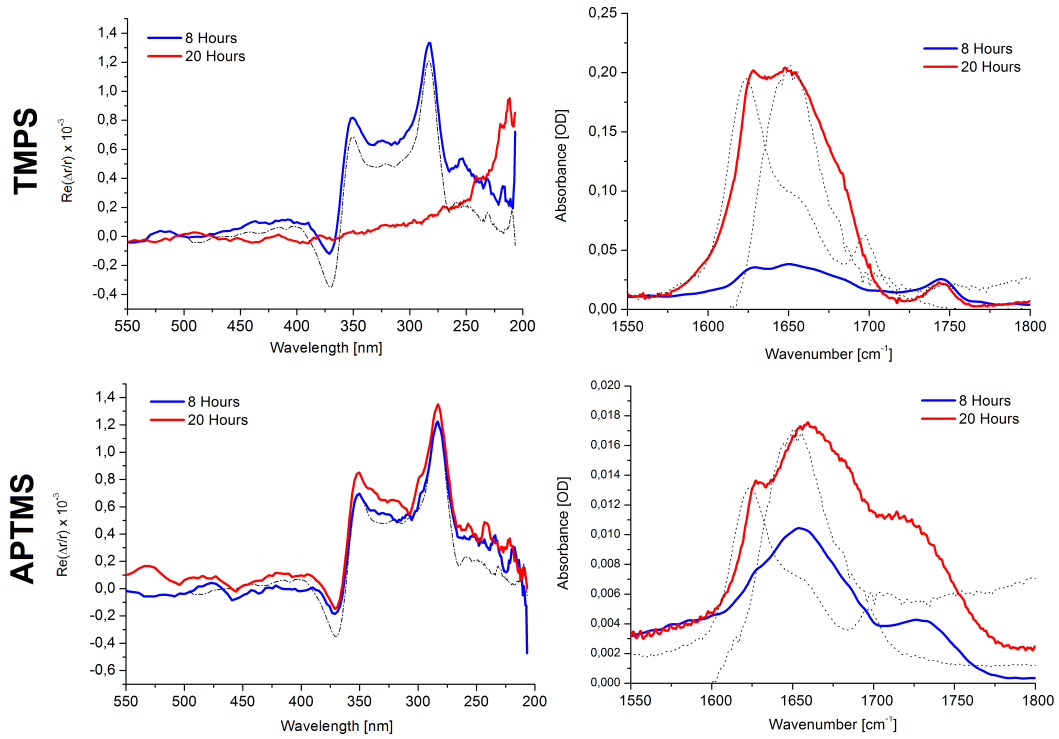

Figure 6.1: *Top: RAS (left) and ATR FT-IR (right) spectra of human insulin adsorbed on TMPS SAMs on silicon. Bottom: RAS (left) and ATR FT-IR (right) spectra of human insulin adsorbed on APTMS SAMs on silicon. Blue lines: HI adsorbed for 8h at 65° C. Red lines: HI adsorbed for 20h at 65° C. Line-Dot lines in RAS plot are the bare Silicon spectrum. Dotted lines in IR spectra are the spectra of native insulin and β-sheets in solution.*

The scenario is quite different for the hydrophobic TMPS functionalised surface after extended heating in insulin solution. After 8 hours, an amount equal to roughly 3 monolayers ($\sim$ 3.2 mgcm^2) of insulin is adsorbed at the surface, as deduced from the IR absorbance. This layer shows a clear β-sheet shoulder in the infrared spectrum, but does again not cause big changes in the RA spectrum, suggesting that the protein is randomly distributed across the surface. After 20 hours, a multi-layer of insulin $\sim$ 24 mgcm^{-2} is present

at the surface, as derived from the IR spectrum absorbance. Now a dramatic change occurs in the RA spectrum, which no longer shows the characteristic peaks of the Si surface at 350 nm (3.5 eV) and 275 nm (4 eV), but instead displays a new feature at ∼220nm.

The disappearance of the silicon peaks can have a number of reasons. Amyloid fibrils are known to be birefringent under visible light [141], even though no detailed spectroscopy of this phenomenon has been carried out to the best of our knowledge.

If one considers a non uniform birefringent layer, incident polarised light will undergo a polarisation rotation according to the local thickness of the layer. Therefore, the initially polarised light, will hit the surface with a random polarisation state, and therefore the RA signal will vanish. A similar phenomenon has been observed in other studies [142] when thick layers of material are present at the surface.

Alternatively, the adsorbed layer affects the in-plane strain of the Si surface to such a degree as to make the surface appear optically isotropic.

The peak at 220 nm suggests the presence of optical anisotropy at the surface. In conjunction with the FT-IR spectrum, we proposes that is signifies the presence of insulin fibrils. Bouchard *et al.* [44] have shown that after heating insulin solutions at high temperatures, a CD peak at 220 nm becomes prominent. They also show electron micrograms of fibrillar structures. Still, CD spectra do not compare directly with RA ones, because the optical interaction that give rise to either CD or LD peaks are different. Where does the linear dichroism we measure originate from? Hicks *et al.* [139] have shown that a β-sheet presents two UV transitions at 221 nm and 219 nm (plus another additional weak one at 195 nm).

To support this idea, we use a method developed by the group of Prof.

Hirst at the University of Nottingham[2] to obtain LD and CD simulated spectra from first principles [143]. We thus get simulated LD spectra of an α-helix and a β-sheet. Simulated CD spectra do not exactly mimic the ones reported in the literature, indicating that these calculation are approximate.

The computed spectra are shown in fig.6.2

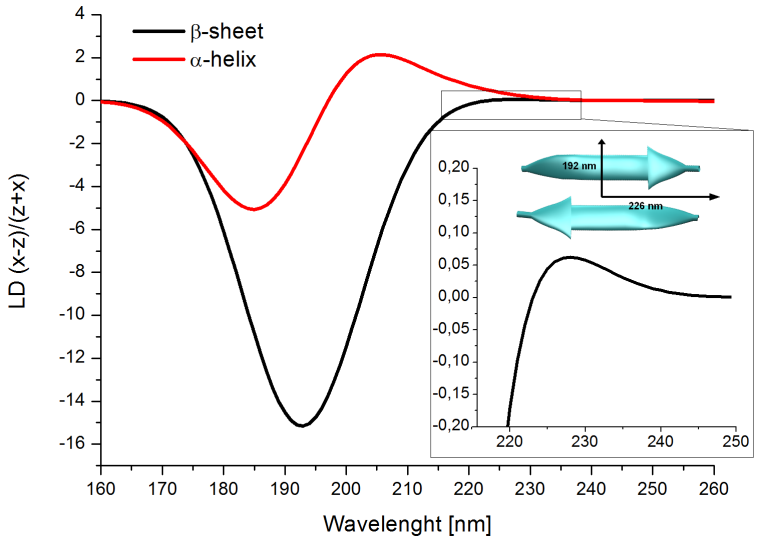

Figure 6.2: *Simulation of Linear Dichroism of an ideal α-helix and an ideal β-sheet.*

α-helices result in two peaks at 184 nm and 204 nm respectively whereas β-sheets present linear dichroism at higher wavelengths, at 192.5 nm and 226 nm. We can therefore assign the RAS peak at 220 nm to the presence of β-sheets throughout the fibrils. Since the RA signal is strictly related to in-plane anisotropy, we propose that insulin fibrils are formed in a ordered manner. Here we have shown that RAS can be used as a fast and user friendly experiment to test fibril formation at surfaces.

[2]http://comp.chem.nottingham.ac.uk/

6.1.3 Preparation of fibrils under harsh temperature conditions and sample aging

To further investigate the temporal stability of insulin fibrils, we also prepared insulin fibrils under high temperature and stirring conditions. Samples are prepared from insulin solutions at 1mg/ml concentration. The temperature is kept just below the boiling point, and the solutions are stirred at a low speed. We measured the freshly prepared samples and we let them age for 3 and 9 months, respectively, to check the stability of insulin fibrils. The experimental results are shown in fig.6.3.

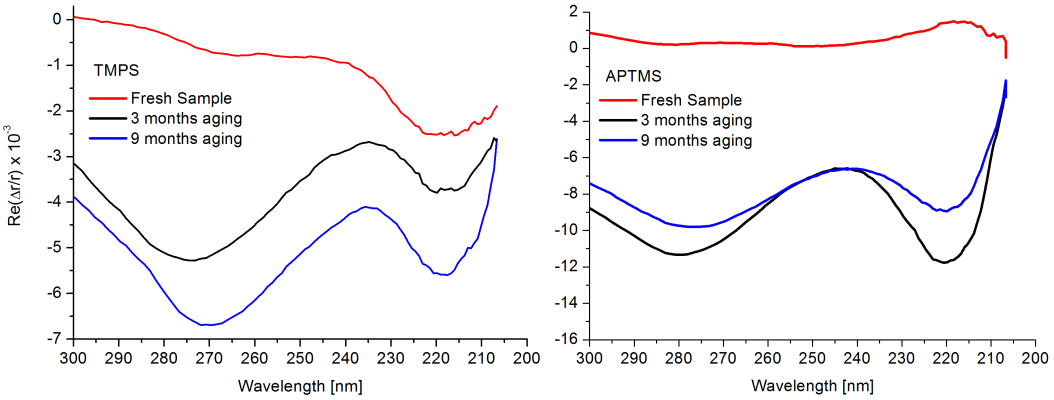

Figure 6.3: *RA spectra of insulin fibrils grown under high temperature conditions on hydrophobic (left) and hydrophilic (right) SAM. Red: fresh samples. Black: samples aged for 3 months. Blue: samples aged for 9 months.*

Even though we did not record FT-IR spectra from fresh samples, we would suspect extended fibril formation on both surfaces simply from the conditions used and the fact that β-sheets are already seen after 20 hours growth at a lower temperature. The RA spectra recorded in the null direction of the silicon surface now show a clear resonant peak at about 220 nm for both SAMs. The fresh samples show anisotropy in this region with opposite

sign, meaning the associated dipole moments are at 90 degrees to each other. As the samples age, a second, much stronger resonance develops at about 280 nm, which has a larger anisotropy for APTMS. While the first resonance is of the right wavelength for the peptide backbone, the second resonance occurs in the wavelength region corresponding to amino acid side chains. Since human insulin does not contain tryptophan and phenylalanine only has a weak absorbance in this range, the most likely source this peak is tyrosine [144]. In linear dichroism spectra, the side chains, create a much weaker anisotropy than the peptide backbone (ca. factor 10). A possible explanation for the much stronger signal in RAS could be that there is on average more order in the side chains than in the backbone when insulin fibrils grow on these surfaces. It is unlikely that the 220 nm resonance belongs to tyrosine, as the transition dipole moment then should be 90 degrees from the 280 nm resonance, ie when one is negative in RAS, the other should be positive.

We recorded FT-IR transmission spectra for the 8 month old TMPS and APTMS samples, as shown in fig. 6.4. Both spectra show a clear feature at 1650 cm^{-1}, indicating a large presence of random coil or alpha-helical structures.

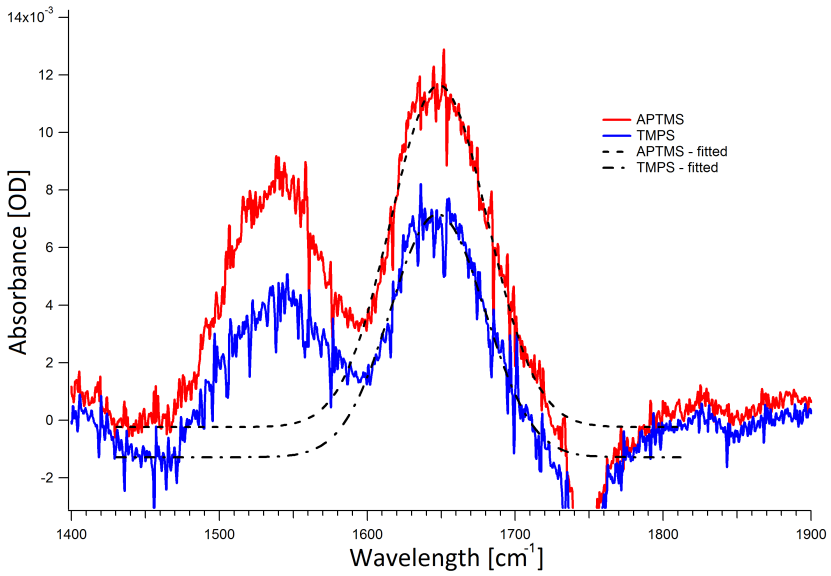

Figure 6.4: *FT-IR spectra of insulin fibrils grown on Si(111)-vicinal substrates after 8 months aging. Blue: hydrophobic. Red: hydrophilic*

These spectra in comparison to those shown for 20 hours (fig. 6.1) suggest significant reordering of the insulin fibrils, either during prolonged heating or during 8 months storage. The RA spectra however only show intensity changes for side chains and backbone and not the peak shift one might expect for a change in secondary structure. While these initial results are promising, clearly more work needs to be done to aid interpretation of the RA spectra. For example, a combination of RAS, FT-IR and AFM should allow a much better interpretation of the results.

Summary and future work

During this PhD project, the interaction of insulin with a variety of surfaces has been studied in a multi-step approach. We aimed at clearing some of the confusions that clutter the literature in this field.

For the first time, insulin amide I spectra at the air/water interface have been measured by SFG. Thanks to these results, we have shown a straight link between the fact that insulin tends to denature (upon shaking) when in contact with air, and the fact that monomers are the responsible species for insulin denaturation.

After we identified the culprit of insulin denaturation, we compared its interaction with hydrophobic versus hydrophilic solid surfaces. To do so, we tackled the problem with a surface science approach, characterising the surfaces in a systematic manner, with a manifold of complementary techniques. We adopted the same approach for studying the protein adsorption onto these surfaces. This revealed surprisingly that adsorption at low pH maintains an almost native structure of insulin on both surfaces at room temperature, and that denaturation occurs only after incubation at higher temperatures.

Finally we followed the formation of insulin fibrils using a novel approach based on Reflection Anisotropy Spectroscopy in combination with FT-IR spectroscopy.

While thesis has answered some questions pertaining the insulin unfold-

ing/aggregation problem, it has also generated some new discussions.

From the biochemical point of view, it would be interesting to study the behaviour of insulin at lipid membranes, because of their higher relevance in a biological context. Would we find the same segregation of the fibril-prone hormones at these interfaces? Would we observe the formation of an undesired oligomeric state of the protein?

For pharmaceutical applications, it would be interesting to explore a larger variety of solid surfaces, to see if a protein-friendly surface can be easily produced. If we could functionalise silicon or glass in such a way that the surface mimics the interfacial hydrophobic patch found in dimers, would these surfaces better prevent aggregation? Or could we make a hydrophilic surface that is less reactive than the amine-terminated one? A polarisation dependent sum frequency study of the amide I region to measure the insulin monomer orientations at various interfaces would be highly valuable here. This would have a great impact on the production and delivery of insulin based drugs.

Even though some of the results we obtained are puzzling, we could demonstrate the potential of reflection anisotropy spectroscopy as an easy alternative to scanning microscopies that are currently used for studying fibril formation and morphology.

Appendix A

From nonlinear hyperpolarizability to nonlinear susceptibility

In this section we show how to relate molecular hyperpolarisabilites and second order nonlinear susceptility in order to derive the molecular orientation of the -CH_3 end groups in the TMPS SAMs. As shown by Wang *et al.* [145] it is possible to calculate the 11 nonvanishing $\beta_{\alpha\beta\gamma}$ for the -CH_3 group with $C_{3\nu}$ symmetry, and they can be written as follows:

$$\begin{aligned}
\beta_{aac} = \beta_{bbc} &= \frac{-3G_{sym}\beta^0_{CH}}{2\omega_s}[(1+r) - (1-r)\cos^2\tau]\cos\tau \\
\beta_{ccc} &= \frac{-3G_{sym}\beta^0_{CH}}{\omega_s}[r + (1-r)\cos^2\tau]\cos\tau \\
\beta_{bcb} = \beta_{cbb} = \beta_{caa} = \beta_{aca} &= \frac{-3G_{deg}\beta^0_{CH}}{\omega_d}(1-r)\sin^2\tau\cos\tau \\
\beta_{aaa} = \beta_{bba} = \beta_{abb} = \beta_{bab} &= \frac{3G_{deg}\beta^0_{CH}}{4\omega_d}(1-r)\sin^3\tau
\end{aligned} \tag{A.1}$$

where τ is the H-C-H bond angle ($\approx 109.5°$) for the -CH_3 group, $G_{sym} = \frac{(1+2\cos\tau)}{M_C} + \frac{1}{M_H}$ and $G_{deg} = \frac{(1-\cos\tau)}{M_C} + \frac{1}{M_H}$ are the inverse effective masses

for the symmetric and asymmetric modes respectively (M_C and M_H are the atomic masses of carbon and hydrogen), ω_s and ω_d are the vibrational frequencies of symmetric and asymmetric stretching modes. Finally r is the single bond polarisability derivative ratio. r=0.14 is the value typically used for a single CH bond in the CH_3 group in alkanes [146]. Even if r varies from molecule to molecule, r=0.14 has been successfully used by Shen *et al.* [147] for interpreting SF spectra of CH_2 at a polyvinyl alcohol surface. Substituting $G_{sym} = \frac{37}{36} = 1.028$, $G_{deg} = \frac{40}{36} = 1.111$ and r=0.14 into eq.A.7 we obtain:

$$\begin{aligned}
\beta_{aac} = \beta_{bbc} &= \beta^0_{CH}\frac{-1.028}{2\omega_s}[(1+0.14)-(1-0.14)\cos^2(109.5)]\cos(109.5) = \\
&= 0.537\frac{\beta^0_{CH}}{\omega_s} \\
\beta_{ccc} &= \beta^0_{CH}\frac{-1.028}{\omega_s}[0.14+(1-0.14)\cos^2(109.5)]\cos(109.5) = \\
&= 0.247\frac{\beta^0_{CH}}{\omega_s} \\
\beta_{bcb} = \beta_{cbb} = \beta_{caa} = \beta_{aca} &= \frac{-1.111\beta^0_{CH}}{\omega_d}(1-0.14)\sin^2(109.5)\cos(109.5) = \\
&= 0.849\frac{\beta^0_{CH}}{\omega_d} \\
\beta_{aaa} = \beta_{bba} = \beta_{abb} = \beta_{bab} &= 3\frac{1.111\beta^0_{CH}}{4\omega_d}(1-0.14)\sin^3(109.5) = \\
&= 0.600\frac{\beta^0_{CH}}{\omega_d}
\end{aligned} \tag{A.2}$$

where ω_d=2958cm^{-1} and ω_s=2877cm^{-1}, $\beta_{aac} = \beta_{bbc}$ and β_{ccc} relates to the symmetric stretching mode, while $\beta_{aca} = \beta_{bcb}$, $\beta_{caa} = \beta_{cbb}$ and $\beta_{aaa} = -\beta_{bba} = -\beta_{abb} = -\beta_{bab}$ are asymmetric elements. Hirose *et al.* [62] have derived all the coefficients to rotate $\beta_{\alpha\beta\gamma}$ into the lab axis frame to obtain β_{ijk}.

$$\beta_{ijk} = \sum u_{ijk:\alpha\beta\gamma}\beta_{\alpha\beta\gamma} \tag{A.3}$$

β_{ijk} further relate to the susceptibility tensor χ^2_{ijk} as follows:

$$\chi_{ijk} = N_s \sum_{ijk} \langle R(\phi)R(\theta)R(\psi)\rangle \, \beta_{\alpha\beta\gamma} \tag{A.4}$$

where $R(\phi)$, $R(\theta)$ and $R(\psi)$ are the elements of the rotational transformation matrix from the molecular to the lab coordinates. For a surface with a C_∞ symmetry, the non-zero elements of $\chi^{(2)}_{ijk}$ are obtained by integrating over the two angles ϕ and ψ.

For the symmetric stretching mode we obtain:

$$\begin{aligned}
\chi^{(2,sym)}_{xxz} = \chi^{(2,sym)}_{yyz} =& \frac{1}{2} N_s \beta_{ccc} \left[(1+R)\langle\cos\theta\rangle - (1-R)\langle\cos^3\theta\rangle\right] \\
\chi^{(2,sym)}_{xzx} = \chi^{(2,sym)}_{zxx} = \chi^{(2,sym)}_{yzy} = \chi^{(2,sym)}_{zyy} =& \frac{1}{2} N_s \beta_{ccc}(1-R)\left[\langle\cos\theta\rangle - \langle\cos^3\theta\rangle\right] \\
\chi^{(2,sym)}_{zzz} =& N_s \beta_{ccc}\left[R\langle\cos\theta\rangle + (1-R)\langle\cos^3\theta\rangle\right]
\end{aligned} \tag{A.5}$$

For the asymmetric mode:

$$\begin{aligned}
\chi^{(2,asym)}_{xxz} = \chi^{(2,asym)}_{yyz} =& -N_s \beta_{aca}\left(\langle\cos\theta\rangle - \langle\cos^3\theta\rangle\right) \\
\chi^{(2,asym)}_{xzx} = \chi^{(2,asym)}_{zxx} = \chi^{(2,asym)}_{yzy} = \chi^{(2,asym)}_{zyy} =& N_s \beta_{aca}\langle\cos^3\theta\rangle \\
\chi^{(2,asym)}_{zzz} =& 2N_s \beta_{aca}\left(\langle\cos\theta\rangle - \langle\cos^3\theta\rangle\right)
\end{aligned} \tag{A.6}$$

In the same way, also the values of $\beta_{\alpha\beta\gamma}$ can be derived for the CH_2 group:

$$\begin{aligned}
\beta_{aac} &= \frac{G_a \beta^0_{CH}}{2\omega_{a1}}[(1+r) - (1-r)\cos\tau]\cos\left(\frac{\tau}{2}\right) \\
&= \frac{0.869\beta^0_{CH}}{2\omega_{a1}} \\
\beta_{bbc} &= \frac{G_a \beta^0_{CH}}{2\omega_{a1}} r\cos\left(\frac{\tau}{2}\right) \\
\beta_{ccc} &= \frac{G_a \beta^0_{CH}}{\omega_{a1}}[(1+r) + (1-r)\cos\tau]\cos\left(\frac{\tau}{2}\right) \\
\beta_{caa} = \beta_{aca} &= \frac{G_a \beta^0_{CH}}{\omega_{a1}}(1-r)\sin\tau\sin\left(\frac{\tau}{2}\right) \\
\beta_{bcb} = \beta_{cbb} &= 0
\end{aligned} \tag{A.7}$$

where $G_a = (1+\cos\tau)/M_c+1/M_H = 1.056$ and $G_b = (1-\cos\tau)/M_c+1/M_H = 1.111$ are the inverse effective mass for the symmetric and asymmetric stretching mode. The $\chi^{(2)}_{ijk}$ tensor elements for the symmetric stretching are:

$$\begin{aligned}
\chi^{(2,sym)}_{xxz} = \chi^{(2,sym)}_{yyz} =& \frac{1}{4}N_s(\beta_{aac}+\beta_{bbc}+2\beta_{ccc})\langle\cos\theta\rangle \\
& \frac{1}{4}N_s(\beta_{aac}+\beta_{bbc}-2\beta_{ccc})\langle\cos^3\theta\rangle \\
\chi^{(2,sym)}_{xzx} = \chi^{(2,sym)}_{zxx} =& \chi^{(2,sym)}_{yzy} = \chi^{(2,sym)}_{zyy} \\
=& \frac{1}{4}N_s(\beta_{aac}+\beta_{bbc}-2\beta_{ccc})\left[\langle\cos\theta\rangle - \langle\cos^3\theta\rangle\right] \\
\chi^{(2,sym)}_{zzz} =& \frac{1}{4}N_s(\beta_{aac}+\beta_{bbc})\langle\cos\theta\rangle - \frac{1}{2}N_s(\beta_{aac}+\beta_{bbc}-2\beta_{ccc})\langle\cos^3\theta\rangle
\end{aligned} \tag{A.8}$$

For the asymmetric mode:

$$\begin{aligned}
\chi^{(2,asym)}_{xxz} = \chi^{(2,asym)}_{yyz} =& -\frac{1}{2}N_s\beta_{aca}\left(\langle\cos\theta\rangle - \langle\cos^3\theta\rangle\right) \\
\chi^{(2,asym)}_{xzx} = \chi^{(2,asym)}_{zxx} = \chi^{(2,asym)}_{yzy} = \chi^{(2,asym)}_{zyy} =& \frac{1}{2}N_s\beta_{aca}\langle\cos^3\theta\rangle \\
\chi^{(2,asym)}_{zzz} =& N_s\beta_{aca}\left(\langle\cos\theta\rangle - \langle\cos^3\theta\rangle\right)
\end{aligned} \tag{A.9}$$

We can write the SF intensity in ppp polarisation combination as (see eq.3.26) $I_{ppp} \propto \left|\chi^{(2)}_{ppp}\right|^2$, where $\chi^{(2)}_{ppp}$ is:

$$\begin{aligned}
\chi^{(2)}_{ppp} =& -L_x(\omega_{SF})K_x(\omega_{VIS})K_z(\omega_{IR})\cos\alpha_{SF}\cos\alpha_{VIS}\sin\alpha_{IR}\chi^{(2)}_{xxz} \\
& -L_x(\omega_{SF})K_z(\omega_{VIS})K_x(\omega_{IR})\cos\alpha_{SF}\sin\alpha_{VIS}\cos\alpha_{IR}\chi^{(2)}_{xzx} \\
& +L_z(\omega_{SF})K_x(\omega_{VIS})K_x(\omega_{IR})\sin\alpha_{SF}\cos\alpha_{VIS}\cos\alpha_{IR}\chi^{(2)}_{zxx} \\
& +L_z(\omega_{SF})K_z(\omega_{VIS})K_z(\omega_{IR})\sin\alpha_{SF}\sin\alpha_{VIS}\sin\alpha_{IR}\chi^{(2)}_{zzz}
\end{aligned} \tag{A.10}$$

Here L and K are Fresnel factors (as discussed in chapter 3) and α are the incident angles of SFG, visible and IR light. We have computed the intensities of the symmetric and asymmetric vibrational mode for both CH_3 and CH_2

groups for the case of the Silicon/air interface as shown in fig. A.2. As discussed by Guo *et al.* [124], the intensity of the symmetric stretch is shared with the Fermi resonance, which leads to an overall intensity reduction in SFG. The appropriate correction can be derived from FT-IR where the total intensity of the symmetric modes is conserved. As the FT-IR spectra in fig. A.1 show, the Fermi resonance at $\sim$2930 cm^{-1} is quite prominent in comparison to the symmetric stretch at $\sim$2873 cm^{-1}. As it strongly overlaps with the asymmetric CH_2 stretch at $\sim$2917 cm^{-1}, the intensity ratio cannot be accurately determined from a fit, but is estimated to be as close to the maximum possible value of 1:1. This reduces the total symmetric SF intensity by a factor 2, which is incorporated in fig. A.2. In the worst case scenario of no intensity sharing, the experimental determined ratio R=1.05 would correspond to a tilt angle of $\sim 40^\circ$ instead of $\sim 33^\circ$ for maximum intensity sharing.

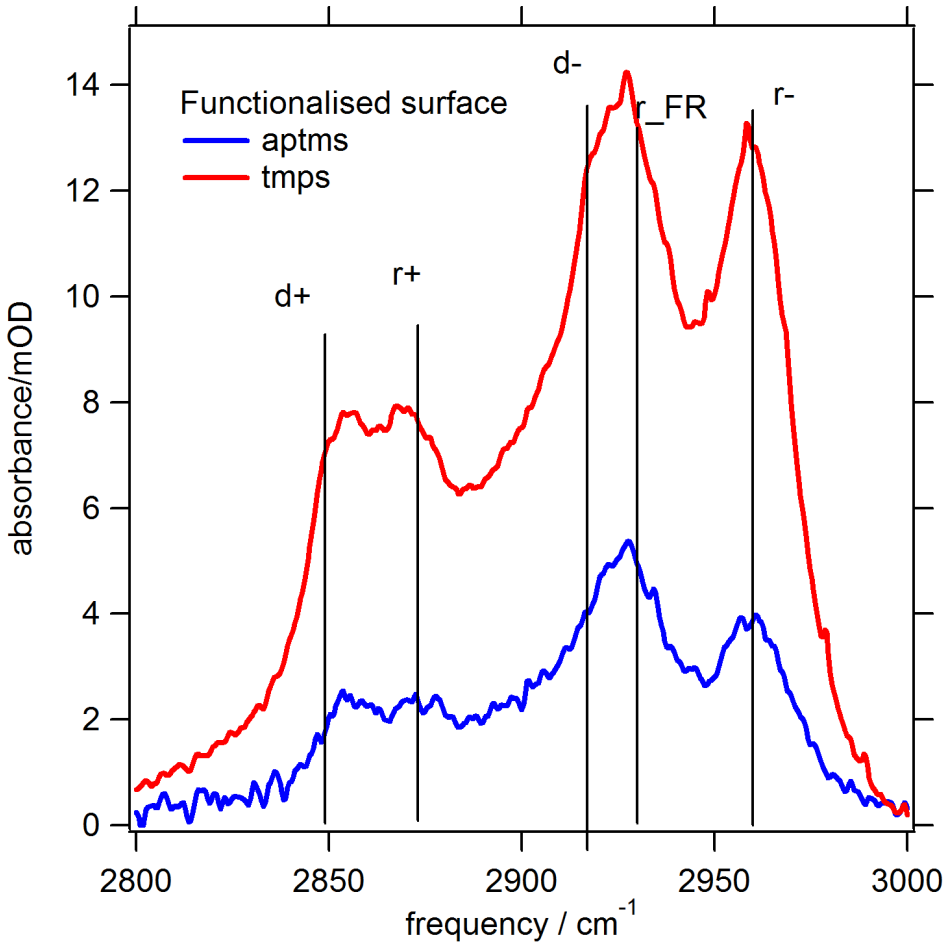

Figure A.1: *ATR FT-IR spectra of TMPS and APTMS monolayers on silicon*

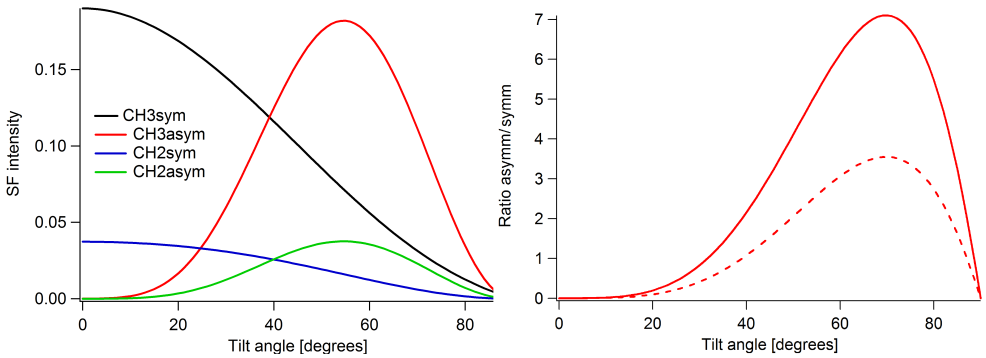

Figure A.2: *Left: SF intensities (ppp polarisation combination) of symmetric and asymmetric CH stretching in methyl and methylene groups plotted versus tilt angle, at the silicon/air interface.Right: asymmetric/symmetric -CH_3 stretching intensity ratio. Solid line: strong Fermi resonance taken into account. Dashed line: Fermi resonance is assumed to be negligible.*

Appendix B

Self-assembled monolayers on gold surfaces

The preparation of alkenethiol self assembled monolayers on gold surfaces was first proposed by Bain [148] [53]. For our sample preparation we do not use a single crystalline gold surface, but we use a commercially available gold substrate (Arrandee - www.arrandee.com) which is known to form Au(111) islands after annealing. The gold substrates (on glass) are briefly rinsed with ethanol after annealing. They are then cleaned in Piranha solution (at 70°C) for 3 minutes, and rinsed with a copious amount of milli-Q ($\sim 18.2 M\Omega$) water. Substrates are then carefully dried under an argon flow. A solution of OctaDecaneThiol (ODT) in ethanol is prepared at a concentration of 5 mM. The glassware used during the SAM preparation is cleaned in Piranha solution (at 70°C) for 15 minute. Glassware is then rinsed with a copious amount of milli-Q water, and finally is rinsed with solvent (ethanol). The gold substrates are left in solution overnight. After the SAM is formed, the samples are sonicated in ethanol for 5 minutes, and finally rinsed thoroughly with milli-Q and dried under a argon flow.

Appendix C

Supporting info for: Insulin Adsorption on model hydrophilic and hydrophobic surfaces

In this section I explain how the techniques presented in chapter 3 are used to quantify the amount of protein adsorbed on solid surfaces as presented in chapter 5. The combination of XPS spectroscopy, IR spectroscopy and QCM-D allows us to quantify both dry and wet mass of adsorbed insulin layers.

Analysis of XPS spectra

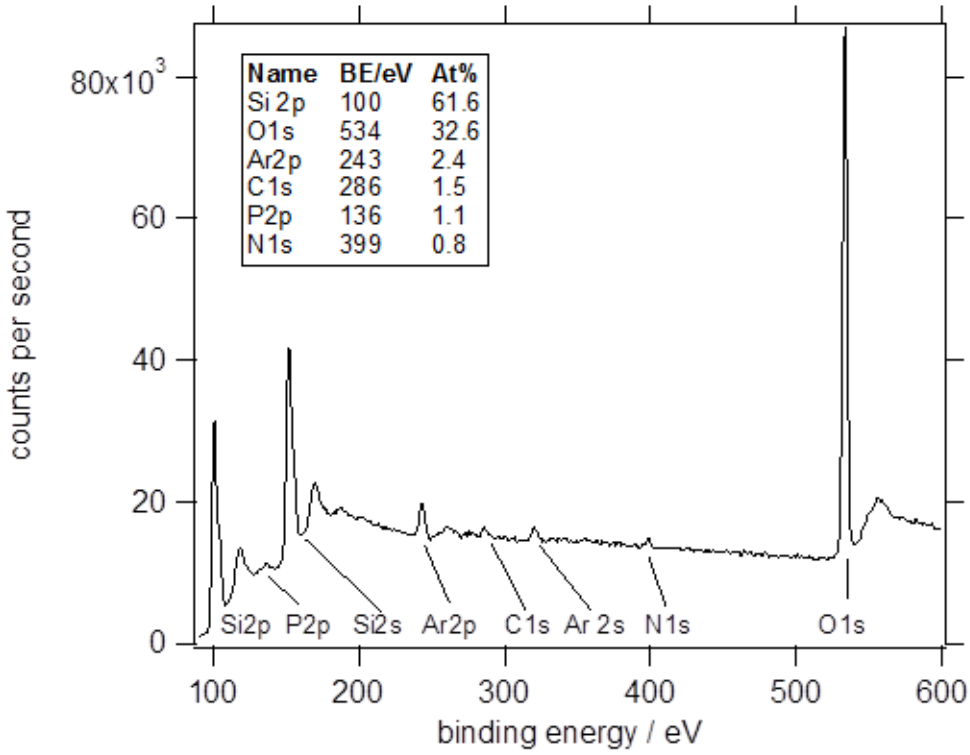

Figure C.1: *Survey spectrum of UV ozone cleaned QCM sensor, showing small residual contamination from carbon. Nitrogen and argon embedded in the crystal stem most likely from the sputter coating process used to manufacture Si coated QCM-D crystals. For this freshly cleaned QCM crystal, the Si:SiO_2 peak ratio is 2.4, which corresponds to a film thickness of 1.2 nm [78], given an electron attenuation length of Si 2p electrons of 3.5 nm [149].*

The oxide and carbon film thicknesses for the various sensors derived using [78] are summarized in Table S1. It can be seen that the values for TMPS are similar to the expected molecular length of 0.73 nm, while APTMS values are lower, indicating a generally lower coverage compared to TMPS. The protein layer thickness is derived from the simplified model by Ray and Shard [79].

Sample pH2.7	Ar 2p energy [eV]	Correction to 241.2 eV
APTMS before	242.60	-1.40
TMPS before	242.58	-1.38
APTMS after	242.40	-1.20
TMPS after	242.25	-1.05

Table C.1: *Energy correction of spectra using Ar* $2p_{\frac{1}{2}}$ *binding energy.*

QCM-D

The data are analysed using Q-Tools (Q-Sense, Goteborg) program, which fits the data using a Voigt model, providing the user with thickness and viscoelastic properties of the adsorbed layer [72]. The resulting units are ng/cm^2, we use a typical protein density of 1.35 g/cm^3 to convert to thickness.

UV-vis absorbance

To determine the concentration of the insulin solution, we used a molar extinction coefficient of 5530 $M^{-1}cm^{-1}$ at 277.5 nm [5], and then recorded FTIR transmission spectra of the amide I band in a transmission cell with a 100 μm spacer.

Infrared ATR analysis

We used a silicon ATR crystal with 80 mm length at the base of the trapezoid, 4 mm thickness and $\gamma = 45°$ face angle. To record spectra of dried insulin

layers, the whole ATR crystal is immersed in insulin solution and then dried, leading to N=19 internal reflections. For wet spectra, insulin is only in contact with the trapezoid base, leading to N=10 internal reflections. We obtain the layer thickness in molcm^{-2} for a dried layer from [47]:

$$\Gamma = \frac{A}{N\cos\gamma} \cdot \frac{1}{\varepsilon} \cdot \frac{n_1}{4n_2} \cdot \sqrt{1 - \frac{n_3^2}{n_1^2}} = 0.0406 \cdot \frac{A}{\varepsilon} \tag{C.1}$$

where we have inserted n_1= 3.4202, the refractive index of Si at 6 μm, n_2=1.5, the refractive index of the protein and n_3=1 as the refractive index of air. A is the measured absorbance and ε is the molar extinction coefficient. The prefactor for water (n_3=1.265) and 10 reflections is 0.0749. For wet spectra, we also have to take into account the absorption by insulin in solution. The penetration depth is:

$$d_p = \frac{\lambda}{2\pi\ n_{si}\sqrt{sin^2(45^\circ) - \left(\frac{n_{HI}^2}{n_{Si}^2}\right)}} = 5.085 \cdot 10^{-5}\text{cm} \tag{C.2}$$

from which we derive a solution absorbance of $dp \times \varepsilon \times \text{conc} = 0.13mOD$ per reflection, ie subtract 1.3mOD from the whole spectrum. This analysis has the disadvantage of simply relying on peak absorbance and does not take the shape of the amide I band into account. We therefore also determined the integrated extinction coefficient from an insulin solution spectrum:

$$\psi = \frac{1}{dc}\int A d\lambda \tag{C.3}$$

And then used a modified version of the equation above:

$$\Gamma = 0.0406 \cdot \frac{\int A d\lambda}{\psi} \tag{C.4}$$

The peak absorbance of a 75μM insulin solution in a liquid cell with 100 μm spacer is 0.025, leading to a peak molar extinction coefficient of ε=33,333

$M^{-1}cm^{-1}$. Integrating the absorbance from 1575 to 1775 cm^{-1}, leads to an integral of 1.55cm^{-1} and therefore an integrated extinction coefficient of ψ=2.07×10^9 cm mol^{-1}. A 1 hr adsorption dried layer on APTMS was found to have an integrated absorbance of 0.62 cm^{-1}, so Γ=0.0406 ×0.62cm^{-1}/2.07 ×10^9 $cmmol^{-1}$=1.22×10^{-11} $molcm^{-2}$=7.1 ×10^{-8} g/cm^2= 0.71 mg/m^2.

Silane at pD=2.7	Integrated absorbance $[cm^{-1}]$	$\Gamma[mg/m^2]$
APTMS	0.61	0.69
TMPS	0.94	1.07

Silane	Time	Integrated absorbance $[cm^{-1}]$	$\Gamma[mg/m^2]$
	30	0.215	0.56
APTMS	90	0.217	0.56
	180	0.247	0.64
	30	0.298	0.77
TMPS	90	0.423	1.10
	180	0.508	1.32

Table C.2: *Coverages derived for dry and wet spectra.*

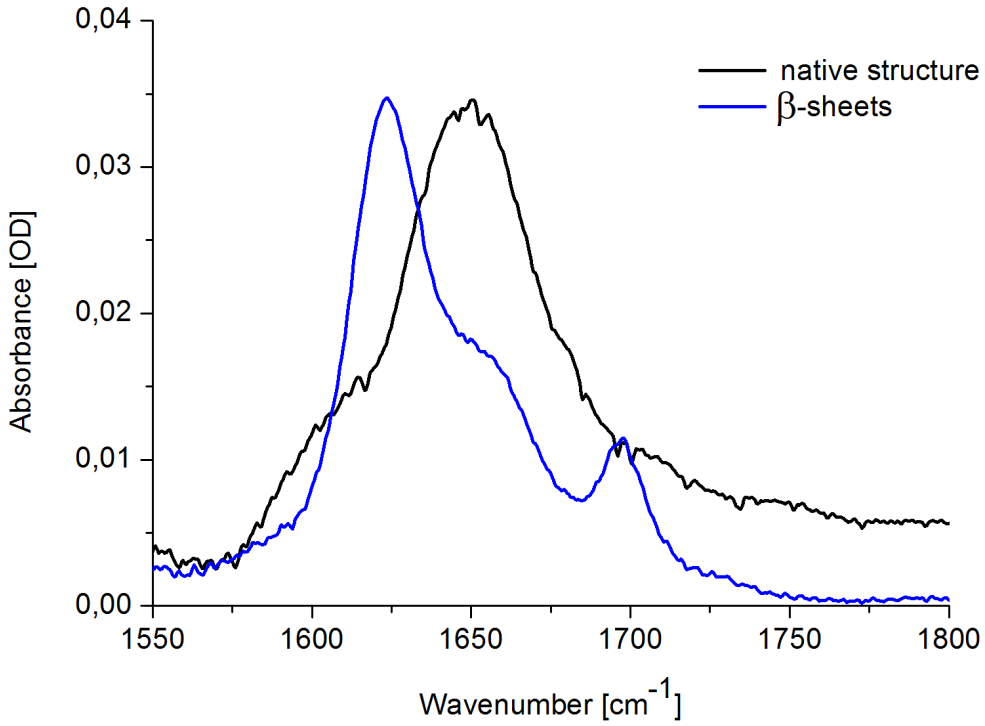

Figure C.2: *FT-IR spectra of humam insulin. Black: insulin in its native form. Blue: insulin β-sheets after incubating solutions at high temperature.*

To determine the net CH_3 and CH_3 orientation of the insulin we used the 2JV1 .pdb file, which incorporates atomic coordinates for hydrogens. For each CH_3 or CH_2 group, we calculate the vector sum of the individual C-H bonds and then divide by 3 or 2, respectively. We add all vectors and then divide by the length of an individual one to obtain the net number of contributing groups and their orientation. We obtain a net contribution of 9.4 CH_3 groups and 6.8 CH_2 groups. We define the direction of the hydrophobic domain that forms part of the dimer interface using the following atoms in the B chain:

ATOM	742	CG	PRO
ATOM	665	O	PHE
ATOM	548	CG	TYR

When we calculate the angle between the normal to the plane spanned by these three atoms and the CH_2 and CH_2 vectors, we obtain an angle of 65 degrees between the normal and the net CH_3 vector and an angle of 32 degrees between the normal and the net CH_2 vector. As the calculation in appendix A shows, the expected CH_2 intensity is quite small for this orientation, while the expected intensity of the asymmetric CH_3 stretch is much larger, which could qualitatively explain the changes we observe in the SFG spectrum of the TMPS-functionalised surface upon insulin adsorption.

Appendix D

Chiral SFG of LKα_{14} peptide

This series of experiments has been done at the Max Planck Institute for Polymer Research, Mainz, Germany, in the group of Prof. Mischa Bonn under the supervision of Dr. Tobias Weidner.

SFG is often used to study protein secondary structure at interfaces [150] [151]. The peak assignment in the amide I region is not straightforward for the same reason as in FT-IR: the amide I band overlaps with the water bending mode, the bands from structural motifs are broad and they overlap (e.g the signal from α-helices and random coils occurs at about the same frequency). From the technical point of view, SFG in the amide region requires the generation of femtosecond pulses in the mid-IR. Optical parametric amplification in the mid-IR is not as efficient as in other spectral regions, and IR light at about 6 μm is strongly absorbed by water vapor. In recent years, chiral SFG has been proposed as a novel technique for identifying protein secondary structure in the amide I and N-H stretching region. This work has mainly been carried out by Yan's grop at Yale Univeristy [152] [97]. They propose that the origin of a protein's chiral SF signal originates from the pro-

tein's backbone. The technique has been tested on model peptides and it has been proposed that α-helices are chiral-SFG active only in the N-H stretching region, while β-sheets show a chiral SFG response only in the amide I region. In fig.D.1 we show the experimental spectra (chiral and achiral) presented by E. Yan *et al.* .

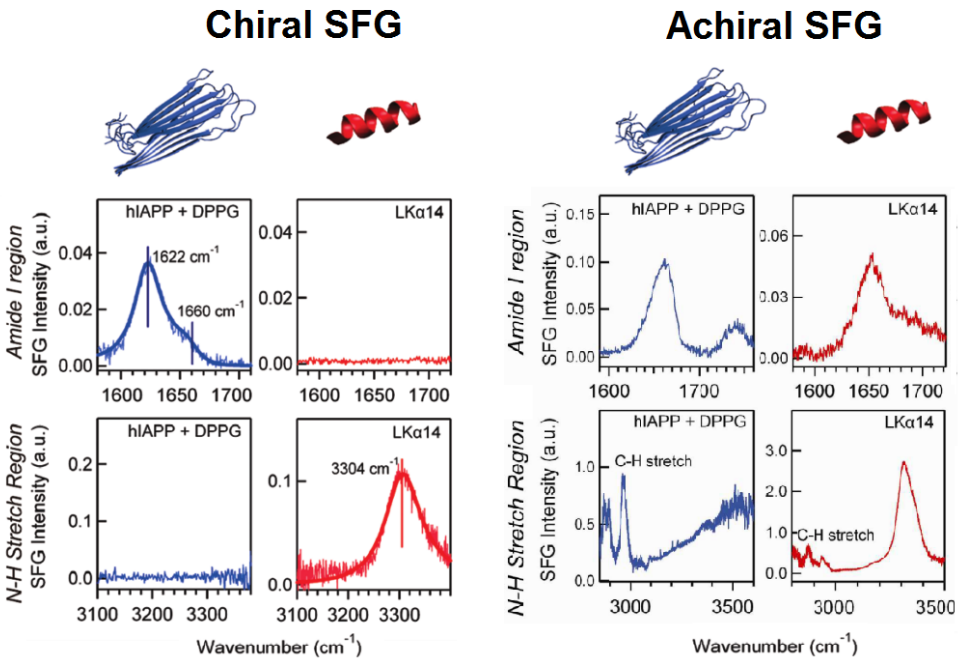

Figure D.1: *Left: Chiral spectra of β-sheets and LKα_{14} in the amide I region (top) and N-H stretching (bottom). Right: Achiral spectra of β-sheets and LKα_{14} in the amide I region (top) and N-H stretching (bottom). Figure adapted from [152] and supporting info*

Here we focus on the LKα_{14} peptide, which is a 14 amino acid peptide made of leucine and lysine residues, arranged in such a way that it folds into a helical shape at hydrophobic interfaces [153]. This peptide has been extensively studied by Somorjai and coworkers [154] on solid surfaces. Weidner *et al.* [155] have also studied the LKα peptide at the air/water interface.

They confirmed the helical structure of the peptide, and in addition showed that the hydrophilic side of the peptide does not order in water. The N-H stretching mode appears only when the peptide is deposited on a hydrophilic surface. They moreover performed an experiment with labeled N atoms in the side chains to prove that the recorded N-H signal can be assigned indeed to the lysine's amino end group [156]. In fig.D.2 we show an overview of the experimental results presented in various works by Somorjai and Weidner.

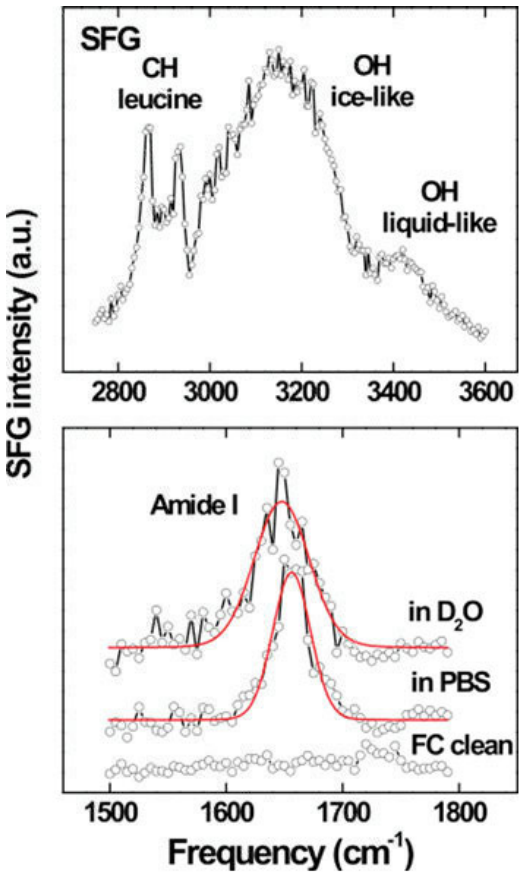

Figure D.2: *Spectra of $LK\alpha_{14}$ at the air/water interface in the -NH stretching region (top) and in the amide I region (bottom). Figures adapted from [157]*

I repeated some of these to test the possibility of doing chiral SFG on more complicated proteins such as insulin. Here I show the spectra (fig.D.3)

obtained of LKα peptide at the air/water interface. Spectra were recorded in both ssp (achiral) and psp (chiral) polarisation combinations. LKα_{14} is dissolved in water at neutral pH at 25μg/ml concentration. Spectra are accumulated for 30 minutes (achiral) and 60 minutes (chiral). The IR light wavelength is set to 3.1μm and the visible wavelength is 803 nm.

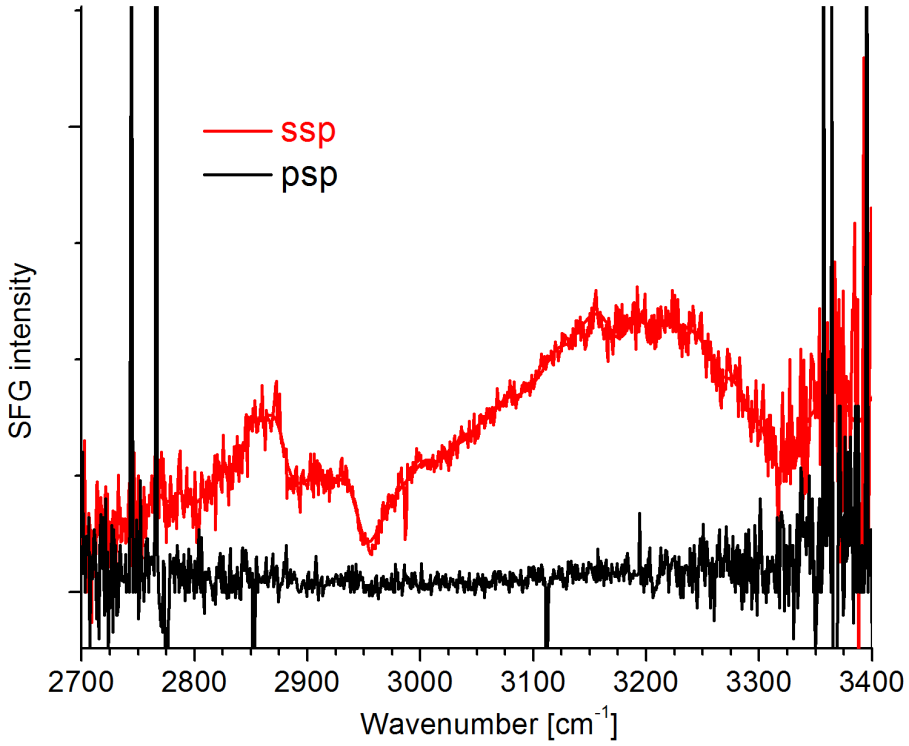

Figure D.3: *Measured spectra of LKα_{14} at the air/water interface in the -NH stretching region.*

The spectral features in the -CH stretching region are due to leucine's side chain orientation into the air phase [155]. The broad peak at around 3200 cm^{-1} arises from the weakly bonded water at the surface. There is a sign of features assignable to -NH stretching vibrations (at around 3300 cm^{1}), but not as prominent as reported by Fu *et al.* [152]. Our measurement, according to the literature [154] [157] does not show any N-H stretching peak,

confirming that no signal from the lysines' side chains, and therefore no N-H backbone signal is recorded. We could not reproduce the chiral spectra of the LKα peptide reported by Yan. Similarly, Engel *et al.* [136] failed to detect a chiral SFG signal of human islet amyloid polypeptide (hIAPP) at the air/water interface in the amide I region. Some of the discrepancies could arise from polarisation leakage in the experiments. The chiral signal is expected to be small, and leakage of light with another polarisation state might generate a nonchiral signal that overwhelms the chiral one, causing spectral misinterpretations. Differences in the behaviour in the different SFG setups could arise from sample heating due to tighter focusing, different pulse energies and repetition rate. Yan's group can afford to loosely focus IR and VIS light at the sample stage thanks to a higher repetition rate and higher pulse energies in their laser setup.

Appendix E

Additional figures

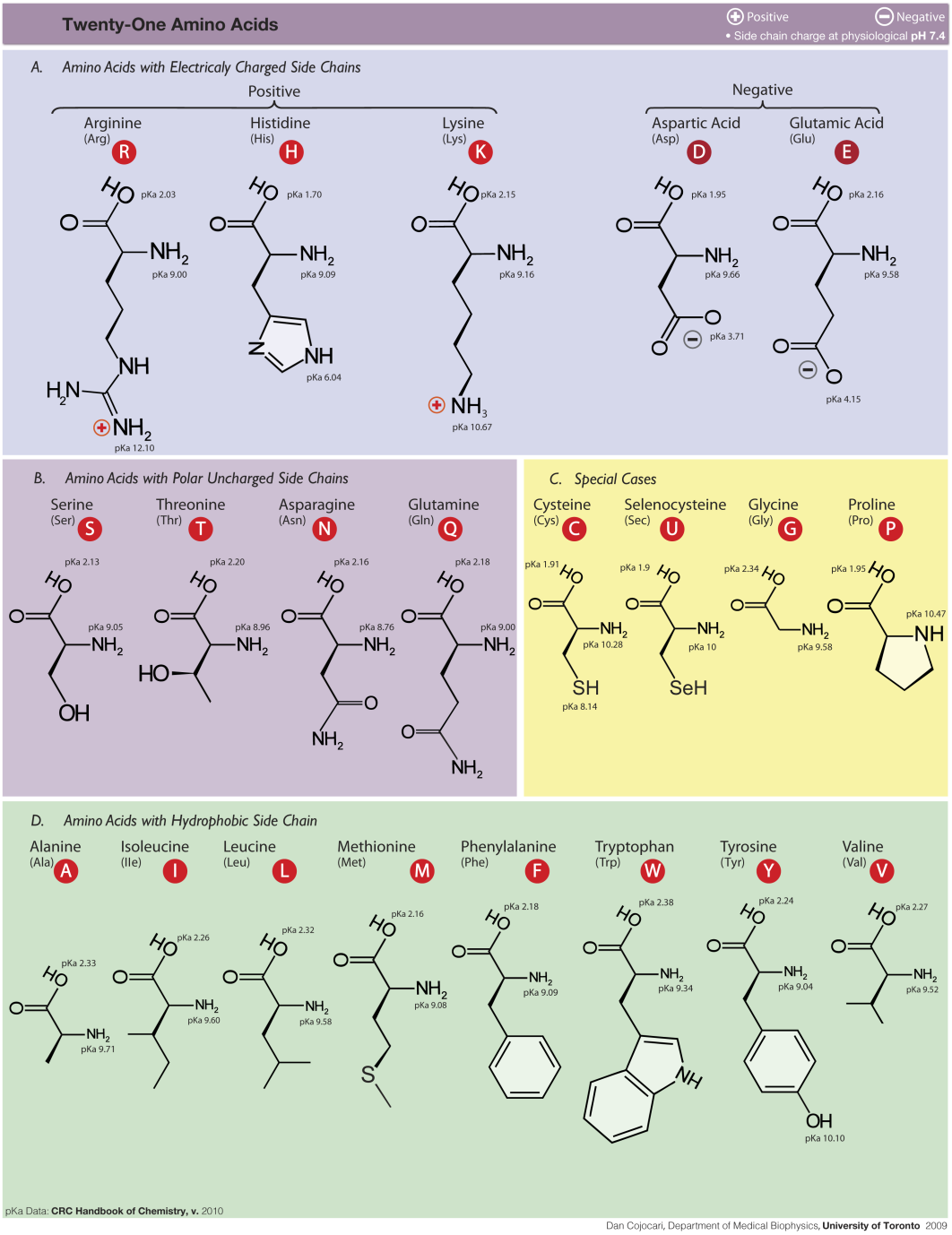

Figure E.1: *Chart of the 21 amino acids available in nature.*

Bibliography

[1] W. Wang, "Instability, stabilization, and formulation of liquid protein pharmaceuticals.," *International Journal of Pharmaceutical*, vol. 185, pp. 129–88, Aug. 1999.

[2] K. A. Dill and D. Shortle, "Denatured states of proteins.," *Annual review of biochemistry*, vol. 60, pp. 795–825, Jan. 1991.

[3] F. Chiti and C. M. Dobson, "Protein misfolding, functional amyloid, and human disease.," *Annual review of biochemistry*, vol. 75, pp. 333–66, Jan. 2006.

[4] G. Taubes, "Misfolding the way to disease.," *Science (New York, N.Y.)*, vol. 271, pp. 1493–5, Mar. 1996.

[5] J. L. Whittingham, D. J. Scott, K. Chance, A. Wilson, J. Finch, J. Brange, and G. Guy Dodson, "Insulin at pH 2: structural analysis of the conditions promoting insulin fibre formation.," *Journal of molecular biology*, vol. 318, pp. 479–90, Apr. 2002.

[6] J. Brange, L. Andersen, E. D. Laursen, G. Meyn, and E. Rasmussen, "Toward understanding insulin fibrillation.," *Journal of pharmaceutical sciences*, vol. 86, pp. 517–25, May 1997.

[7] L. Nielsen, S. Frokjaer, J. F. Carpenter, and J. Brange, "Studies of the structure of insulin fibrils by Fourier transform infrared (FTIR) spectroscopy and electron microscopy.," *Journal of pharmaceutical sciences*, vol. 90, pp. 29–37, Jan. 2001.

[8] P. E. R. M. Claesson, T. Arnebrant, B. Bergenstahl, and T. Nylander, "Direct Measurements of the Interaction between Layers of Insulin Adsorbed on Hydrophobic Surfaces Preparation of Hydrophobed Mica Muscovite mica is a layered aluminosilicate," *Journal of Colloid and Interface Science*, vol. 130, no. 2, pp. 457–466, 1989.

[9] P. Nilsson, T. Nylander, and S. Havelund, "Adsorption of insulin on solid surfaces in relation to the surface properties of the monomeric and oligomeric forms," *Journal of Colloid and Interface Science*, vol. 144, pp. 145–152, June 1991.

[10] L. Nault, P. Guo, B. Jain, Y. Bréchet, F. Bruckert, and M. Weidenhaupt, "Human insulin adsorption kinetics, conformational changes and amyloidal aggregate formation on hydrophobic surfaces.," *Acta biomaterialia*, vol. 9, pp. 5070–9, Feb. 2013.

[11] L. Nault, C. Vendrely, Y. Bréchet, F. Bruckert, and M. Weidenhaupt, "Peptides that form β-sheets on hydrophobic surfaces accelerate surface-induced insulin amyloidal aggregation.," *FEBS letters*, vol. 587, pp. 1281–6, May 2013.

[12] E. Buxbaum, *Fundamentals of Protein Structure and Function*. Boston, MA: Springer US, 2007.

[13] A. Sakula, "Paul Langerhans (1847-1888): a centenary tribute," *Journal of the Royal Society of Medicine*, vol. 81, pp. 414–415, Jan. 1988.

[14] C. Funk, "The Isolation of Insulin," *Science*, vol. 63, no. 1633, pp. 401–2, 1926.

[15] F. Sanger and H. Tuppy, "The amino-acid sequence in the phenylalanyl chain of insulin," *Biochemical Journal*, vol. 49, pp. 463–481, 1951.

[16] T. Blundell, J. Cutfield, S. Cutfield, E. Dodson, G. Dodson, D. Hodgkin, D. Mercola, and M. Vijayan, "Atomic positions in rhombohedral 2-zinc insulin crystals," *Nature*, vol. 231, 1971.

[17] R. R. Bowsher, R. A. Lynch, P. Brown-Augsburger, P. F. Santa, W. E. Legan, J. R. Woodworth, and R. E. Chance, "Sensitive RIA for the specific determination of insulin lispro," *Clinical chemistry*, vol. 45, pp. 104–10, Jan. 1999.

[18] P. Jeffrey and J. Coates, "An equilibrium untracentrifuge study of self-association of bovine insulin," *Biochemistry*, vol. 5, no. 2, pp. 489–&, 1966.

[19] B. Frank, A. Veros, and A. Pekar, "Insulin and proinsulin conformation in solution," *Diabetes*, vol. 21, no. 2, pp. 486–&, 1972.

[20] M. L. Blader and M. F. Dunn, "Insulin hexamers: new conformations and applications," *Trends in Biochemical Sciences*, vol. 16, pp. 341–345, Jan. 1991.

[21] S. Yumlu, R. Barany, M. Eriksson, and C. Röcken, "Localized insulin-derived amyloidosis in patients with diabetes mellitus: a case report.," *Human pathology*, vol. 40, pp. 1655–60, Nov. 2009.

[22] Y. Shikama, J.-I. Kitazawa, N. Yagihashi, O. Uehara, Y. Murata, N. Yajima, R. Wada, and S. Yagihashi, "Localized Amyloidosis at

the Site of Repeated Insulin Injection in a Diabetic Patient," *Internal Medicine*, vol. 49, no. 5, pp. 397–401, 2010.

[23] J. D. Sipe and A. S. Cohen, "Review: history of the amyloid fibril.," *Journal of structural biology*, vol. 130, pp. 88–98, June 2000.

[24] C. A. Ross and M. A. Poirier, "Protein aggregation and neurodegenerative disease.," *Nature medicine*, vol. 10 Suppl, pp. S10–7, July 2004.

[25] A. T. Petkova, Y. Ishii, J. J. Balbach, O. N. Antzutkin, R. D. Leapman, F. Delaglio, and R. Tycko, "A structural model for Alzheimer's beta -amyloid fibrils based on experimental constraints from solid state NMR.," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 99, pp. 16742–7, Dec. 2002.

[26] J. F. Smith, T. P. J. Knowles, C. M. Dobson, C. E. Macphee, and M. E. Welland, "Characterization of the nanoscale properties of individual amyloid fibrils.," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 103, pp. 15806–11, Oct. 2006.

[27] M. I. Ivanova, S. A. Sievers, M. R. Sawaya, J. S. Wall, and D. Eisenberg, "Molecular basis for insulin fibril assembly.," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 106, pp. 18990–5, Nov. 2009.

[28] S. Tzotzos and A. J. Doig, "Amyloidogenic sequences in native protein structures.," *Protein science*, vol. 19, pp. 327–48, Feb. 2010.

[29] J. L. Jiménez, E. J. Nettleton, M. Bouchard, C. V. Robinson, C. M. Dobson, and H. R. Saibil, "The protofilament structure of insulin amyloid fibrils.," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 99, pp. 9196–201, July 2002.

[30] I. Amenabar, S. Poly, W. Nuansing, E. H. Hubrich, A. A. Govyadinov, F. Huth, R. Krutokhvostov, L. Zhang, M. Knez, J. Heberle, A. M. Bittner, and R. Hillenbrand, "Structural analysis and mapping of individual protein complexes by infrared nanospectroscopy.," *Nature communications*, vol. 4, p. 2890, Jan. 2013.

[31] W. Nuansing, D. Frauchiger, F. Huth, A. Rebollo, R. Hillenbrand, and A. M. Bittner, "Electrospinning of peptide and protein fibres: approaching the molecular scale," *Faraday Discussions*, vol. 166, pp. 209–221, 2013.

[32] J. J. Gray, "The interaction of proteins with solid surfaces.," *Current opinion in structural biology*, vol. 14, pp. 110–5, Feb. 2004.

[33] E. Dickinson, "Adsorbed protein layers at fluid interfaces: interactions, structure and surface rheology," *Colloids and surfaces B: Biointerfaces*, vol. 15, pp. 161–176, 1999.

[34] A. Barth, "Infrared spectroscopy of proteins.," *Biochimica et biophysica acta*, vol. 1767, pp. 1073–101, Sept. 2007.

[35] A. Barth and C. Zscherp, "What vibrations tell about proteins," *Quarterly Reviews of Biophysics*, vol. 35, pp. 369–430, Nov. 2002.

[36] W. Surewicz, H. Mantsch, and D. Chapman, "Determination of protein secondary structure by Fourier transform infrared spectroscopy: a critical assessment," *Biochemistry*, vol. 32, no. 2, 1993.

[37] J. Bandekar, "Amide modes and protein conformation.," *Biochimica et biophysica acta*, vol. 1120, pp. 123–43, Apr. 1992.

[38] N. Nevskaya and Y. Chirgadze, "Infrared spectra and resonance interactions of amide I and II vibrations of αhelix," *Biopolymers*, vol. 15, pp. 637–648, 1976.

[39] Y. Chirgadze and N. Nevskaya, "Infrared spectra and resonance interaction of amide I vibration of the antiparallel chain pleated sheet," *Biopolymers*, vol. 15, pp. 607–625, 1976.

[40] Y. Chirgadze and N. Nevskaya, "Infrared spectra and resonance interaction of amide I vibration of the parallel chain pleated sheet," *Biopolymers*, vol. 15, pp. 627–636, 1976.

[41] E. Goormaghtigh, V. Cabiaux, and J. M. Ruysschaert, "Determination of soluble and membrane protein structure by Fourier transform infrared spectroscopy. III. Secondary structures.," *Sub-cellular biochemistry*, vol. 23, pp. 405–50, Jan. 1994.

[42] H. Susi, D. M. Byler, and J. M. Purcell, "Estimation of beta-structure content of proteins by means of deconvolved FTIR spectra.," *Journal of biochemical and biophysical methods*, vol. 11, pp. 235–40, Oct. 1985.

[43] J. L. Arrondo, A. Muga, J. Castresana, and F. M. Goñi, "Quantitative studies of the structure of proteins in solution by Fourier-transform infrared spectroscopy.," *Progress in biophysics and molecular biology*, vol. 59, pp. 23–56, Jan. 1993.

[44] M. Bouchard, J. Zurdo, E. J. Nettleton, C. M. Dobson, and C. V. Robinson, "Formation of insulin amyloid fibrils followed by FTIR simultaneously with CD and electron microscopy.," *Protein science : a publication of the Protein Society*, vol. 9, pp. 1960–7, Oct. 2000.

[45] W. Dzwolak, A. Loksztejn, and V. Smirnovas, "New insights into the self-assembly of insulin amyloid fibrils: an HD exchange FT-IR study," *Biochemistry*, no. 45, pp. 8143–8151, 2006.

[46] J. Vörös, "The density and refractive index of adsorbing protein layers," *Biophysical journal*, vol. 87, pp. 553–61, July 2004.

[47] K. K. Chittur, "FTIR/ATR for protein adsorption to biomaterial surfaces," *Biomaterials*, vol. 19, pp. 357–69, Mar. 1998.

[48] N. Harrick and F. du Pre', "Effective thickness of bulk materials and of thin films for internal reflection spectroscopy," *Applied Optics*, vol. 5, no. 11, pp. 1739–1743, 1966.

[49] H. Tompkins, "Physical basis for analysis of depth of absorbing species using internal-reflection spectroscopy," *Applied Spectroscopy*, vol. 28, no. 4, pp. 335–341, 1974.

[50] K. Ohta and R. Iwamoto, "Lower limit of the thickness of the measurable surface layer by Fourier transform infrared attenuated total reflection spectrometry," *Analytical Chemistry*, vol. 57, pp. 2491–2499, Nov. 1985.

[51] K. Ohta and R. Iwamoto, "Experimental proof of the relation between thickness of the probed surface-layer and absorbance in FT-IR ATR spectroscopy," *Applied Specctroscopy*, vol. 39, no. 3, pp. 418–425, 1985.

[52] L. M. Haupert and G. J. Simpson, "Chirality in nonlinear optics.," *Annual review of physical chemistry*, vol. 60, pp. 345–65, Jan. 2009.

[53] C. Bain, P. Davies, T. Ong, R. Ward, and M. Brown, "Quantitative analysis of monolayer composition by sum-frequency vibrational spectroscopy," *Langmuir*, vol. 7, pp. 1563–1566, 1991.

[54] Y. Shen, *The principles of nonlinear optics*. Wiley classics library, Wiley-Interscience, 2003.

[55] R. Boyd, *Nonlinear Optics*. Nonlinear Optics Series, Elsevier Science, 2008.

[56] G. Cerullo and S. De Silvestri, "Ultrafast optical parametric amplifiers," *Review of Scientific Instruments*, vol. 74, no. 1, p. 1, 2003.

[57] A. Lagutchev, S. A. Hambir, and D. D. Dlott, "Nonresonant Background Suppression in Broadband Vibrational Sum-Frequency Generation Spectroscopy," *The Journal of Physical Chemistry C*, vol. 111, pp. 13645–13647, Sept. 2007.

[58] A. G. Lambert, P. B. Davies, and D. J. Neivandt, "Implementing the Theory of Sum Frequency Generation Vibrational Spectroscopy: A Tutorial Review," *Applied Spectroscopy Reviews*, vol. 40, pp. 103–145, May 2005.

[59] C. Bain, "Sum-frequency vibrational spectroscopy of the solid/liquid interface," *J. Chem. Soc., Faraday Trans.*, vol. 91, no. 9, pp. 1281–1296, 1995.

[60] C. Yang, L. Richter, J. Stephenson, and K. Briggman, "In situ, vibrationally resonant sum frequency spectroscopy study of the self-assembly of dioctadecyl disulfide on gold," *Langmuir*, no. 27, pp. 7549–7556, 2002.

[61] B. Bourguignon, W. Zheng, S. Carrez, A. Ouvrard, F. Fournier, and H. Dubost, "Deriving the complete molecular conformation of self-assembled alkanethiol molecules from sum-frequency generation vibrational spectra," *Physical Review B*, vol. 79, p. 125433, Mar. 2009.

[62] C. Hirose, N. Akamatsu, and K. Domen, "Formulas for the analysis of surface sum-frequency generation spectrum by CH stretching modes of methyl and methylene groups," *The Journal of Chemical Physics*, vol. 96, no. 2, p. 997, 1992.

[63] E. Bulard, M.-P. Fontaine-Aupart, H. Dubost, W. Zheng, J.-M. Herry, M.-N. Bellon-Fontaine, R. Briandet, and B. Bourguignon, "The Effect of Bacterial Adhesion on Grafted Chains Revealed by the Non-Invasive Sum Frequency Generation Spectroscopy," *Spectroscopy: An International Journal*, vol. 27, no. 5-6, pp. 571–579, 2012.

[64] C. O´ Sullivan and G. Guilbault, "Commercial quartz crystal microbalances - theory and applications," *Biosensors and Bioelectronics*, vol. 14, pp. 663–670, Dec. 1999.

[65] A. Janshoff, H. Galla, and C. Steinem, "Piezoelectric Mass-Sensing Devices as Biosensors-An Alternative to Optical Biosensors?," *Angewandte Chemie ivnternational edition*, vol. 39, pp. 4004–4032, Nov. 2000.

[66] M. Rodahl, F. Höök, A. Krozer, P. Brzezinski, and B. Kasemo, "Quartz crystal microbalance setup for frequency and Q-factor measurements in gaseous and liquid environments," *Review of Scientific Instruments*, vol. 66, no. 7, p. 3924, 1995.

[67] M. Tanahashi and T. Matsuda, "Surface functional group dependence on apatite formation on self-assembled monolayers in a simulated body fluid.," *Journal of biomedical materials research*, vol. 34, pp. 305–15, Mar. 1997.

[68] G. Sauerbrey, "Verwendung von Schwingquarzen zur Wägung dünner Schichten und zur Mikrowägung," *Zeitschrift für Physik*, vol. 222, 1959.

[69] M. V. Voinova, M. Rodahl, M. Jonson, and B. Kasemo, "Viscoelastic Acoustic Response of Layered Polymer Films at Fluid-Solid Interfaces : Viscoelastic Acoustic Response of Layered Polymer Films at Fluid-Solid Interfaces : Continuum Mechanics Approach," *Physica Scripta*, vol. 59, pp. 391–396, 1999.

[70] A. Domack, O. Prucker, J. Rühe, and D. Johannsmann, "Swelling of a polymer brush probed with a quartz crystal resonator," *Physical Review E*, vol. 56, no. 1, pp. 680–689, 1997.

[71] M. Voinova, M. Jonson, and B. Kasemo, "'Missing mass' effect in biosensor's QCM applications," *Biosensors and Bioelectronics*, vol. 17, pp. 835–841, 2002.

[72] F. Höök, B. Kasemo, T. Nylander, C. Fant, K. Sott, and H. Elwing, "Variations in coupled water, viscoelastic properties, and film thickness of a Mefp-1 protein film during adsorption and cross-linking: a quartz crystal microbalance with dissipation monitoring, ellipsometry, and surface plasmon resonance study.," *Analytical chemistry*, vol. 73, pp. 5796–804, Dec. 2001.

[73] H. Lüth, *Solid Surfaces, Interfaces and Thin Films*. Graduate Texts in Physics, Berlin, Heidelberg: Springer Berlin Heidelberg, 2010.

[74] K. Siegbahn, "Electron spectroscopy for atoms, molecules, and condensed matter.," *Science*, vol. 217, pp. 111–21, July 1982.

[75] C. Fadley, R. Baird, and W. Siekhaus, "Surface analysis and angular distributions in x-ray photoelectron spectroscopy," *Journal of Electron Spectroscopy and related phenomena*, vol. 4, pp. 93–137, 1974.

[76] R. Flitsch and S. Raider, "Electron mean escape depths from x-ray photoelectron spectra of thermally oxidized silicon dioxide films on silicon," *Journal of Vacuum Science and Technology*, vol. 12, p. 305, Jan. 1975.

[77] H. Ebel, M. F. Ebel, P. Baldauf, and A. Jablonski, "The energy dependence of attenuation lengths in elements," *Surface and Interface Analysis*, vol. 12, pp. 172–173, July 1988.

[78] P. Cumpson, "The Thickogram: a method for easy film thickness measurement in XPS," *Surface and interface analysis*, vol. 406, no. April, pp. 403–406, 2000.

[79] S. Ray and A. G. Shard, "Quantitative analysis of adsorbed proteins by X-ray photoelectron spectroscopy.," *Analytical chemistry*, vol. 83, pp. 8659–66, Nov. 2011.

[80] P. Weightman, D. S. Martin, R. J. Cole, and T. Farrell, "Reflection anisotropy spectroscopy," *Reports on Progress in Physics*, vol. 68, pp. 1251–1341, June 2005.

[81] P. D. Lane, G. E. Isted, D. S. Roseburgh, and R. J. Cole, "Azimuth dependent reflection anisotropy of oriented thin films," *Applied Physics Letters*, vol. 95, no. 14, p. 141907, 2009.

[82] B. F. Macdonald, J. S. Law, and R. J. Cole, "Azimuth-dependent reflection anisotropy spectroscopy," *Journal of Applied Physics*, vol. 93, no. 6, p. 3320, 2003.

[83] T. Yasuda, D. Aspnes, D. Lee, C. Bjoikman, and G. Lucovsky, "Optical anisotropy of singular and vicinal Si-SiO_2 interfaces and H-terminated Si surfaces," *Journal of Vacuum Science and Technology A: Vacuum, Surfaces, and Films*, vol. 12, p. 1152, July 1994.

[84] S. D. Silaghi and D. R. Zahn, "Monitoring the ordering in biomolecular films on vicinal silicon surfaces by reflectance difference/anisotropy spectroscopy," *Applied Surface Science*, vol. 252, pp. 5462–5465, May 2006.

[85] K. Hingerl, R. Balderas-Navarro, A. Bonanni, P. Tichopadek, and W. Schmidt, "On the origin of resonance features in reflectance difference data of silicon," *Applied Surface Science*, vol. 175-176, pp. 769–776, May 2001.

[86] F. M. Green, I. S. Gilmore, J. L. S. Lee, S. J. Spencer, and M. P. Seah, "Static SIMS-VAMAS interlaboratory study for intensity repeatability, mass scale accuracy and relative quantification," *Surface and interface analysis*, vol. 42, no. 3, pp. 129–138, 2010.

[87] V. Sluzky, J. a. Tamada, a. M. Klibanov, and R. Langer, "Kinetics of insulin aggregation in aqueous solutions upon agitation in the presence of hydrophobic surfaces.," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 88, pp. 9377–81, Nov. 1991.

[88] Venyaminov SY and N. N. Kalnin, "Quantitative IR spectrophotometry of peptide compounds in water (H2O) solutions. II. Amide absorption bands of polypeptides and fibrous proteins in alpha-, beta-, and random coil conformations.," *Biopolymers*, vol. 30, pp. 1259–71, Jan. 1990.

[89] Z. Ganim, K. C. Jones, and A. Tokmakoff, "Insulin dimer dissociation and unfolding revealed by amide I two-dimensional infrared spectroscopy.," *Physical chemistry chemical physics : PCCP*, vol. 12, pp. 3579–88, Apr. 2010.

[90] Y. Pocker and S. B. Biswas, "Conformational dynamics of insulin in solution. Circular dichroic studies," *Biochemistry*, vol. 19, pp. 5043–5049, Oct. 1980.

[91] S. Johnson, W. Liu, G. Thakur, A. Dadlani, R. Patel, J. Orbulescu, J. D. Whyte, M. Micic, and R. M. Leblanc, "Surface chemistry and spectroscopy of human insulin Langmuir monolayer.," *The journal of physical chemistry. B*, vol. 116, pp. 10205–12, Aug. 2012.

[92] A. K. Attri, C. Fernández, and A. P. Minton, "pH-dependent self-association of zinc-free insulin characterized by concentration-gradient static light scattering.," *Biophysical chemistry*, vol. 148, pp. 28–33, May 2010.

[93] S. Hvidt, "Insulin association in neutral solutions studied by light scattering," *Biophysical chemistry*, vol. 39, pp. 205–213, 1991.

[94] R. D. Wampler, A. J. Moad, C. W. Moad, R. Heiland, and G. J. Simpson, "Visual methods for interpreting optical nonlinearity at the

molecular level.," *Accounts of chemical research*, vol. 40, pp. 953–60, Oct. 2007.

[95] A. J. Moad, C. W. Moad, J. M. Perry, R. D. Wampler, G. S. Goeken, N. J. Begue, T. Shen, R. Heiland, and G. J. Simpson, "NLOPredict: visualization and data analysis software for nonlinear optics.," *Journal of computational chemistry*, vol. 28, pp. 1996–2002, Sept. 2007.

[96] E. F. Pettersen, T. D. Goddard, C. C. Huang, G. S. Couch, D. M. Greenblatt, E. C. Meng, and T. E. Ferrin, "UCSF Chimera–a visualization system for exploratory research and analysis.," *Journal of computational chemistry*, vol. 25, pp. 1605–12, Oct. 2004.

[97] L. Fu, Z. Wang, and E. C. Y. Yan, "Chiral vibrational structures of proteins at interfaces probed by sum frequency generation spectroscopy.," *International journal of molecular sciences*, vol. 12, pp. 9404–25, Jan. 2011.

[98] C. Pinholt, R. A. Hartvig, N. J. Medlicott, and L. Jorgensen, "The importance of interfaces in protein drug delivery - why is protein adsorption of interest in pharmaceutical formulations?," *Expert opinion on drug delivery*, vol. 8, pp. 949–64, July 2011.

[99] D. Chandler, "Interfaces and the driving force of hydrophobic assembly.," *Nature*, vol. 437, pp. 640–7, Sept. 2005.

[100] L. Jorgensen, P. Bennedsen, S. V. Hoffmann, R. L. Krogh, C. Pinholt, M. Groenning, S. Hostrup, and J. T. Bukrinsky, "Adsorption of insulin with varying self-association profiles to a solid Teflon surface–influence on protein structure, fibrillation tendency and thermal stability.," *Eu-*

ropean journal of pharmaceutical sciences, vol. 42, pp. 509–16, Apr. 2011.

[101] J. S. Sharp, J. a. Forrest, and R. a. L. Jones, "Surface denaturation and amyloid fibril formation of insulin at model lipid-water interfaces.," *Biochemistry*, vol. 41, pp. 15810–9, Dec. 2002.

[102] M. I. Smith, J. S. Sharp, and C. J. Roberts, "Nucleation and growth of insulin fibrils in bulk solution and at hydrophobic polystyrene surfaces.," *Biophysical journal*, vol. 93, pp. 2143–51, Sept. 2007.

[103] S. H. Mollmann, U. Elofsson, J. T. Bukrinsky, and S. Frokjaer, "Displacement of Adsorbed Insulin by Tween 80 Monitored Using Total Internal Reflection Fluorescence and Ellipsometry," *Pharmaceutical Research*, vol. 22, pp. 1931–1941, Aug. 2005.

[104] S. H. Mollmann, L. Jorgensen, J. T. Bukrinsky, U. Elofsson, W. Norde, and S. Frokjaer, "Interfacial adsorption of insulin conformational changes and reversibility of adsorption.," *European journal of pharmaceutical sciences*, vol. 27, pp. 194–204, Feb. 2006.

[105] S. Hsieh, C.-W. Hsieh, H.-H. Chou, C.-W. Chang, and L.-Y. Chu, "Effect of the surface chemistry of insulin fibrils on the aggregation rate.," *ChemPhysChem : a European journal of chemical physics and physical chemistry*, vol. 15, pp. 76–9, Jan. 2014.

[106] D. Kurouski, T. Deckert-Gaudig, V. Deckert, and I. K. Lednev, "Structure and composition of insulin fibril surfaces probed by TERS.," *Journal of the American Chemical Society*, vol. 134, pp. 13323–9, Aug. 2012.

[107] Y. Yang, A. M. Bittner, S. Baldelli, and K. Kern, "Study of self-assembled triethoxysilane thin films made by casting neat reagents in

ambient atmosphere," *Thin Solid Films*, vol. 516, pp. 3948–3956, Apr. 2008.

[108] J. Brange, *Galenics of Insulin*. Springer-Verlag, 1987.

[109] C. Reichhart and C. Czeslik, "Native-like structure of proteins at a planar poly(acrylic acid) brush.," *Langmuir*, vol. 25, pp. 1047–53, Jan. 2009.

[110] H. S. Mansur, R. L. Oréfice, W. L. Vasconcelos, Z. P. Lobato, and L. J. C. Machado, "Biomaterial with chemically engineered surface for protein immobilization.," *Journal of materials science. Materials in medicine*, vol. 16, pp. 333–40, Apr. 2005.

[111] W. M. Lau, I. Bello, L. J. Huang, X. Feng, M. Vos, and I. V. Mitchell, "Argon incorporation in Si(100) by ion bombardment at 15-100 eV," *Journal of Applied Physics*, vol. 74, no. 12, p. 7101, 1993.

[112] W. Dzwolak, V. Smirnovas, R. Jansen, and R. Winter, "Insulin forms amyloid in a strain-dependent manner: An FT-IR spectroscopic study," *Protein science*, pp. 1927–1932, 2004.

[113] J. Kim, G. J. Holinga, and G. A. Somorjai, "Curing Induced Structural Reorganization and Enhanced Reactivity Observations of Two Types of Chemically and Structurally Unique amino groups on the surface," *Langmuir*, vol. 27, pp. 5171–5175, 2011.

[114] T. Z. Mengistu, V. Goel, J. H. Horton, and S. Morin, "Chemical force titrations of functionalized Si(111) surfaces.," *Langmuir*, vol. 22, pp. 5301–7, June 2006.

[115] G. C. Allen, F. Sorbello, C. Altavilla, A. Castorina, and E. Ciliberto, "Macro-, micro- and nano-investigations on 3-aminopropyltrimethoxysilane self-assembly-monolayers," *Thin Solid Films*, vol. 483, pp. 306–311, July 2005.

[116] Y. Xu, Y. Yan, D. Seeman, L. Sun, and P. L. Dubin, "Multimerization and aggregation of native-state insulin: effect of zinc.," *Langmuir*, vol. 28, pp. 579–86, Jan. 2011.

[117] J. M. Jones and T. Coffey, "Using the quartz crystal microbalance to study macro- and nanoscale bubbles at solid-liquid interfaces," *Journal of Physics: Conference Series*, vol. 100, p. 072026, Mar. 2008.

[118] M. H. Ahmed, J. A. Byrne, J. McLaughlin, and W. Ahmed, "Study of Human Serum Albumin Adsorption and Conformational Change on DLC and Silicon Doped DLC Using XPS and FTIR Spectroscopy," *Journal of Biomaterials and Nanobiotechnology*, vol. 04, no. 02, pp. 194–203, 2013.

[119] H. Schiffter, J. Condliffe, and S. Vonhoff, "Spray-freeze-drying of nanosuspensions: the manufacture of insulin particles for needle-free ballistic powder delivery.," *Journal of the Royal Society, Interface / the Royal Society*, vol. 7 Suppl 4, pp. S483–500, Aug. 2010.

[120] D. Harrison and C. Garrat, "The accurate measurement of Insulin Molarity," *Biochemical Journal*, pp. 733–734, May 1969.

[121] F. Evers, C. Reichhart, R. Steitz, M. Tolan, and C. Czeslik, "Probing adsorption and aggregation of insulin at a poly(acrylic acid) brush.," *Physical chemistry chemical physics : PCCP*, vol. 12, pp. 4375–82, May 2010.

[122] A. D. Quast, N. C. Wilde, S. S. Matthews, S. T. Maughan, S. L. Castle, and J. E. Patterson, "Improved assignment of vibrational modes in sum-frequency spectra in the CH stretch region for surface-bound C18 alkylsilanes," *Vibrational Spectroscopy*, vol. 61, pp. 17–24, July 2012.

[123] Y. Liu, L. K. Wolf, and M. C. Messmer, "A Study of Alkyl Chain Conformational Changes in Self-Assembled n -Octadecyltrichlorosilane Monolayers on Fused Silica Surfaces," *Langmuir*, vol. 17, pp. 4329–4335, July 2001.

[124] Z. Guo, W. Zheng, H. Hamoudi, C. Dablemont, V. Esaulov, and B. Bourguignon, "On the chain length dependence of CH3 vibrational mode relative intensities in sum frequency generation spectra of self assembled alkanethiols," *Surface Science*, vol. 602, pp. 3551–3559, Dec. 2008.

[125] Y. Tong, E. Tyrode, M. Osawa, N. Yoshida, T. Watanabe, A. Nakajima, and S. Ye, "Preferential adsorption of amino-terminated silane in a binary mixed self-assembled monolayer.," *Langmuir*, vol. 27, pp. 5420–6, May 2011.

[126] J. Wang, X. Chen, M. L. Clarke, and Z. Chen, "Vibrational spectroscopic studies on fibrinogen adsorption at polystyrene/protein solution interfaces: hydrophobic side chain and secondary structure changes.," *The journal of physical chemistry. B*, vol. 110, pp. 5017–24, Mar. 2006.

[127] J. Weber, T. Balgar, and E. Hasselbrink, "Conformational disorder in alkylsiloxane monolayers at elevated temperatures.," *The Journal of chemical physics*, vol. 139, p. 244902, Dec. 2013.

[128] V. V. Naik, M. Crobu, N. V. Venkataraman, and N. D. Spencer, "Multiple Transmission-Reflection IR Spectroscopy Shows that Surface Hydroxyls Play Only a Minor Role in Alkylsilane Monolayer Formation on Silica," *The Journal of Physical Chemistry Letters*, vol. 4, pp. 2745–2751, 2013.

[129] G. J. Holinga, R. L. York, R. M. Onorato, C. M. Thompson, N. E. Webb, A. P. Yoon, and G. A. Somorjai, "An SFG study of interfacial amino acids at the hydrophilic SiO2 and hydrophobic deuterated polystyrene surfaces.," *Journal of the American Chemical Society*, vol. 133, pp. 6243–53, Apr. 2011.

[130] D. Phillips, R. York, O. Mermut, K. McCrea, R. Ward, and G. A. Somorjai, "Side Chain, Chain Length, and Sequence Effects on Amphiphilic Peptide Adsorption at Hydrophobic and Hydrophilic Surfaces Studied by Sum-Frequency Generation Vibrational Spectroscopy and Quartz Crystal Microbalance," *Journal of Physical Chemistry C*, vol. 111, pp. 255–261, Jan. 2007.

[131] E. Bulard, Z. Guo, W. Zheng, H. Dubost, M.-P. Fontaine-Aupart, M.-N. Bellon-Fontaine, J.-M. Herry, R. Briandet, and B. Bourguignon, "Non-invasive vibrational SFG spectroscopy reveals that bacterial adhesion can alter the conformation of grafted "brush" chains on SAM.," *Langmuir*, vol. 27, pp. 4928–35, Apr. 2011.

[132] M. L. Clarke and Z. Chen, "Polymer surface reorientation after protein adsorption.," *Langmuir*, vol. 22, pp. 8627–30, Oct. 2006.

[133] A. Barth, "The infrared absorption of amino acid side chains.," *Progress in biophysics and molecular biology*, vol. 74, no. 3-5, pp. 141–73, 2000.

[134] A. Nikolic, S. Baud, S. Rauscher, and R. Pomès, "Molecular mechanism of β-sheet self-organization at water-hydrophobic interfaces.," *Proteins*, vol. 79, pp. 1–22, Jan. 2011.

[135] C. Bryant, D. Spencer, and A. Miller, "Acid stabilization of insulin," *Biochemistry*, vol. 32, pp. 8075–8082, 1993.

[136] M. F. M. Engel, C. C. VandenAkker, M. Schleeger, K. P. Velikov, G. H. Koenderink, and M. Bonn, "The polyphenol EGCG inhibits amyloid formation less efficiently at phospholipid interfaces than in bulk solution.," *Journal of the American Chemical Society*, vol. 134, pp. 14781–8, Sept. 2012.

[137] G. Zandomeneghi, M. R. H. Krebs, M. G. McCammon, and M. Fändrich, "FTIR reveals structural differences between native beta-sheet proteins and amyloid fibrils.," *Protein science*, vol. 13, pp. 3314–21, Dec. 2004.

[138] B. M. Bulheller, A. Rodger, and J. D. Hirst, "Circular and linear dichroism of proteins.," *Physical chemistry chemical physics : PCCP*, vol. 9, pp. 2020–35, May 2007.

[139] M. R. Hicks, J. Kowalski, and A. Rodger, "LD spectroscopy of natural and synthetic biomaterials.," *Chemical Society reviews*, vol. 39, pp. 3380–93, Sept. 2010.

[140] D. R. T. Zahn, S. D. Silaghi, C. Cobet, M. Friedrich, and N. Esser, "Spectroscopic ellipsometry and reflectance anisotropy spectroscopy

of biomolecular layers on silicon surfaces," *Physica Status Solidi (B)*, vol. 242, pp. 2671–2680, Nov. 2005.

[141] C. Du, G. Falini, S. Fermani, C. Abbott, and J. Moradian-Oldak, "Supramolecular assembly of amelogenin nanospheres into birefringent microribbons.," *Science*, vol. 307, pp. 1450–4, Mar. 2005.

[142] R. Forker and T. Fritz, "Optical differential reflectance spectroscopy of ultrathin epitaxial organic films.," *Physical chemistry chemical physics : PCCP*, vol. 11, pp. 2142–55, Apr. 2009.

[143] B. M. Bulheller and J. D. Hirst, "DichroCalc–circular and linear dichroism online.," *Bioinformatics*, vol. 25, pp. 539–40, Feb. 2009.

[144] F. Evers, C. Reichhart, R. Steitz, M. Tolan, and C. Czeslik, "Probing adsorption and aggregation of insulin at a poly(acrylic acid) brush.," *Physical chemistry chemical physics : PCCP*, vol. 12, pp. 4375–82, May 2010.

[145] H.-F. Wang, W. Gan, R. Lu, Y. Rao, and B.-H. Wu, "Quantitative spectral and orientational analysis in surface sum frequency generation vibrational spectroscopy (SFG-VS)," *International Reviews in Physical Chemistry*, vol. 24, pp. 191–256, Apr. 2005.

[146] K. M. Gough, "Theoretical analysis of molecular polarizabilities and polarizability derivatives in hydrocarbons," *The Journal of Chemical Physics*, vol. 91, no. 4, p. 2424, 1989.

[147] X. Wei, S.-C. Hong, X. Zhuang, T. Goto, and Y. R. Shen, "Nonlinear optical studies of liquid crystal alignment on a rubbed polyvinyl alcohol surface," *Phys. Rev. E*, vol. 62, pp. 5160–5172, Oct 2000.

[148] C. Bain, E. Troughton, and Y. Tao, "Formation of monolayer films by the spontaneous assembly of organic thiols from solution onto gold," *Journal of the American Chemical Society*, no. 111, pp. 321–335, 1989.

[149] J. E. Fulghum, R. Stokell, G. McGuire, B. Patnaik, N. Yu, Y. Zhao, and N. Parikh, "Determination of Si2p electron attenuation lengths in SiO2," *Journal of Electron Spectroscopy and Related Phenomena*, vol. 60, pp. 117–125, Oct. 1992.

[150] K. T. Nguyen, J. T. King, and Z. Chen, "Orientation determination of interfacial beta-sheet structures in situ.," *The journal of physical chemistry. B*, vol. 114, pp. 8291–300, July 2010.

[151] X. Chen, J. Wang, J. J. Sniadecki, M. a. Even, and Z. Chen, "Probing alpha-helical and beta-sheet structures of peptides at solid/liquid interfaces with SFG.," *Langmuir : the ACS journal of surfaces and colloids*, vol. 21, pp. 2662–4, Mar. 2005.

[152] L. Fu, J. Liu, and E. C. Y. Yan, "Chiral sum frequency generation spectroscopy for characterizing protein secondary structures at interfaces.," *Journal of the American Chemical Society*, vol. 133, pp. 8094–7, June 2011.

[153] W. DeGrado and J. Lear, "Induction of peptide conformation at apolar water interfaces. 1. A study with model peptides of defined hydrophobic periodicity," *Journal of the American Chemical Society*, no. 10, pp. 7684–7689, 1985.

[154] G. a. Somorjai, H. Frei, and J. Y. Park, "Advancing the frontiers in nanocatalysis, biointerfaces, and renewable energy conversion by in-

novations of surface techniques.," *Journal of the American Chemical Society*, vol. 131, pp. 16589–605, Nov. 2009.

[155] T. Weidner, N. T. Samuel, K. McCrea, L. J. Gamble, R. S. Ward, and D. G. Castner, "Assembly and structure of alpha-helical peptide films on hydrophobic fluorocarbon surfaces.," *Biointerphases*, vol. 5, pp. 9–16, Mar. 2010.

[156] T. Weidner, N. F. Breen, G. P. Drobny, and D. G. Castner, "Amide or amine: determining the origin of the 3300 cm(-1) NH mode in protein SFG spectra using 15N isotope labels.," *The journal of physical chemistry. B*, vol. 113, pp. 15423–6, Nov. 2009.

[157] T. Weidner, N. F. Breen, K. Li, G. P. Drobny, and D. G. Castner, "Sum frequency generation and solid-state NMR study of the structure, orientation, and dynamics of polystyrene-adsorbed peptides.," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 107, pp. 13288–93, July 2010.