UNIVERSITÀ DEGLI STUDI DI TRIESTE

XXXI CICLO DEL DOTTORATO DI RICERCA IN
NANOTECNOLOGIE

Tesi di dottorato in

Conformational characterization of functional proteins in physiological environment by Collective Enhancement IR Absorption (CEIRA) microscopy

Settore scientifico-disciplinare: FIS/07

DOTTORANDO
PAOLO ZUCCHIATTI

COORDINATORE
PROF. LUCIA PASQUATO

SUPERVISORE DI TESI
DR. LISA VACCARI

CO-SUPERVISORE DI TESI
PROF. ALESSANDRO BARALDI

ANNO ACCADEMICO 2017/2018

Abstract

The functionality of a protein is interwoven with its structure, small variations in the latter may result in the arise of resistances to drugs or severe pathologies. This is one of the main reasons why the study of protein structure, field of science called structural biology, is of paramount importance and requires a multidisciplinary/multi-technique approach to understand protein structure, function and dynamics. All these analytical tools have their pros and cons, and notably no one technique alone can provide all the answers to the question of structural biology. In this field, a pivotal role is played by those techniques that elucidate the secondary structure of proteins, such as Fourier Transform Infrared (FTIR) spectroscopy. However, despite its several advantages, FTIR dramatically suffers for sensitivity limitations.

The aim of my PhD is to overcome this drawback by exploiting Collective Enhanced Infrared Absorption (CEIRA) and apply it to the investigation of secondary structure of proteins of biological and biomedical relevance in aqueous environment and at ultralow concentrations.

At first the main efforts were devoted to the optimization of fabrication protocols for the of CEIRA plasmonic devices. These chips are made by ordered arrays of gold nanoantennas, whose dimensions can be tuned in order to give signal enhancement in different Regions of Interest (ROIs) in the Mid-IR. Then, by using model proteins, we proceeded with the implementation of efficient strategies for the measurements of protein solutions in buffer environment and at nanomolar concentrations, in both static and dynamic conditions.

Once everything was established, our objective was to study the structural modifications that the Epithelial Growth Factor Receptor, EGFR, and in particular its Kinase Domain, undergoes when in contact to a specific ligand: the tyrosine-kinase inhibitor (TKI), Lapatinib, validating the great potentialities of our technique for protein conformational analyses.

In conclusion, we developed a non-destructive label-free technology for *in vitro* detection of structural information of protein monolayers in their native states giving the possibility to follow the dynamics of the protein binding to the surface. Our CEIRA devices are a versatile tool for further understanding of the fundamental principles underlying the mechanism on the base of protein-ligand interaction. Moreover, we believe that a platform with this kind of sensibility will open new possibilities in the study of proteins of biomedical interest, in particular for those proteins involved in diseases where missense mutations induce local conformational changes and possible misfolding, giving rise to new possibilities in the understanding of disease-related conformational changes in proteins.

Table of contents

Introduction

The aim of my PhD is to exploit Collective Enhanced Infrared Absorption (CEIRA) microscopy for the investigation of secondary structure of proteins of biological and biomedical relevance in aqueous environment and at physiological concentrations. This goal has been achieved (i) by optimizing protocols for the fabrication of CEIRA plasmonic devices, made by ordered arrays of gold nanoantennas allowing the Mid-IR signal enhancement in different Regions of Interest (ROIs), tuned to the main IR absorption bands of proteins; (ii) by improving methodologies for the effective anchoring of proteins onto gold-nanonatennas; (iii) by implementing strategies for the efficient measurements of protein solutions in buffer environment and at physiological concentrations, in both static and dynamic conditions; (iv) by capitalizing on the acquired experience to assess the subtle secondary structure modifications undergone by Epidermal Growth Factor Receptor, EGFR, interacting with tyrosine-kinase inhibitor (TKI), Lapatinib.
The large majority of my PhD activity has been carried out at the Chemical and Life Sciences branch of SISSI (Synchrotron Infrared Source for Spectroscopy and Imaging) beamline at Elettra Sincrotrone Trieste, where I performed both CEIRA measurements and more conventional FTIR spectroscopy and microscopy analysis. However, due to the interdisciplinary nature of the PhD project, strong collaboration has been established with Plasmon Nanotechnology Research Laboratory at the Istituto Italiano di Tecnologia (IIT) in Genova under the supervision of Dr. Andrea Toma, where I had the opportunity to gain some experience on nanofabrication technologies and to fabricate on myself the plasmonic devices used for the PhD experiments.

The present manuscript summarizes the performed activities and the outcomes of my PhD project.

Chapter 1 is intended to provide the reader background information on the state of the art of structural characterization of proteins, with both standardized and more innovative methods. The structural characterization of proteins is of paramount importance for understanding the molecular mechanisms which tune their activity and, consequently, their

proper function. X-ray crystallography plays the main role in the 3D characterization of proteins at atomic resolutions, while, for the investigation of protein structure and dynamics under nearly physiological conditions, Nuclear Magnetic Resonance (NMR) spectroscopy is the methodology of choice. However, severe limitations to protein structural characterization are imposed by proteins that are challenging to crystallize (i.e. membrane proteins), or available in limited quantities, (i.e. mutated proteins involved in rare diseases). In the last decades, different techniques have been developed and optimized for protein conformational analysis, such as Small Angle X-ray scattering (SAXS) or cryo-Electron Microscopy (cryo-EM). However, all these analytical tools have their pros and cons, and notably no one technique alone can provide all the answers to the questions of structural biology (*section 1.1*). This has motivated researchers to combine different techniques to understand protein structure, function and dynamics. Therefore, techniques which can provide complementary information to the aforementioned structural methods are of great interest for integrated structural biology studies. In this field, a pivotal role is played by those techniques that elucidate the secondary structure of proteins, such as Fourier Transform Infrared (FTIR) spectroscopy. Specifically, Mid-infrared spectroscopy is a widely employed tool for protein secondary structure analyses, since it is label-free and it is particularly sensitive to little protein conformational changes. Despite its potential, IR measurements on small quantities of molecules using conventional approaches are extremely challenging due to limitations in sensitivity of the technique: for diluted solutions (micromolar and sub-micromolar concentrations) and thin samples (such as monolayers), the IR signal becomes prohibitively weak (*section 2.2*).

A challenging method for increasing FTIR sensitivity is to take advantage from the electric field enhancement associated with the excitation of plasmonic resonances when light is incident on *ad-hoc* shaped metallic structures. This method is called Surface Enhanced InfraRed Absorption (SEIRA). “Conventional” SEIRA substrates, usually made by roughened gold or silver islands, provide only a modest Mid-IR signal enhancement, in the range of 1-2 orders of magnitude. The exploitation of nanofabrication technologies and electron beam lithography, in particular, allows the fabrication of controlled nanostructures that can produce enhancement factors of several orders of magnitude, for the so-called “resonance SEIRA”. Comparable enhancements can be obtained for nanoantennas arrays through the mutual tuning of the single antenna parameters (shape and material) and array lattice parameters (periodicity), making Collective Enhanced IR Absorption (CEIRA)

suitable for protein structural characterization at very low concentration, up to nanomolar detection sensitivity (section 1.3). We would like to mention at this stage that values reported in literature concerning the magnitude enhancement of plasmonic devices, and the consequent sensitivity should be carefully assessed, since strongly influenced by the protein deposition or anchoring method, that could lead to the formation of protein layers also outside the nanoantenna "hot-spots". Section 3.1 of this thesis draw out on this issue.

Chapter 2. The second chapter of the thesis concentrates on the strategy adopted for optimizing the CEIRA devices that has been used for accomplish PhD project. The simulated design and the fabrication procedure followed for producing the CEIRA substrates will be presented in *section 2.1* and *section 2.2* respectively. CEIRA devices have been produced during the three years of PhD activity with two different geometries, namely *line-to-line* and *cross*-geometry. The characterizations of the morphology of the two forms of devices and of their infrared response will be reported in *section 2.3* and *section 2.4* respectively. Finally, the interpretation of the enhanced protein signals will be discussed in *section 2.5*.

Chapter 3. Chapter 3 reviews the experiments done at the Chemical and Life Sciences branch of SISSI beamline at Elettra with CEIRA substrates for protein conformational analysis. The chapter is divided in five main sections. In *section 3.1* will be described the strategies that have been tested for anchoring the protein target molecules at the hot-spots of the CEIRA arrays, validating their performances by using *line-to-line* devices. In the second section (*section 3.2*) the improved performances of second generation *cross*-geometry CEIRA devices with both conventional and infrared synchrotron radiation (IR-SR) will be examined. *Section 3.3* will report the main results achieved in the secondary structure analysis of thin dried layers of model proteins, Bovine Serum Albumin (BSA) and Concanavalin A (ConA). S*ection 3.4* will focus on the difficulties being addressed in protein secondary structure assessment in aqueous buffer conditions. Plasmonic Internal Reflection (PIR) approach, developed by Adato and co-workers, will be introduced, and the results of its exploitation for hydrated BSA and ConA conformational analysis will be discussed. Finally, in *section 3.5* the use of PIR-CEIRA technique will be considered for the secondary structure characterization of a protein of biomedical interest at ultra-low concentrations. Static and dynamic experiments aimed to investigate both wild-type and mutate form of EGFR, and to assess the minimal conformational changes undergone by EGFR, interacting Lapatinib will be described.

The outcomes of the PhD thesis will be then summarized in **Chapter 4**, **Conclusions**, aiming attention at the points of novelty and perspectives applications of the strategies developed during the PhD thesis.

The present manuscript incudes also an appendix, **Appendix A**, which summarizes the main results obtained during these three years on a side project on biospectroscopy, developed in collaboration with the group of inelastic ultraviolet scattering (IUVS) beamline at Elettra Sincrotrone Trieste. This minority activity was focused onto the development of two different tailored protocols which enables to obtain Nucleic Acids solutions extracted from cell cultures with a spectroscopic quality satisfactory for routine FTIR and UV Raman analysis.

Chapter 1

This chapter will introduce the reader to the context of protein structure analysis. After a general introduction to protein structure, the first part will provide an overview on the principal techniques devoted to their structural analysis, highlighting potentials and limitations of each technique (*section 1.1*). In the second section (*section 1.2*), we will focus on the use of conventional infrared spectroscopy for protein analyses, explaining the basic principles of Fourier transform infrared (FTIR) spectroscopy, paying particular attention to the advantages offered by this technique for conformational analyses of biomolecules, but also to the technique limitations. The third part (*section 1.*3) will contain a short overview on plasmonic optics, focusing then on the surface-enhanced infrared absorption (SEIRA), which underpins nanoantenna-assisted plasmonic effect that will be used to enhance the vibrational modes of proteins, allowing conformational studies with few attomoles of sample.

1.1 Structural analyses of proteins

1.1.1 Protein structure

Proteins are the most abundant and versatile biological macromolecules in the cell, and they constitute most of its dry mass. They are cellular building blocks and execute nearly all the cell's functions: they are involved in catalysis, transport, immune response and mechanical support. They also provide cellular motility, transmit nerve impulses and promote cell replication, controlling growing and differentiation [1], [2].

Proteins are organic compounds composed of long chains of amino acids folded into complex structures. There are 20 types of amino acids in proteins, each with different chemical properties. Amino acids sub-units are made by the same core element: onto the so-called Cα are bound a carboxylic group, an amino group, a hydrogen and a side chain which is different for each amino acid and which governs the chemical property of the monomer (Figure 1.1a). In the macromolecule, amino acids monomers are linked by covalent peptide bonds as shown in Figure 1.1b.

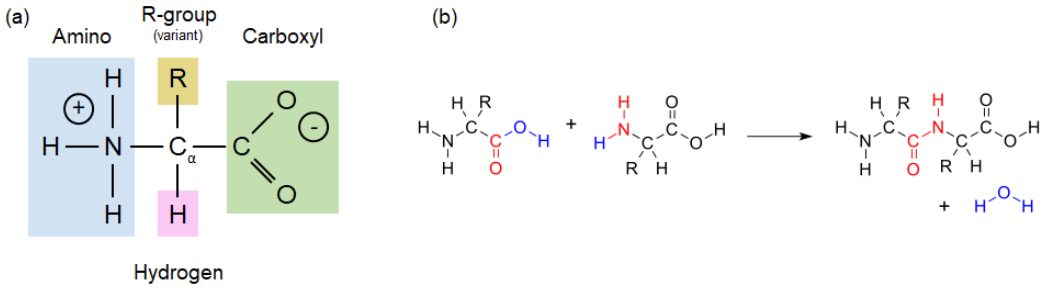

Figure 1.1: (a) Chemical structure of a representative amino acid in its zwitterionic state. (b) The peptide bond: the condensation of the carboxylic group with the amino group of two neighboring amino acids results in the formation of a covalent amidic bond.

The repetitive sequence of atoms along the core of the polypeptide chain is the polypeptide backbone, while those portions of the amino acids that are not involved in making a peptide bond are the amino acid side chains (Figure 1.2). The sequence of the amino acids, their nature and number, is unique for each protein and determine the so-called protein primary structure.

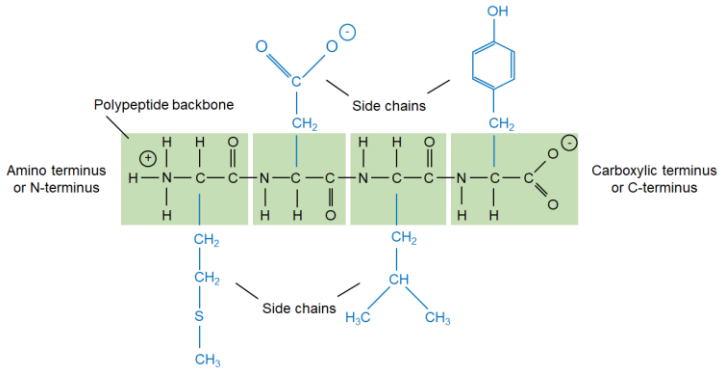

Figure 1.2: Structural components of a protein: a protein consists of a polypeptide backbone "decorated" with side chains. Image adapted from [2].

The two ends of a polypeptide chain are chemically different: the end carrying the free amino group (in its neutral, NH_2, or protonated form, NH_3^+) is the amino terminus (or N-terminus), and that carrying the free carboxyl/carboxylate group ($COOH/COO^-$) is the carboxyl terminus (or C-terminus). The amino acid sequence of a protein is always presented in the N-to-C direction, reading from left to right. However, proteins are not simple linear molecules, but they have highly complex folding patterns which depend on the amino acid sequence and on environmental influences [3], [4]. Protein folding is one of the fundamental examples of self-assembly in nature and it is determined by different sets of weak noncovalent bonds. Three types of noncovalent bonds are involved: hydrogen bonds, ionic bonds, and van der Waals interaction. As a result of all of these interactions, each type of

protein has a particular three-dimensional structure. The final folded structure adopted by any polypeptide chain is generally the one in which the free energy is minimized [2].

Conventionally, biologists distinguish four levels of organization in protein structure. The amino acid linear sequence is considered the primary structure, as already mentioned, while the stretches of the polypeptide resulting from intra-chain interactions constitute the protein secondary structure. Such interactions arise from hydrogen-bonding between the N–H and C=O groups in the polypeptide backbone. The two main recognized classes of secondary structures are the α helix, where the backbone is arranged in a spiral-like coil, and the β sheet, where the polypeptide backbone is arranged side by side [5] (see Figure 1.3a). An α helix is generated when a single polypeptide chain turns around on itself to form a rigid cylinder. In this conformation, a hydrogen bond is made between every fourth peptide bonds, linking the C=O of one peptide bond to the N–H of the other, giving rise to a regular helix with a complete turn every 3.6 amino acids. Other transient helix structures have been found, the most representative of which is the 3^{10} helix (nearly 10–15% of all helices in protein secondary structures) [6], [7]. Such structures differ from α helix for the backbone of the hydrogen-bonding network and usually are made by no more than 7–11 residues [7]. β sheets can form either from adjacent chains that run in the same direction (parallel chains) or from strands that folds back and forth upon itself, with sections running in opposite directions (antiparallel chains). Both types of β sheet produce a very rigid structure. The helix and sheet structures are joined with connecting loops, known as random coils, that do not fall into the previous defined categories. These secondary structural elements are represented in ribbon models of proteins as shown in Figure 1.3b.

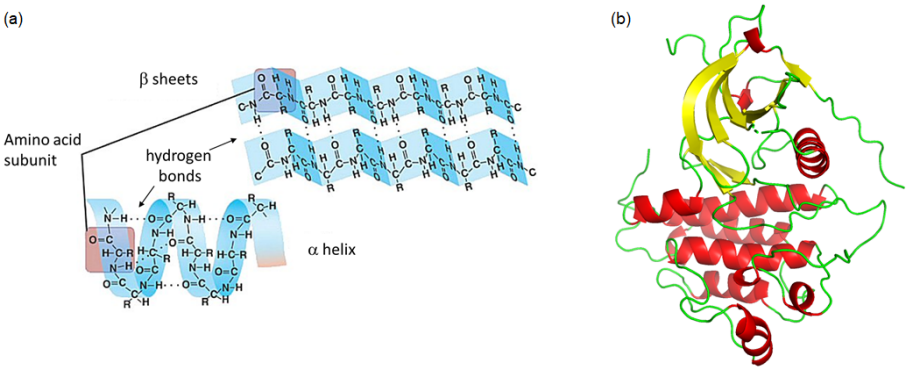

Figure 1.3: (a) Regular conformation of the polypeptide backbone in the α helix and the β sheet stretches, highlighting the hydrogen bonding network. Image adapted from [2]. (b) Ribbon model of the kinase domain of Epidermal Growth Factor Receptor (EGRF), from [8]. The main secondary structural elements of proteins are showed with differently colored graphical elements: red ribbons – α helix; yellow arrows– β sheet; green line– random coils.

The organization of all the secondary structural elements in a full three-dimensional is described as the tertiary structure of the protein and defines it (see ribbon model in Figure 1.3b). Hydrogen bonding and ionic interactions between side chains mostly influence protein secondary structure. Instead, hydrophobic force is an essential contributor that drives the protein 3D folding. Generally, the nonpolar side chains in a protein tend to cluster inside the macromolecule, in order to limit the contact with the surrounding polar environment (namely, cellular water). On the contrary, polar side chains have the tendency to arrange themselves near the outside of the molecule, where the formation of hydrogen bonds with water and with other polar molecules is allowed. Notably, many proteins are either attached to the outside of the cell membrane or secreted as part of the extracellular matrix, and are directly exposed to extracellular conditions. In order to maintain their structures in such environment, the polypeptide chains are often stabilized by covalent cross-linkages. The most common cross-linkage in proteins is the covalent sulfur–sulfur bond, which forms between two pairs of –SH groups of cysteine side chains that are adjacent in the folded protein, and act as atomic staples to reinforce the protein most favored conformation. It is noteworthy that disulfide bonds generally fail to form in the cell cytosol, where a high concentration of reducing agents converts S–S bonds back to cysteine –SH groups [2].

When different individual protein macromolecules, called subunits, are arranged into a singular functional complex, the complete structure is defined as the quaternary structure [9]. Biologist recognized another unit of organization distinct from the four just described, called *protein domain* [10], which is a substructure produced by any part of a polypeptide chain that can fold independently into a compact, functional, stable structure [11]. The diverse domains of a protein are often associated with different functions. For instance, protein kinases-like superfamily (PK-like), a class of protein effectors essential for the maintenance of cell cycle control, are characterized by a homologous catalytic domain, called Kinase Domain (KD) [12]. Such domain is commonly made by 250-300 amino acid residues and it contains the active site of the protein, by which PKs control the activation of target proteins thanks to their catalytic activity [13].

1.1.2 Conventional techniques for protein structural analyses

It is a well-established paradigm that the function of a protein is related to its structure. Therefore, structural characterization of proteins is of paramount importance for understanding the molecular mechanisms which define their activity. Such analyses are used

also to recognize the molecular basis of diseases, leading, for instance, to the development of novel drugs [14]–[17]. Structural biology is the branch of life science which focuses on this topic [14]. The two principal techniques employed for the 3D characterization of proteins at atomic resolutions are X-ray crystallography [18]–[20] and Nuclear Magnetic Resonance (NMR) spectroscopy [21]–[23].

X-ray crystallography is the most powerful analytical technique for structural biologists, and it is the key tool for providing high resolution protein 3D structure [18] having been used for about the 90% of the total protein structures nowadays deposited in the Protein Data Bank (PDB) [24]. By using this technique, the structure of a protein can be reconstructed from its X-ray diffraction pattern [25], obtained from an ordered crystal of the protein. A scheme of the workflow for protein structure determination using X-ray crystallography is reported in Figure 1.4. By X-ray crystallography, the entire three-dimensional structure of a protein (tertiary structure in case of single polypeptides, quaternary structure in case of multiple polypeptides) can be resolved at very high resolution, down to 0.5 Å [18], [26].

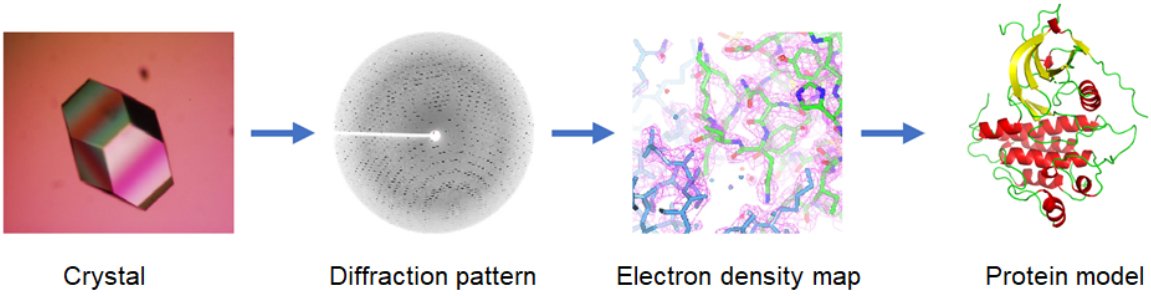

Figure 1.4: Workflow for protein structure determination using X-ray crystallography. When proteins crystals are illuminated with X-ray radiation in conditions satisfying the Bragg's law [18], light is diffracted due to the ordered arrangement of atoms within the crystal. The repetitive nature of structural units within a protein crystal produces a diffraction pattern, formed by the interference of the diffracted X-ray beams, from which is possible to obtain the 3D image of the electron density within the crystal, getting the position of atoms and chemical bonds within the structure to be determined. Image adapted from [27].

X-ray crystallography is widely employed also for visualizing the interactions between the protein and exogenous molecules [19], both natural ligands [20] or synthetic drugs such as enzyme inhibitors [28]. This is possible after co-crystallization of proteins with their target molecules [19], providing accurate and highly informative structural data on the mechanisms of such interactions. Despite its great potentials for 3D structural determination and drug discovery, X-ray crystallography suffers from some intrinsic drawbacks mostly related to the sample preparation. Making well-diffracting crystals is mandatory for performing such analyses [18], and crystal preparation is often time consuming and expensive in terms of

protein and laboratory consumables. Indeed, it is the major bottleneck in protein structural determination [29], [30]. Not all proteins can be easily produced in such ordered crystals. Although it is not possible *a priori* to recognize problem cases, it is known that some classes of protein are hard to crystallize [31], such as membrane proteins [32], protein-ligand complexes with ligands having low affinity [33] and large protein aggregates. In particular, for membrane proteins the use of detergents, required for protein isolation from the cellular membrane, induces the formation of micelles around the protein hydrophobic regions, significantly reducing the possibility of forming crystals [31]. Also, even when crystals can be obtained, they are often small and diffract poorly. Moreover, it is a matter of fact that proteins in crystals are not in their physiological conditions and their movement is limited, thus not allowing to obtain information on protein dynamic behavior.

The second principal technique for protein 3D structural studies at high resolution is Nuclear Magnetic Resonance (NMR) [21], [22]. This technique offers several advantages, since it does not require to prepare protein in crystalline form for being analyzed and it can work either in solid state [22] or in solution [23]. NMR spectroscopy is the only method that allows the determination of the 3D structure of protein molecules at atomic resolution in the liquid phase, and to follow protein dynamic behavior [1]. Despite these clear advantages for protein structural determination, its application is usually limited to small proteins (less than 30 kDa), due to the technique intrinsic limitations and the complexity of the spectral data provided [21]. Moreover, usually, expensive media are required, because of the cost of isotopes used for protein preparation [23]. Since only certain isotopes are capable of intense NMR signals, proteins require to be labeled with such isotopes, such as Hydrogen-2 (^{2}H), Carbon-13 (^{13}C) and Nitrogen-15 (^{15}N) [21]. Moreover, due to the inherent insensitivity of the technique, usually high concentrated solutions (about 1 mM) are required for obtaining intense profiles.

In the last decades, structural biology had a substantial progress in terms of available instrumentation, techniques and protocols allowing to get further insight of larger molecules and multi-protein complexes [34]. In this context, solution-based small- and wide-angle X-ray scattering (SAXS/WAXS) are other possible methods for observing protein structure and dynamics [35]. Such techniques are not highly resolved, but they can provide a way to explore the full extent of the conformational landscape as a function of a wide range of solution conditions [36]. More significantly, solution scattering represents a straightforward way to perform time-resolved studies [37]. Such analyses are employed in particular for

disordered systems. However, SAXS cannot provide structural details with high resolutions and require protein concentration in the range of mg/mL (in the order of tens of μM).

While NMR and X-ray crystallography are limited to structural studies of certain classes of proteins, most of the time small ones, cryo-Electron Microscopy (cryo-EM) is an emerging tool for the structural determination of large proteins or multi-protein complexes [38]. Furthermore, recent developments in cryo-EM combined with image analysis provided unique information on the conformational variability of macromolecular complexes [39]. However, the biochemical quality of the sample remains the major bottleneck to routine cryo-EM-based determination of structures at very high resolution (less than 10 Å) [40]. In addition, this technique is difficult to apply to molecules smaller than approximately 50 kDa, a cut-off larger than the average size of cellular proteins [41], while it performs better for proteins in the range of 160-300 kDa [42].

As seen, all these analytical tools have their pros and cons, and notably the requests of structural biology cannot be satisfied by a single technique. Moreover, they cannot be applied to all kinds of proteins and systems. This has motivated researchers to combine different techniques in order to fully understand protein structure, its function and dynamics [34]. Thus, methods offering complementary information to the aforementioned conventional methods are of great interest for integrated structural biology studies. In this field, a pivotal role is played by those techniques that can provide valuable insight to secondary structure analysis of proteins, such as Circular dichroism (CD) [43]–[45] and Fourier Transform Infrared (FTIR) spectroscopy [46]–[49]. These are the tools most commonly applied as complementary to X-ray crystallography and are particularly useful for analyzing proteins in a variety of environments, not being limited to the study of crystalline proteins. CD is generally employed to determine whether an expressed and purified protein is folded, or if a mutation affects its conformation, stability [50] and for folding thermodynamics [51]. Protein CD spectroscopy is also used for secondary structure estimation [50], but secondary structure content cannot be accurately calculated from CD spectra for the majority of proteins [44]. In fact, a reliable secondary structure content can be calculated only from CD spectra of all-α proteins and all-β proteins [44]. On the other hand, Fourier Transform InfraRed (FTIR) spectroscopy [52] is a widely used technique for analyzing the protein structure [49], since it is label free, not damaging and sensitive to subtle protein conformational changes [43], [53]. Due to these advantages, FTIR spectroscopy is especially useful for probing proteins that are difficult to be measured by other spectroscopic

techniques. In the next section, the basic principles for FTIR protein measures and structural analysis are presented.

1.2 FTIR spectroscopy for protein analysis

This section aims to provide a brief introduction to FTIR spectroscopy and underline the basic principles of the technique. Different sampling methods are discussed and their application for protein conformational studies is presented.

1.2.1 Infrared spectroscopy basic principles

In general, spectroscopy deals with the interaction of electromagnetic (EM) radiation and matter. It is based on the analysis of the energy of the transmitted or reflected light after illuminating a sample with a definite radiation source. IR light was discovered for the first time by Frederick William Herschel (1738-1822) basing on the studies of Sir Isaac Newton (1643-1727). In 1800, Herschel performed a famous experiment where, by promoting the dispersion of light with a prism, he tried to measure the temperature of different colors of the visible spectrum by placing a thermometer on each color. In this experiment he discovered that the warmest part of the spectrum was where there was no color at all, below the red. Therefore, this hot light became known as InfraRed (IR). In March of the same year, Herschel made another discovery: by placing a sample in the path of the IR light, he observed that by changing the part of the spectrum that passed through the sample, in some points the temperature rapidly decreased, concluding that the decreasing in temperature was consequent to the absorption of IR light from the sample. On this evidence, he defined IR spectroscopy as the "measurement of light absorption in the infrared", and this definition is still perfectly correct [54].

Actually, the InfraRed (IR) radiation covers a wide range of the EM spectrum, extending from visible to microwaves (Figure 1.5). Photons in this range have energies which span from 1.25 meV to 1.7 eV. Such energies cannot generally excite electronic transition, but they can induce vibrational excitation of covalently bonded atoms and groups.

The IR range is conventionally divided into three spectral regimes, named accordingly to their distance from visible light:

- ✓ near-infrared regime (NIR): 14000–4000 cm^{-1} (0.8–2.5 μm);
- ✓ mid-infrared regime (Mid-IR or MIR): 4000–400 cm^{-1} (2.5–25 μm);
- ✓ far-infrared regime (FIR): 400–10 cm^{-1} (25–1000 μm).

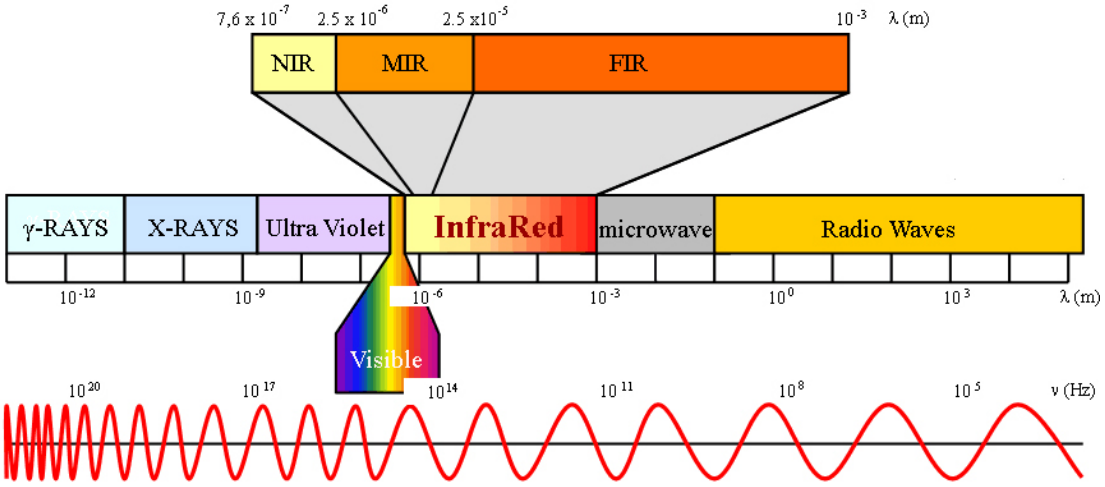

Figure 1.5: Electromagnetic Spectrum of light. The IR range and its sub-regions are highlighted.

In this thesis we will focus on the Mid-IR regime, which is the most useful IR region for analyzing organic compounds, since their transition energies are from about 44 meV to 440 meV (1-10 Kcal/mol), and therefore in the same range of the vibrational energies of covalent bonds [55].

Covalent bonds in molecules are usually simplified as rigid sticks or rods, but actually, they behave more like stiff springs that can be stretched and bent. At temperatures above 0 K, all the atoms in a molecule are continuously vibrating with respect to each other in a variety of ways. IR photons with appropriate energy can be absorbed by the molecule, inducing transitions between vibrational energy states. The frequency of the transition depends on different factors, the main of which is the type of vibration, the bond strength and the masses of the involved atoms. For instance, in order to stretch (or compress) a bond, higher energies are required than to bend it. Accordingly, the frequency that characterizes the stretching vibration of a given bond is higher than the one of the bending modes.

The stretching frequency of a diatomic molecule can be approximated by the Hooke's Law: Hooke's model treats two atoms and their connecting bond as a simple harmonic oscillator composed of two masses joined by a spring. According to Hooke's law, the frequency of the vibration of the spring (ν) depends on its force constant and the masses involved, as summarized in the Eq. 1.1:

$$\nu = \frac{1}{2\pi}\sqrt{\frac{k}{\mu}} \qquad \text{[Eq. 1.1]}$$

where ν is the frequency, *k* the spring constant of the bond and μ the reduced mass of the system, which, in the case of a diatomic molecule, is calculated as follows [Eq. 1.2]:

$$\mu = \frac{m_1 m_2}{m_1 + m_2} \quad \text{[Eq. 1.2]}$$

where m_1 and m_2 are the masses of the two atoms of the diatomic molecule.

As can be seen, equation 1.1 provides a link between the strength of the bond among two atoms (or molecular fragments), the mass of the interacting atoms (or molecular fragments) and the frequency of a vibration. Consequently, an increase in bond strength leads to equivalent frequency increase, while an increase in the atom mass corresponds to frequency decrease. The potential energy (E) of a simple harmonic oscillator is given by Equation 1.3:

$$E = \frac{1}{2} kx^2 = h\nu \quad \text{[Eq. 1.3]}$$

where *x* represents the displacement of the spring from its equilibrium position and *h* is the Plank's constant.

However, at the atomic scale, molecular systems cannot assume a continuous energy profile as predicted by the classical "balls-and-spring" model [Eq. 1.3]. Indeed, quantum mechanics [56] describes a diatomic molecule in terms of discrete energy levels E_n, defined as in Eq. 1.4.

$$E_n = \left(n + \frac{1}{2}\right) h\nu \quad \text{[Eq. 1.4]}$$

where ν is the frequency, *n* the vibrational quantum number (0, 1, 2, 3, . . .) and *h* the Plank's constant. Furthermore, for harmonic potentials, only transitions between neighboring energy levels that determine a net change in the dipole moment are allowed ($\Delta n = \pm 1$, selection rule).

In Figure 1.6a, a schematic representation of the potential energy for the harmonic model is reported. Despite this simple model is in good agreement with experimental data for transitions between the ground state of a molecule (n=0) and the first vibrational quantum level, it fails for transitions extending beyond. For a more accurate modeling, limits in bond compression and extension need to be considered. All the deviations from the simple harmonic oscillator model bring actually to a more appropriate anharmonic oscillator model, as shown in Figure 1.6b. This model account also for overtone ($\Delta n = \pm n$, with n integer and $n > 1$) and combination ($\sim \nu_n + \nu_m$) bands often detected in the NIR for real molecules.

Let's now move from the "balls-and-spring" model, which defines the basis for the simple vibration of a diatomic molecule, to models describing molecules made by a number of

atoms (N) higher than two. For such complex systems, there is a wide number of possible vibrations and all of them can be described as the combination of the motions of each atom in the 3D space. Each atom in a molecule has three degrees of freedom, corresponding to motions along each of the three Cartesian axes (x, y, and z). Consequently, a polyatomic molecule of N atoms has 3N total degrees of freedom. Three degrees of freedom are required to describe the translation motion of the entire molecule in space and three additional describe the whole molecule rotation. The remaining 3N – 6 degrees of freedom are the fundamental vibrations for nonlinear molecules, each of them describing the motion of all its atoms at the same frequency and with a fixed phase relation. Linear molecules possess 3N – 5 fundamental vibrational modes because only 2 degrees of freedom are sufficient to describe the entire molecule rotation. However, among the 3N – 6 (or 3N – 5) fundamental vibrations (also known as normal modes of vibration), only those that produce a net change in the dipole moment of the molecule are IR active [57].

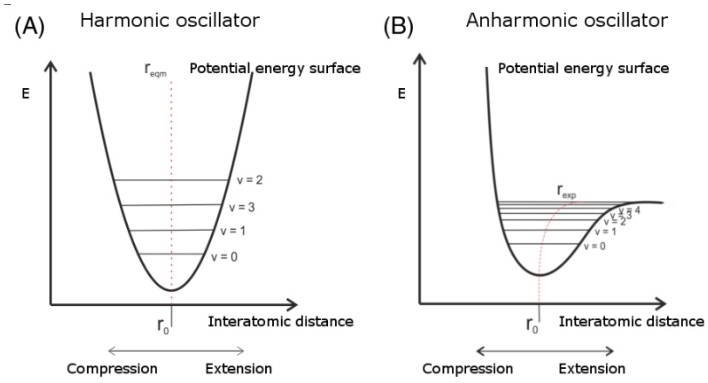

Figure 1.6: Representation of harmonic (a) and anharmonic (b) models of the vibrational energy levels of a diatomic molecule. r_0 is the equilibrium distance that minimize the Potential Energy (E). Adapted from [58].

According to the normal mode description, the vibrations of the two (or more) bonds are coupled. As one bond contracts, the other bond can either contract (in-phase) or expand (out-of-phase) in a vibration known as stretching mode. The stretching is defined symmetric (ν_s) when the bonds stretch in-phase, while it is defined asymmetric when it is out-of-phase (ν_{as}). Vibrations involving a change in bond angle are known as bending. The bending vibration that determines a variation of the angle between two bonds is called scissoring or deformation. Rocking, wagging and twisting are bending vibrations that produce a change in angle between the plane defined by the group of atoms involved in the vibrational mode and the rest of the molecule. Some molecules with a preferred orientation cannot freely rotate. They can however turn slightly back and forth with respected to the fixed orientation

in a motion known as libration. This is the case of liquid water, where the network of hydrogen bonding imposes a preferential orientation to the molecules. In order to better understand these concepts, in Figure 1.7 the vibrational modes of water are shown.

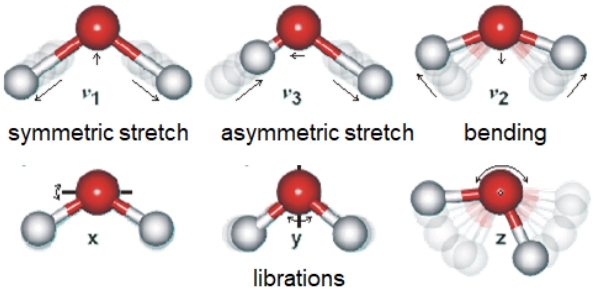

Figure 1.7: Vibrational modes of water. Image adapted from [59].

The number of experimentally observed IR absorption bands is generally different from the total number of active fundamental vibrations, since more than one mode of vibration can be characterized by the same frequency. Conversely, additional bands can be seen. These are overtones, which are integral multiples of the fundamental absorption frequencies, and combination bands, originating from the combination, either sum or difference, of fundamental frequencies.

Summarizing, when IR light hits a sample, absorption occurs if the frequency of the IR photons matches the one of a mode of vibration. The frequencies at which absorption of IR radiation occurs are directly correlated with the types of chemical moieties in the analyzed sample, and information about the molecular architecture of the sample can be retrieved. The superimposition of all the IR absorption bands generates a unique IR spectrum for each compound which can be considered as the fingerprint of the molecule.

1.2.2 Fourier Transform InfraRed (FTIR) spectroscopy

First IR experiments were made with the so called “dispersive spectrophotometers”, which exploited dispersive optical elements in order to monochromatize either illuminating or transmitted IR light. Conversely, modern instrumentation is based on interferometric systems. In contrast to dispersive spectrophotometers, Fourier Transform InfraRed (FTIR) interferometers perform the spectral splitting with the help of an interferometer, allowing wavelength dependent radiation modulation. The simplest interferometer is the Michelson one, shown in Figure 1.8. In the figure caption, a brief description of the working principle of the Michelson interferometer is also given.

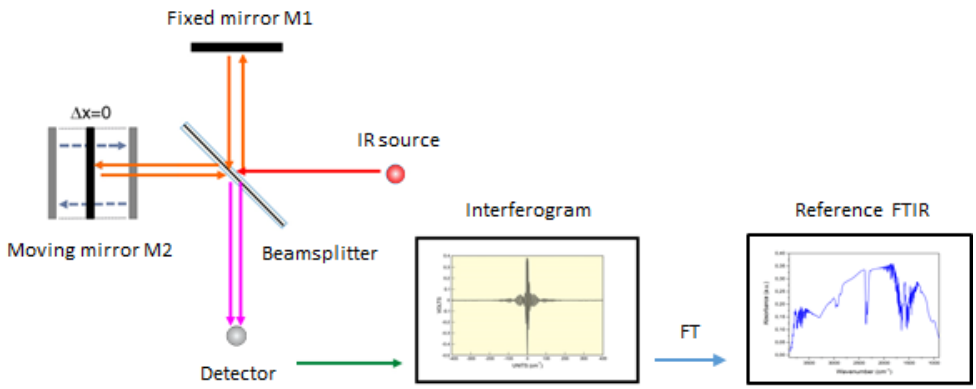

Figure. 1.8: Scheme of Michelson interferometer. Light coming from the IR source is partially transmitted and partially reflected by the beamsplitter, illuminating both M1 (fixed) and M2 (moving) mirrors. The shift in position of M2 generates a difference in the optical path modulated by the velocity of the mirror shifting. Therefore, for each wavelength of the source, constructive/destructive interference conditions are created over time. This generates and interferogram as a function of the movable mirror displacement from the zero-path-difference position (x, [cm]), characterized by an intense central burst for a continuous polychromatic source. The interferogram, recorded by the detector, is converted by Fourier Transformation to a spectrum in the x^{-1} space, also known as wavenumber, $\bar{\upsilon}$ [cm^{-1}].

Among the advantages provide by FTIR interferometers with respect to dispersive systems, we should highlight the faster data acquisition and the improved signal to noise (S/N) ratio by a factor of √M (where M is the number of sampled points comprising the spectrum), both deriving from the simultaneous observation of all the wavelengths at once. This advantage, known as "multiplex advantage" is one of the reason why the widespread diffusion of infrared spectroscopy is consequent to the development of FTIR interferometers.

1.2.3 FTIR Sampling techniques

Whether a sample is in the optical path between the beamsplitter and the detector (see Figure 1.8), the interferogram of the sample is collected, and the transmission spectrum, T [%], of the samples can be retrieved as the ratio between its spectrum to the spectrum of the background or reference. Alternatively, the spectrum can be plot as $-\log_{10}(T)$, called Absorbance (A) which is expressed in arbitrary unit [a.u.].

A wide number of sampling techniques are available for performing FTIR measurements. This versatility in sampling allows the technique to be applied to solids, liquids and gases making it a powerful tool for a label-free identification of molecular species [60]. In this paragraph, the focus will be on the techniques functional for the understanding of the measurements done during this thesis, namely transmission and Attenuated Total Reflectance (ATR) of liquid samples and FTIR microscopy in the Mid-IR.

(i) Transmission FTIR spectroscopy

The most common FTIR sampling method of FTIR is the transmission, where IR light passes through the sample before detection of its attenuation. The extent of attenuation obeys to Beer-Lambert law [Eq. 1.5]:

$$A = \varepsilon bC \qquad \text{[Eq. 1.5]}$$

Where the absorption intensity, A [a.u.], is proportional to the molar concentration of the sample, C [M], and the optical path, b [cm], through a constant known as the molar extinction coefficient, ε [$cm^{-1}M^{-1}$]. According to the molar extinction coefficient of the specific band of sample under investigation, the sample concentration and optical path need to be accorded in order to avoid the total absorption of light by the sample, i.e. to reach the regime of non-linear detector response. Considering liquid samples, they are usually confined within two IR transparent windows (commonly Calcium Fluoride, Barium Fluoride, or Zinc Selenide in the Mid-IR) spaced apart by a spacer, either commercial or fabricated [61]–[63]. The path lengths are typically in the range of few microns between the two IR transparent windows [60], and this constraint is particularly severe in the case of aqueous samples (see below in section 1.2.4).

(ii) Attenuated total reflection (ATR)-FTIR spectroscopy

Attenuated Total Reflection (ATR) [46], [53] is a sampling method that overcomes the path length limitations of conventional transmission measurements. The basic principle of ATR technique is illustrated in Figure 1.9.

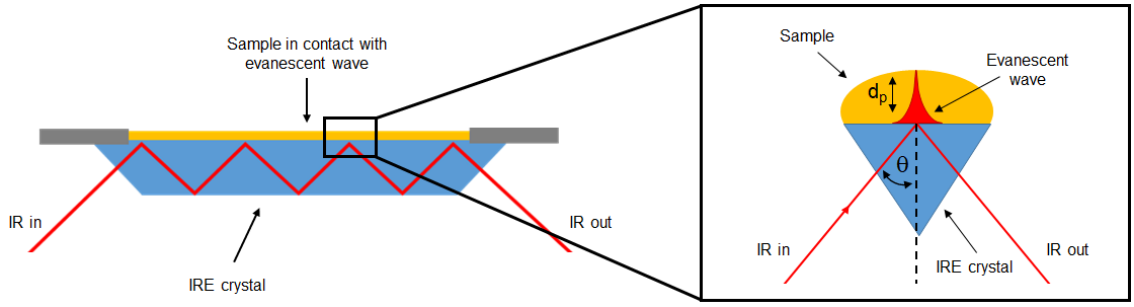

Figure. 1.9: Scheme of infrared path through an ATR crystal where the angle of incidence of the IR beam (θ) is greater than the critical angle. In the inset, a zoom of the propagating evanescent electric field at IRE/sample interface.

In this method, the IR beam is directed to a prism made of an IR transparent material with a high refractive index (n_1), higher than the refractive index of the sample (n_2). Such crystals are known as Internal Reflection Elements (IREs). Germanium (n_1=4 at 1000 cm^{-1}), Silicon (n_1=3.4 at 1000 cm^{-1}), ZnSe (n_1=2.4 at 1000 cm^{-1}) and diamond (n_1=~2.4 at 1000cm^{-1}) are

among the most common IRE crystals. The critical angle (θ_{crit}) is dependent on the refractive indices of the internal reflection element (n_1) and the sample (n_2), according to Snell's law [Eq. 1.6].

$$\sin\theta_{crit} = \frac{n_2}{n_1} \qquad \text{[Eq. 1.6]}$$

Upon reflection at the interface between sample and crystal, an evanescent field is generated that propagates into the sample. The evanescent field has the same frequency of the incoming light but the its amplitude decays exponentially with the distance from the interface. The sample absorbs IR radiation, attenuating the evanescent wave and this attenuation is measured. The penetration depth (d_p) of the evanescent electric field depends on the wavelength (λ), on the angle of incidence and on the ratio of the refractive indices of crystal and sample, as defined in Eq. 1.7.

$$d_p = \frac{\lambda}{2\pi(n_1^2 sin^2\theta - n_2^2)^{1/2}} \qquad \text{[Eq. 1.7]}$$

For the most common IRE materials, considering an average refractive index of the sample about 1.5, the penetration depth varies from hundreds of nanometers to few microns in the MIR (i.e. 0.65 μm for Germanium IRE or 1.65 μm for Diamond crystals at 1000 cm^{-1}).
ATR-FTIR spectroscopy has been extensively used for protein studies. Usually a protein film is prepared on the crystal surface [64]–[66], often by controlled drying/hydration cycles for creating a protein film onto the IRE and keeping the film hydration constant. [49]. Since the penetration depth of the technique is ruled by equation 1.7, for relatively thick protein layers, the volume of the buffer layer on top of the protein film does not dramatically influences the measured spectrum [46]. ATR-FTIR is a method relatively simple and very flexible. However, the preparation of a stable film can be difficult and sometimes impossible. Furthermore, the relatively high concentration of the protein at the IRE interface can drive aggregation phenomena not easy to control and identify, inducing to possible spectra misinterpretation.

(iii) FTIR microscopy

FTIR microscopy allows correlating morphological information on the sample to its chemistry/biochemistry, by recording spatially resolved vibrational spectra at specific sample locations. Morpho-chemical correlation is possible thanks to the coupling of FTIR interferometers with specially designed Visible-Infrared (Vis-IR) microscopes. The Vis-IR microscope looks like a standard visible microscope, but it uses basically all-reflective

optical elements and Cassegrain focusing elements, achromatic optics able to focus both visible and infrared light at the same sample location. The optical scheme of Cassegrain objective is reported in Figure 1.10a, while in Figure 10b the photo of a 15X objective is shown. The photo of the Vis-IR microscope of the Chemical-Life Sciences branch at SISSI beamline at Elettra, where FTIR measurements reported in this thesis were done, is displayed in Figure 1.10c.

Different kinds of objectives are available for different applications, from standard normal incidence objectives, as well as grazing incidence (GI) and attenuated total reflection (ATR). In general, they have Numerical Apertures (NAs) ranging from 0.4 to 0.65 and working distances varying accordingly from less than 1 mm to few centimeters. Therefore, spatially resolved studies can be done sampling in transmission, reflection and ATR. In Figure1.11 a schematic representation of different sampling techniques is reported.

The light transmitted or reflected by the sample is then collected by an IR detector. For Mid-IR, conventionally Mercury Cadmium Telluride (MCT) detectors are employed. They can be made as a single point detector or arranged in an array, linear or bi-dimensional, commonly known as Focal Plane Array (FPA) detector. FTIR hyperspectral images can be obtained either by FTIR mapping of the sample with MCT single-point detector or by direct imaging of larger areas with bi-dimensional FPA one. The workflow is summarized in Figure 1.12.

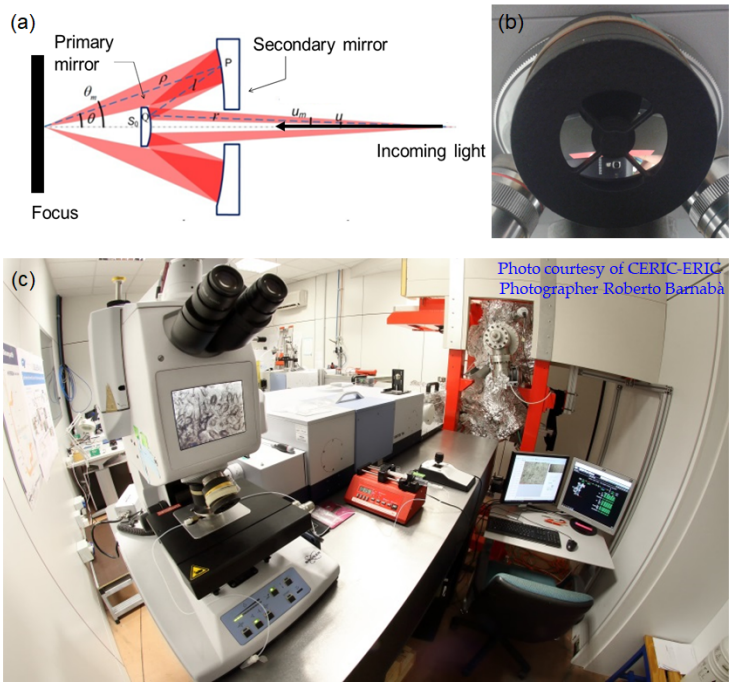

Figure 1.10: (a) Scheme of the optical path of a Cassegrain objective, also known as Schwarzschild objective. Working principle: Cassegrain objective is composed by two concentric spherical mirrors: the "primary" mirror

has a smaller diameter, while the "secondary" mirror has a larger diameter and presents a center aperture, through which visible and IR light enters the objective. Light is back reflected by the convex primary mirror, and is focused by the concave secondary one. (b) The 15X Schwartzschild objective at SISSI beamline. (c) Bruker Hyperion 3000 Vis-IR microscope, equipped with both single-point MCT and 64x64 FPA imaging detectors, and the vertex 70v interferometer of the Chemical-Life Sciences end station at SISSI beamline.

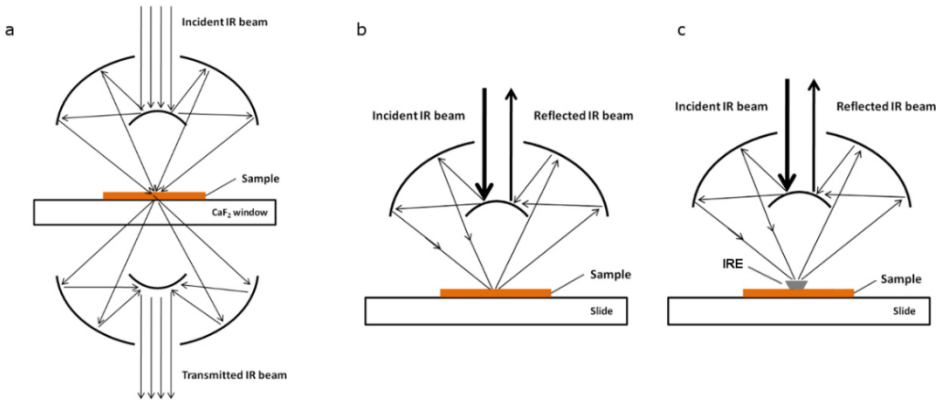

Figure 1.11: Schematic representation of different FTIR microscopy sampling techniques: (a) Transmission: two faced Cassegrain optics are used as focusing condenser and collecting objective; (b) Reflection: For both normal and grazing incidence, only one objective is used for both sample illumination and reflected light collection. This dual operational mode is made possible by the insertion of one splitting mirror along the optical path, which blocks almost 50% of the incoming light; (c) ATR: ATR micro-sampling is done in reflection geometry, but the incoming light is focused at the interface between the sample and the IRE.

In the FTIR mapping mode, the image pixel size is defined by closing knife-edge apertures that, for the Bruker Hyperion 3000 microscope used during this thesis, are placed after the sample. Isolated single spectra or maps can be measured by raster scanning the sample translating the microscope stage. On the contrary, in FTIR imaging thousands of FTIR spectra from a large sample area are collected at once and the pixel resolution is determined by the individual detector element size and by the magnification of the combined optical system. FPA detector available at SISSI- Chemical and Life Science branch is made by 64x64 pixels, but larger bi-dimensional arrays exist.

It is important to notice that knife-aperture setting and FPA sensitive element size represent the pixel resolution of hyperspectral images, while the sampling lateral resolution is mostly limited by diffraction phenomena. In far-field microscopy, the lateral resolution is imposed by the Rayleigh criterion.

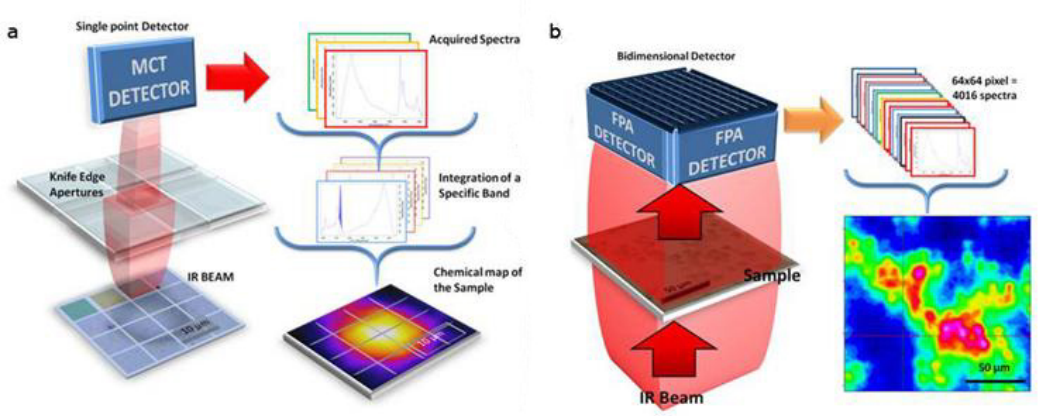

Figure 1. 12: a) Simplified scheme of the IR mapping procedure: first, a visible image is recorded and divided in a matrix of points. Subsequently, the knife-edge apertures of the microscope are closed to the desired pixel resolution and the sample is raster-scanned in order to collect an infrared spectrum every point (hyperspectral image). In the simplest case, by integrating specific absorption bands of the acquired spectra, an image of the chemical distribution of the molecule/moiety of interest is obtained. This procedure is known as univariate functional group mapping. b) Process flow for the IR imaging procedure. The bi-dimensional detector can be seen as a matrix of sensitive elements, made by 64x64 or 128x128 pixels. At once, 64^2 or 128^2 spectra can be acquired. The chemical image of the sample is then obtained as in the case of IR mapping.

As can be deduced by Eq. 1.8, two objects can be resolved at a fixed wavelength when their distance, δ, is at least comparable with the probing wavelength traveling in air (refractive index n=1), stated the numerical aperture, NA, of the used objective in the range of 0.6. Since NA of IR objectives fits in this range, the diffraction limit imposes lateral resolutions in the range of several microns for Mid-IR microscopy.

$$\delta = \frac{0.61 \times \lambda}{NA} \qquad \text{[Eq. 1.8]}$$

The diffraction limit is the principal limitation to IR microscopy lateral resolution, while deviation of optical elements from ideal and factors affecting the signal to noise spectral ratio, S/N, such as source brightness, detector sensitivity, or sampling geometry, further reduce the effective resolution of IR microscopy. In practice, diffraction-limited lateral resolution is hardly achievable employing a conventional thermal Mid-IR source. For thin and/or poor absorbing samples, diffraction-limited measures can be performed at synchrotron facilities, thanking advantage from the brightness benefit offered by InfraRed Synchrotron Radiation (IR-SR). SISSI (Synchrotron Infrared Source for Spectroscopy and Imaging) [67] is the IR beamline at Elettra Sincrotrone Trieste, and the FTIR measurements performed during this thesis has been done at Chemical and Life Sciences branch of SISSI with the instrumentation shown in Figure 1.10c.

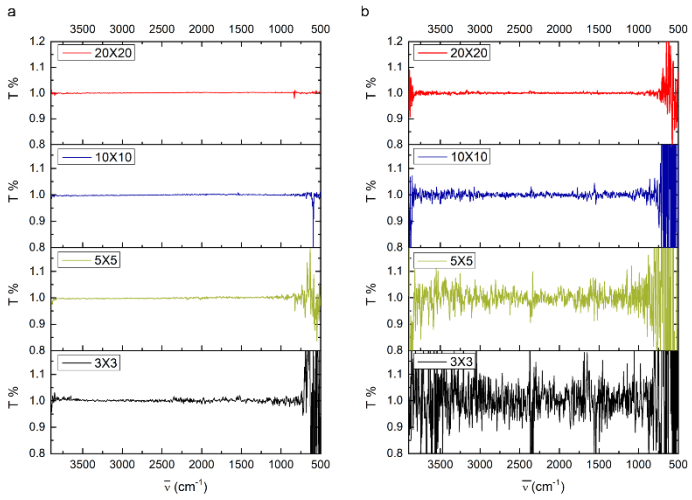

Figure 1.13: Comparison between air transmittance spectra acquired with IR-SR, panel a, and globar source, panel b, by decreasing the pixel size from 20x20 to 3x3 μm^2. Spectra have been acquired in transmission mode using 15X condenser/objective with a scanner velocity of 40 kHz for both globar and IR-SR, averaging the same number of scans (128) at 4 cm^{-1} spectral resolution. The better quality of IR-SR spectra is also associated to a shorter acquisition time (about 1/3).

It is far beyond the scope of this thesis to detail on the IR-SR properties and on the SISSI beamline design and performances, while the reader is refereed to [67] and cited literature herein for more details. However, in order to give to the reader insight into the S/N advantage offered by IR-SR in comparison to conventional thermal Mid-IR source (also known as globar), Figure 1.13 reports the comparison between transmission spectra in air acquired from 20x20 to 3x3 μm^2 pixel resolution averaging 128 scans at 4 cm^{-1} spectral resolution (40KHz scanner velocity) in the 4000-500 cm^{-1} spectral range at SISSI - Chemical al Life Sciences branch with IR-SR (panel a) and globar source (panel b). As can be easily appreciated by the spectral quality, the brightness advantage offered by IR-SR reflects in improved S/N ratio, that can be estimated as the ratio between the detected intensity I and the rms of the 100% transmission line (I/RMS). At 10X10 μm^2, I/RMS evaluated in the 1200-1000 cm^{-1} spectral range, sets at about 10^4 (a.u.) and 10^1 (a.u.) for IR-SR and globar source respectively, showing a S/N improvement of about 3 orders of magnitude.

1.2.4 FTIR spectroscopy of proteins

The non-damaging nature of IR light had boosted its application as label-free, non-destructive technique for measuring biological samples in the last decades. FTIR spectroscopy and microscopy have been widely employed for probing the molecular composition and the structure of biological samples, from single biomolecules [49], [68], [69] and simplified mixtures [70], to complex systems such as cells and tissues [71]–[73].

Mid-IR measurements are nowadays employed in automated tissue classification [74], cancer [75] or disease identification [76], cell sorting [77] and several other applications [78]. Specifically, FTIR spectroscopy is one of the oldest and well-established analytical methods for the assessment of protein secondary structure. One of the strengths of infrared spectroscopy for protein studies is that it is sensitive to a variety of sample forms including solid films or powder [47], solutions [79], or liquid crystals [80], and therefore it is able not only to provide information about protein structure in their native environment, but also giving insight into conformational alternations associated with environmental changes, such as pH, temperature, pressure, solvent and hydration state.

As reported in the section 1.1.1, a protein is defined as a repetition of amino acids sub-units bounded each-other by the amide bonds. This series of amide groups composes the protein backbone which folds, giving the polypeptide structure characterizing each protein. The polypeptide chain gives rise to characteristic normal modes of proteins, the so called amide bands. The stretching and bending modes of C=O, C-N, N-H and O-C-N groups of the amide bond originate nine amide bands, called Amide A, Amide B and Amides I-VII in order of decreasing wavenumber [81]. Amide A band (~3310-3270 cm^{-1}) results from the N-H stretching vibration, and it is insensitive to the conformation of the polypeptide backbone. The amide B band is the second protein mode associated to NH stretching, that absorbs weakly between 3100 and 3030 cm^{-1} [82],[83]. The IR spectrum of the Bovine Serum Albumin (BSA) is reported in Figure 1.14, and is used as representative of protein vibrational profile.

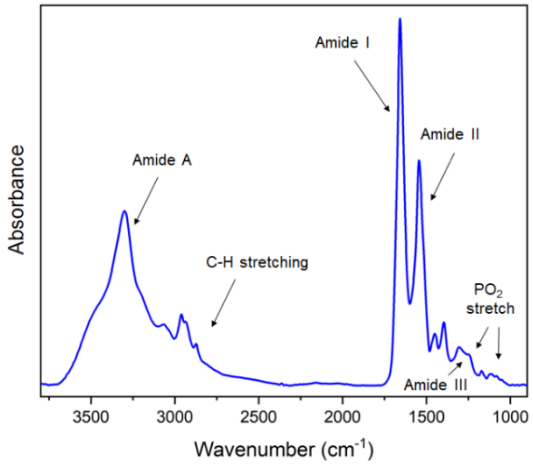

Figure. 1.14: Typical protein absorbance spectrum in the Mid-IR regime. In the picture the absorbance spectrum of Bovine Serum Albumin (BSA) protein is reported. The spectrum has been acquired by transmission microscopy at SISSI beamline, onto the coffee ring [84] formed upon dehydration of a drop of 20 µM BSA solution onto a CaF_2 window. The most relevant protein spectral features are highlighted.

The Amide A band is pointed out by an arrow in the figure, together with two other intense and well-defined bands, called Amide I (1700-1610 cm^{-1}) and Amide II (1595-1480 cm^{-1}). The Amide I vibration arises mainly from the C=O stretching vibration (~80%), with minor contributions from the out-of-phase CN stretching vibration (~10%), the CCN deformation and the NH in-plane bend (~10%) [49]. Among the amide bands sensitive to protein secondary structure, Amide I band is the most intense, and therefore, it is the most commonly used feature for secondary structure analysis (see below). The amide II mode (1595-1480 cm^{-1}) is the out-of-phase combination of the N-H in plane bending (~60%) and the CN stretching (~40%) vibrations with smaller contributions from the CO in plane bend and the CC and CN stretching vibrations. The Amide II vibration is definitely less sensitive than Amide I band to protein secondary structure, and its spectral profile is also more affected by spectral contributions of amino acid side chain vibrations [85]. The Amide III mode (1400-1200-cm^{-1}) is the in-phase combination of the NH bending and the CN stretching vibration with small contributions from the CO in plane bending and the CC stretching. The Amide III band is very sensitive to protein conformation, since the different secondary structures of proteins have more resolved differences in their Amide III spectra with respect to Amide I [86]. However, despite this, it is not often used for the purpose of secondary structure determination, due to its relatively low intensity [86], [87]. Amides IV to VII, that are in the spectral range from ~800 cm^{-1} to ~200 cm^{-1}, are much less intense and basically neglected in protein studies.

From a more general point of view, the contribution of amino acid side chains to the protein spectrum is usually negligible, with the exception of the methyl/methylene stretching and bending modes, at 3000-2800 cm^{-1} and 1480-1360 cm^{-1} respectively. Another region that can be relevant for protein analyses is that of stretching of phosphate moieties (asymmetric stretching, between 1280-1191 cm^{-1}; symmetric stretching between 1150-990 cm^{-1}) [72].

As stated, Amide I is the most exploited band for the characterization of the secondary protein structure. Its position is sensitive to the network of hydrogen bonds involving the carbonyl group and to the coupling of transition dipoles, thus to the protein secondary structure. This band originates from the overlapping of the different contributions coming from helices, sheets, turns and random structures that are present in the secondary structure of each protein. Roughly speaking, stronger is the hydrogen bonding network involving the peptide carbonyl groups, weaker will be the C=O bond, and consequently lower the energy of its vibration. The spectral interval of Amide I sub-components is summarized in Table

1.1, while a representative deconvolution of such contributions for the Amide I bands of Concanavalin A (ConA) is presented in Figure 1.15.

Functional group	Vibrational mode	Frequency range (cm^{-1})	Comments
R-CONHR	Amide I (C=O stretching ~80%; CN stretching ~10%); CCN deformation and the NH in-plane bend ~10%)	1610−1695	Whole range; sensitive to hydration and H-bonding
		1692-1681	β-Turns and loops
		1671-1679	antiparallel pleated sheets and β -turns of proteins
		1671-1658	Turns and Bends
		1657-1651	α-helical structures
		1647-1641	Random coil
		1640-1631	Native β-sheet
		1630-1620	Inter-chain β-sheet typical of aggregates

Table 1.1: Amide I band components, diagnostic for proteins structure and conformation accordingly to [88]. The assignments reported in literature have been obtained by theoretical calculations and by spectra-structure correlations established from experimentally measured model peptides and proteins of well-known 3D structure.

By the direct observation of the Amide I band, the individual sub-components cannot be identified. This is due to the fact that the spectral width of each constitutive component of the Amide I is usually larger than the separation between the maxima of adjacent peaks. In order to perform secondary structure analysis from protein spectra, spectral manipulation is mandatory. Since 1986, a large number of methods to estimate protein secondary structure content via the analysis of FTIR spectra have been reported [48].

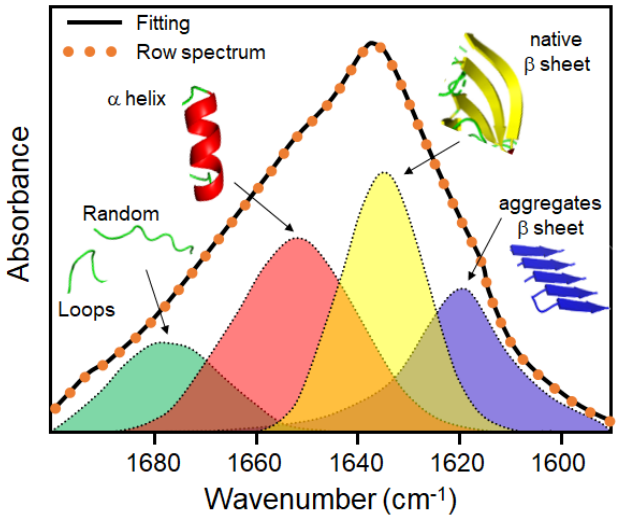

Figure 1.15: Fitting of the Amide I band of Concanavalin A (ConA) protein with its sub-components. Vibrational modes are associated in the picture to their structural elements. ConA is mainly β-sheet folded, as highlighted by the main peak at 1637 cm^{-1}. Figure adapted from [89].

The easier and possibly more direct method for identifying the components of Amide I band by increasing the separation of the overlapping bands is based on derivation procedure, most commonly on second derivative analysis [53], [90]. [91]. An example of the details that can be obtained by second derivative analysis is shown in Figure 1.16, where the ATR-FTIR absorbance spectrum of dry BSA and its second derivative are plotted. In second derivative, the number and position of negative peaks indicates the number and the type of the main contributions to the broad Amide I absorption band, allowing to recognize the different secondary structural elements of the protein. Specifically, BSA second derivative in Figure 1.16a highlights that in addition to the main peak centered at 1656 cm^{-1}, assigned to α helix structures, there is also a spectral component centered at about 1686 cm^{-1}, commonly assigned to loops and turns in proteins (see Table 1.1). It important to remember that, for derivative analysis high quality spectra with good S\N ratio are required, since derivation procedures increase the noise significantly. It is a matter of fact that spectral noise can introduce additional peaks which leads to "over-deconvolution", and therefore misidentification of structural contributions [92]. For this reason, usually routines for performing second derivative analyses includes smoothing functions, such as the Savitzky–Golay smoothing filter [93], in order to counteract the tendency for derivatization to enhance noise. Finally, we should highlight that second derivative analyses can provide only qualitative information on the secondary structure of the protein.

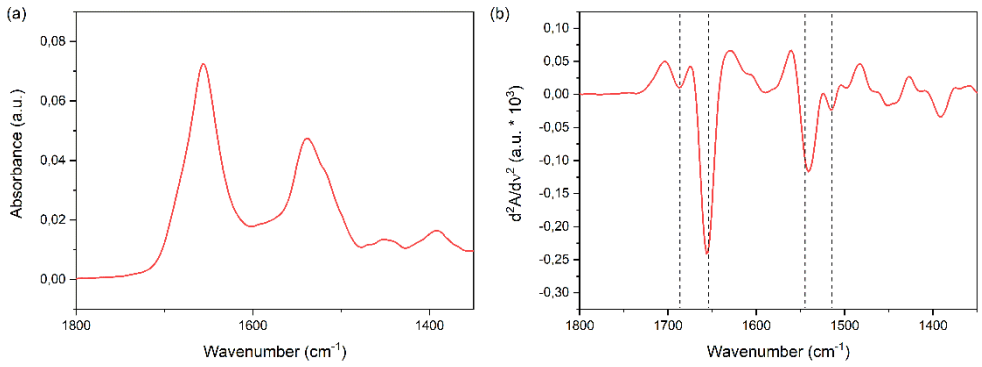

Figure 1.16: ATR-FTIR spectrum and its second derivative of dehydrated BSA (a) FTIR-ATR absorbance spectrum of dry BSA in the amide spectral window (b) Second derivative of BSA protein spectrum in (a). Second derivative analysis allows to identify the position of hidden peaks in the Amide I and Amide II bands. The second derivative peaks relevant for BSA secondary structure are highlighted by vertical dashed lines and discussed in the text.

Comparable considerations can be done also for another method employed for resolution enhancement, known as Fourier Self-Deconvolution (FSD) [94], not used for the analysis of

the data presented in this thesis. Once defined the components of the band, quantitative analysis of their contribution to protein secondary structure can be done by Amide I curve fitting [95], exploiting the results of derivation or FSD [96] for selecting input fitting parameters, such as number of components and their position (See Figure 1.15). Despite considered quite accurate, providing results within an accuracy of 2-3% with respect to X-ray structures, curve fitting procedures always have a certain degree of uncertainty. The subjectivity in the choice of the input parameters, first of all the shape of the constitutive bands (Gaussian curve, Lorentz curve or a mix of the two), and the large number of parameters to be fitted (at least three for each contributing band: position height and width) do not guarantee a unique solution for the fitting. In addition, the band fitting implicitly assumes that the molar absorptivity of the different conformers of Amide I band is the same, but this assumption is not true. Last, but not least, the accurate estimation of the protein secondary structure would imply uniqueness in band assignment. This is clearly not the case, since the same structural components, such as helices and sheets, can assume quite different position depending on the specific protein.

Due to all the aforementioned limitations and doubts, another method can be chosen for detecting changes in protein conformations, namely Difference Spectroscopy. In this approach, spectra of the same protein into two different states are subtracted, to yield a difference spectrum that represents only the structural modifications, disregarding the unaffected protein conformation. Since the conformational variations produces very weak signals in comparison to the bulk protein, very accurate measurements have to be done for exploiting difference spectroscopy, possibly without changing the measurement setup, while performing dynamic experiments in well-controlled environments. If the aforementioned requirements are fulfilled, the method allows to follow small transitions such as changes in functional groups and hydrogen bonding network [97]. Molecular information on protein-protein or protein-ligand interactions have been retrieved on the base of this analyses [49], [87]. Especially, in recent works [97], [98] it has been demonstrated the capability to analyze the conformational effects of a series of enzyme inhibitors with distinct conformational binding modes (helix and loop binder) on target proteins, by using such data processing [98].

1.2.5 Issues for FTIR protein measurements

Despite the great potential of FTIR spectroscopy for protein analyses, this technique suffers two main limitations for the analysis of diluted protein solutions: (i) the so called “water absorption barrier” and (ii) the sensitivity limit.

(i) *The water absorption barrier*: A huge limitation of FTIR spectroscopy of aqueous solutions is the strong absorbance of water in the Mid-IR spectral region. Liquid water has three normal modes of vibration: (i) the asymmetric stretching mode, ν_3, centered at about 3490 cm^{-1}; (ii) the symmetric stretching mode, ν_1, centered at about 3280 cm^{-1}; (iii) the bending mode of water, ν_2, centered at about 1640 cm^{-1} [99]. Figure 1.17 shows the ATR-FTIR spectrum of water in the 3800-900 cm^{-1} spectra region. It is possible to appreciate that the stretching modes of water give a very broad contribution, extending from about 3600 to 2800 cm^{-1}, that also include the overtone of the water bending mode. Moreover, a further band characteristic of liquid water centered at ~2127 cm^{-1} can be seen, that results from the combination of ν_2 with the librational bands [100].

Figure 1.17 plots also the spectrum of hydrated BSA collected in the same experimental conditions. Comparing the BSA spectrum here reported with the one of dry BSA in Figure 1.14, it is easy to deduce that the study of aqueous protein samples is challenging because of the overwhelming of the protein modes by those of water in all the Mid-IR range. Especially, the overlapping of the bending mode with the Amide I band in the 1700-1600 cm^{-1} spectral range is the major drawback for the analysis of the secondary structure of proteins in water solution. It is noteworthy to mention at this stage that water extinction coefficient at ~1650 cm^{-1} is about 20 $L \cdot mol^{-1} \cdot cm^{-1}$. Indeed, this value is on average 20 times smaller than the protein extinction coefficient, but the water concentration is much higher than the protein one. As reported in [100], considering a protein solution 0.3 mol/L and that the concentration of pure water is about 55 mol/L, the bending water mode and Amide I band intensities are 9:1 in proportion, meaning that 90% of the protein spectrum will be dominated by water contribution in the Amide I band region. This can be easily appreciated in Figure 1.17.

The band overlapping can be overcome through the use of deuterated water (^2H_2O or D_2O), since D_2O bending mode is downshifted to ~1210 cm^{-1}. However, the use of D_2O can affect the native secondary structure of the protein and it modifies the spectral behavior of proteins. For example, N-deuteration converts the Amide II mode to largely a CN stretching vibration at 1490–1460 cm^{-1} (named Amide II′ mode) [49].

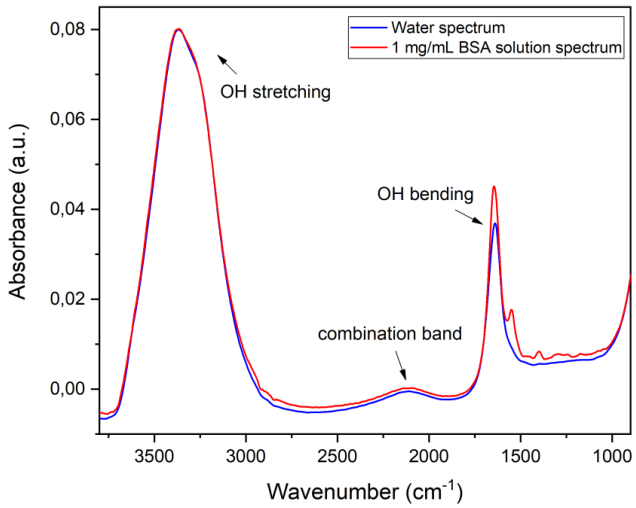

Figure 1.17: ATR-FTIR (Ge IRE) raw spectrum of BSA water solution 1 mg/mL (red line) and the spectrum of milliQ water in the Mid-IR regime (blue line). Main water features are highlighted. As shown, they dominate both spectra and protein contributions can be barely discerned.

On the other hand, water contribution can be potentially subtracted. Several papers present guidelines for accurate subtraction of the water band, in order to avoid over subtraction issues [53], [81]. It is important to notice that, due to the strong water absorption, short path length for aqueous samples, typically less than 10 μm, are demanded in order to avoid HOH bending saturation. ATR-FTIR sampling partially addressed this issue, due to the limited penetration depth of the evanescent filed, but it does not have the adequate sensitivity for probing diluted samples.

(ii) *FTIR sensitivity limitation*: FTIR sensitivity is limited as a result of the Beer's law. Due to the low molecular absorption cross-sections of Mid-IR vibrations ($\sigma_{abs} \approx 10^{-20}$ cm^2), a considerably large amount of material, or relatively high concentrations, are needed for obtaining spectra with good S/N. Typical concentrations in infrared spectroscopy are 0.1 to 1 mM for proteins and 1 to 100 mM for small molecules [49], [101]. A major problem for such concentrated protein solutions is the tendency for the proteins to aggregate in a non-specific way at this concentration levels [57]. Virtually, monolayers of proteins anchored onto a flat surface can be measured using either ATR or polarized-dependent grazing angle incidence setups for ultrathin films on metallic surfaces [102]. Nevertheless, both these techniques require the fully homogenous coverage of large areas, from few millimeters up to centimeters, causing the consumption of large quantity of sample. Indeed, in biology, for an ideal screening process, the amount of used material should be minimized, due to the time and the costs of protein production/expression and the large number of tests to be performed.

In conclusion, without good signal amplification, spectroscopic characterization of minute amounts of analytes is prevented, as in the case of protein monolayers or protein solutions at typical concentration employed for biological assays, in the order of micromolar (μM) regime, but often even lower.

1.3 Overcoming the FTIR detection limit

A promising strategy for overcoming the sensitivity limitations of conventional FTIR technique is to take advantage from the collective electron oscillations at a metal-dielectric interface when they are excited by an incident light [103]. The discipline which exploits such interactions is called plasmonics, which is an emerging area of research in the field of nanophotonics, with many applications in spectroscopy [104], [105], fluorescence microscopy [106], [107], biosensing [108], [109], non-linear [110], [111] and quantum optics [112]. Plasmonics is based on the interaction between electromagnetic (EM) radiation and conduction electrons at metallic interfaces leading to enhanced optical near-fields confined in subwavelength dimensions. Under appropriated conditions, collective oscillations of the electron plasma can be resonantly excited in close proximity of metallic nanostructures. Such localized oscillations, which are called localized surface plasmon resonances (LSPRs), strongly depend on the nanostructure geometry and material. When these oscillations are propagating along metal/dielectric or metal/air interfaces, they are defined surface plasmon polaritons (SPPs) [113], [114]. The collective/resonant nature of plasmon excitation is associated to highly confined (down to the nanometer scale) and enhanced EM fields. As a consequence, strongly localized EM radiation can be used to boost light-matter interactions. The ability to increase the signal with field enhancements is the basis for surface enhanced infrared absorption (SEIRA) spectroscopy [115], in analogy to surface-enhanced Raman scattering (SERS) [116]. For SERS, the EM field enhancement is effective for both the probing and the scattered radiation; it is proportional to 4^{th} power of $|E|$, and usually it takes advantage from resonant wavelengths in the visible [108]. In the case of SEIRA, the EM field enhancement is effective only for the probing radiation, and it is proportional to the 2^{nd} power of $|E|$[117]. Although the enhancement factor of SEIRA is smaller than that of SERS, the cross-section for IR absorption is several orders of magnitude higher than the corresponding Raman cross-section and SEIRA does not suffer of fluorescence background effects often detected in Raman spectroscopy [118].

At the beginning of 1980s, SEIRA effect obtained with rough metal islands films was identified [119]. The so called SEIRA-active surfaces used for conventional SEIRA

spectroscopy are made in a similar way: they consist of large areas of random arrangements of metallic particles, usually silver or gold nanoparticles. By using such systems, no more than three orders of magnitude of direct vibration signal enhancement have been achieved [119], [120], [121]. In recent years, a significant stronger IR enhancement, was obtained by using the so-called Mid-IR resonant nanoantennas [122], [123], which offer the opportunity to resonantly tune the SEIRA response in specific regions of interest in the Mid-IR regime.

In the following subsections a brief overview on conventional SEIRA substrates will be given. Next, the prerequisites for SEIRA with resonant nanoantennas will be discussed. Finally, advantages offered by the so-called Collective Enhanced IR Absorption (CEIRA) microscopy will be presented.

1.3.1 Conventional SEIRA

A SEIRA-like effect was described for the first time in 1980 by Hartstein *et al*. They obtained an IR enhancement by taking advantage from the electric field enhancement provided by collective electron resonances associated with the island nature of the thin metal films coated with their target molecules [119]. Conventional SEIRA spectra were typically measured in transmittance mode on transparent substrates (Figure 1.18 a) or in ATR geometries, with the metal particles placed on the ATR prism (Figure 1.18 b). The multiple reflection ATR geometry [124] was successfully applied for biomolecular sensing and for monitoring molecular reactions on a microsecond time scale [125].

It is important to notice that conventional SEIRA was found to be strongly dependent on the metal surface morphology [117], [128], [129]. Since conventional SEIRA experiments adopted randomly grown island films, the local field enhancement spatially varied on the layer with respect to the morphology of the substrate as a consequence of the random nature of such structures. Due to the distribution of particle sizes and distances, the resulting IR resonant spectra were significantly broadened [130]. The corresponding enhancement factors, defined as the ratio between enhanced vibrational signal and unenhanced reference measurement, were limited to 1-3 orders of magnitude [117], [124], [129]. It is important to notice that it has been demonstrated that the Mid-IR signal enhancement can also appear for metal islands which have the main absorption in the NIR regime, since the resonance tail fall in the Mid-IR range [126], [119]. Thus, due to the very broad plasmonic absorption, the importance of resonance tuning for “conventional” SEIRA enhancement is low.

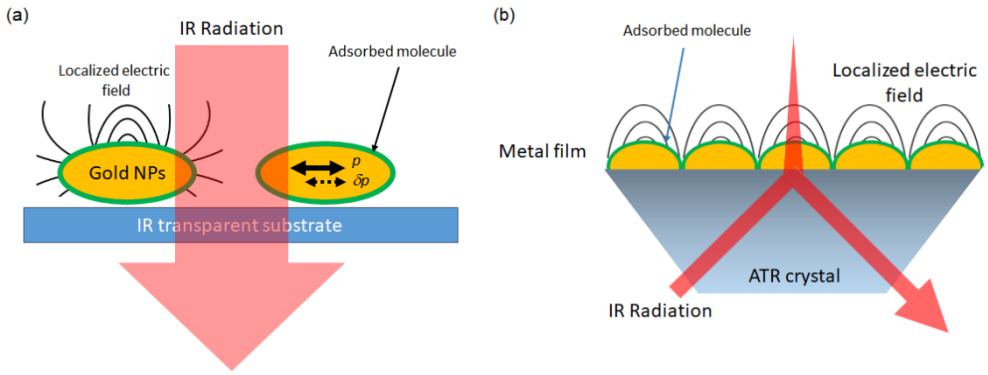

Figure 1.18: Schematic representation of conventional SEIRA acquisition modes: (a) Transmission mode: figure shows the EM mechanism of SEIRA on metal island films acquired in transmission mode. Incident infrared photon field excite the collective electron resonance (or localized plasmon) modes. The dipole *p* induced in an island generates a local EM field stronger than the incident infrared photon field around the island. (Adapted from [126]). (b) ATR mode. Metal islands films are spread onto the ATR crystals. Tailored protocols are used for immobilize the molecules of interest. (Adapted from [127]).

At least two different mechanisms are speculated to contribute to the total SEIRA enhancement: (i) the electromagnetic effect, that is the plasmonic effect, related to the free electron collective/resonant response previously mentioned; and (ii) a chemical effect, that is related to charge transfer upon adsorption. The chemical effect has been first deduced by the fact that molecules chemisorbed on a metal surface had a larger enhancement than physisorbed molecules, suggesting that the absorption coefficients of chemisorbed molecules are larger than those of condensed overlayers. More details on this topic can be found in the ref [103]. In order to fully exploit the chemical enhancement, a well-defined and perfect functionalization of the metal surface is required. Unfortunately, metal roughness hampers the formation of such ordered layers of functional molecules leading, most of the times, to the suppression of the enhancement effect.

Despite their limitation and low enhancement factor, metal-island films are still interesting structures for surface-enhanced vibrational spectroscopy since they are much easier to produce and cheaper than those which requires nanolithographic methods for being fabricated, as in the case of resonant nanoantennas for SEIRA.

1.3.2 Resonant nanoantennas for SEIRA applications

Recent advances in nanofabrication techniques have allowed nanoscale engineering of the optical properties of metallic nanostructures from the UV to the FIR, enabling different applications in many interdisciplinary fields [131]. By changing size and shape of the

plasmonic nanoparticles and their inter-particle separation, the LSPR spectral position can be finely tuned/manipulated across a wide spectral range [109].

It has been demonstrated that strong plasmonic enhancement of optical near fields can also be extended to the Mid-IR regime by using plasmonic nanoantennas [35]. In many studies, simple linear nanorods with micrometer lengths and nanometer cross-sectional geometries were used as model systems [123], [132], [133]. Such linear antennas give rise to pronounced plasmonic resonances in the infrared spectral region caused by collective oscillations of electrons, enabling dramatic signal enhancement for IR spectroscopy [122], [123], [109]. The possibility to finely tune such narrow resonances with the vibrational fingerprints of biomolecules allowed to obtain signal enhancement of several orders of magnitude when they are arranged in arrays (see section 1.3.3 "*Interaction between Resonant Nanoantennas: the Collective enhancement*"). In 2008, Neubrech *et al.* demonstrated for the first time the possibility to perform SEIRA based on sharp plasmonic resonances provided by such nanostructures [122]. They described the phenomenon in terms of plasmonic resonance and antenna resonance, stating that both terms describe the identical physical phenomena, where with the term "plasmonic" they aimed to emphasizes the electronic nature of the system, while the term "antenna" represented the photonic one.

In the IR regime, for a half-wave dipole antenna, the resonance frequency (λ) is given by [Eq. 1.7]:

$$\lambda = \frac{1}{m}(2n_{eff}L)a + C \qquad \text{[Eq. 1.7]}$$

where m is the mode number and n_{eff} is the refractive index of the surrounding medium. The constant C accounts for the phase associated with the reflection at the antenna end, and the parameter a depends on the antenna's geometry and its material [108]. It is important to note that the effective index, n_{eff} is a function of the surface plasmon wave vector supported by the nanorod and therefore it depends both from the background refractive index and the material and geometrical properties of the antenna [122]. The possibility to scale the resonance by modifying the geometry of the nanostructures in a controlled way is ideal for a practical SEIRA spectroscopy platforms, allowing to tune the resonance of the system to a vibrational mode of interest. This tunability provides the opportunity to maximize field enhancements demonstrating the ability of such structures to act as plasmonic biosensors for measuring the absorption spectra of a protein monolayer [123]. A scheme of the principle of resonant SEIRA is reported in Figure 1.19. More details on the pre-requisites for resonant

SEIRA as well as on the fundamental properties of this approach, including details on the resonant coupling mechanism and the signal enhancement, can be found in references [103], [109].

In their simplest form, Mid-IR SEIRA platforms consist of multiple arrays of rod-shaped gold nanoantennas, nanopatterned on substrates and engineered for plasmonic resonant excitations of their dipole mode in the mid-IR frequency range. Conventional FTIR interferometers, coupled to IR-Vis microscopes, are employed to perform such measurements, eventually with the implementation of a microfluidic system to deliver the analyte to the sensing area [134].

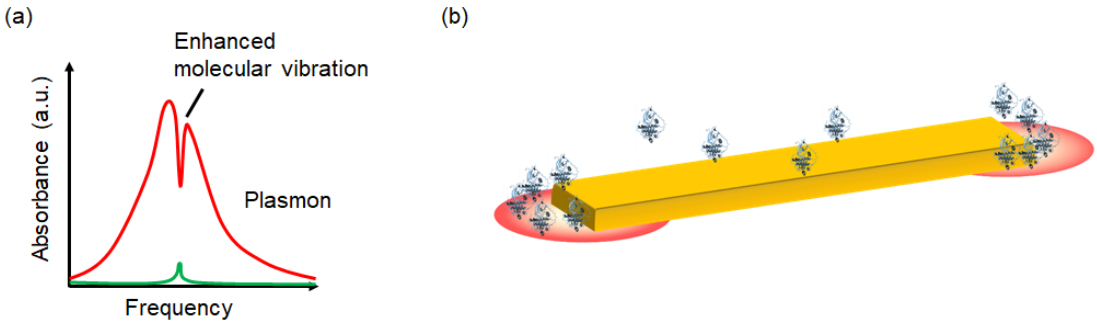

Figure 1.19: Principle of resonant surface-enhanced infrared spectroscopy (SEIRA). (a) The IR vibrations of molecules located in the enhanced electromagnetic near-field of a plasmonic nanostructure (nanoantenna) are enhanced when the resonance plasmonic response of the nanoantenna (red) is resonantly matched to the molecular vibration (green). (b) Sketch of the localization of target molecules onto an individual nanoantenna. Only signals of protein in the enhanced field can be detected. The enhanced field is represented by the red area at the nanoantenna apexes. Image adapted from ref [103].

1.3.3 Interaction between Resonant Nanoantennas: the Collective enhancement

The sensitivity of resonant SEIRA can be further increased if interacting nanostructures are used. Commonly, it is possible to distinguish between near-field interaction of pairs of nanoantennas separated by nanometer sized gaps (dimers) [133], [135] and the excitation of collective plasmonic oscillations in arrays of periodically arranged nanostructures, in the so called Collective Enhanced IR Absorption (CEIRA) microscopy [123].

Near-field coupling offers high field intensities for structures separated with small gaps and filling those gaps, also called “hot-spots”, with target molecules [132], [133]. This phenomenon was predicted by numerical calculations as reported in Figure 1.20a and was first reported for SEIRA by Dregely et al. [135] and fully demonstrated by Huck et al. [133]. An increase of 1 order of magnitude of near-field intensity enhancement was found when

decreasing the gap size from 50 to 3 nm, corresponding to enhancement factors on the order of 10^5, demonstrating that this approach is particular efficient for spectroscopic applications.

On the other side, in arrays of plasmonic structures, another type of interaction mediated by the far-field can provide a further increase the near field intensity, as shown in Figure 1.20b [109]. Far field interaction requires an ordered and periodic arrangement of nanostructures with distances in the order of the incident wavelength, enabling the excitation of collective oscillations [103], [109], [123]. Adato and coworkers defined this surface enhanced spectroscopy technique, based on collective plasmonic excitations, CEIRA (Collective Enhanced InfraRed Absorption). They demonstrated that such collective interaction on the far-field can provides enhancement factors of several orders of magnitude of the vibrational absorption signatures of the layered protein, with high signal to noise ratios [123].

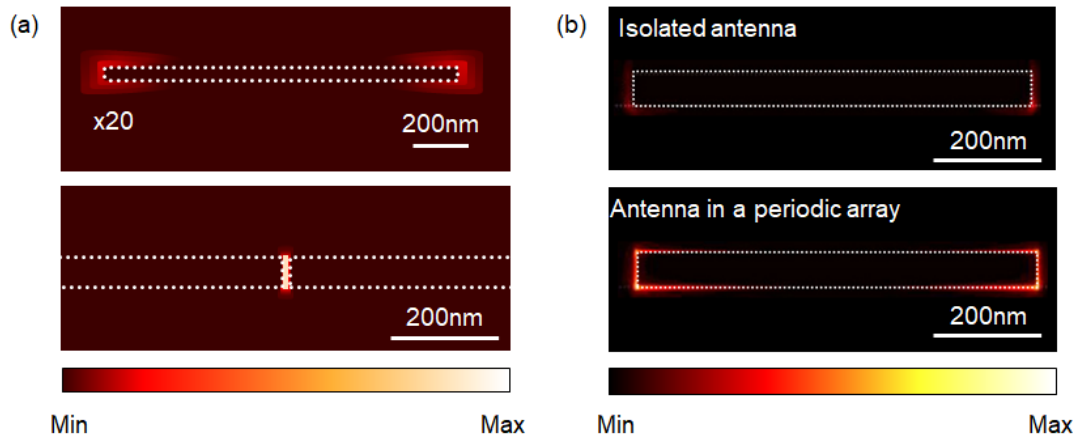

Figure 1.20: (a) Nanoantennas dimers: Representation of near-field distributions of a single antenna and an antenna dimer with a gap size of 10 nm at respective resonance frequencies obtained from FDTD simulations. The intensity of the single antenna was multiplied by a factor of 20. Adapted from ref [133]. Copyright 2014 American Chemical Society. (b) Numerically calculated distribution of near-field intensity for an isolated antenna and a periodic antenna array (d = 1.6 μm). Collective resonances in periodic arrangements of nanoantennas can give rise to nearly an order of magnitude larger near-field intensity enhancements compared with the isolated antenna. Adapted from [123]. Copyright 2009 National Academy of Sciences.

1.3.4 The Fano-type effect

As stated, in the enhancement mechanism of resonant SEIRA, the plasmonic excitation interact with molecular vibrational dipoles. When vibrational dipoles are in resonance with the plasmonic excitation, they strongly modify the plasmonic extinction and, most of the times, the enhanced vibrational signatures do not appear as conventional Lorentzian-type line shape, but the show asymmetric line shape with a Fano-type profile [103]. The

modification of the line shape is due to the interaction with the narrowed vibration of molecules and the broad plasmonic resonance. For giving an intuitive understanding of the experimental SEIRA data, different models have been developed.

In particular, the Fano formalism model, states that Fano-resonances features with an asymmetric profile (Fano-profile) can be described by the expression [eq 1.8]:

$$I \propto \frac{(q\gamma + \omega - \omega_0)^2}{(\omega - \omega_0)^2 + \gamma^2} \qquad \text{[eq. 1.8]}$$

where ω_0 is the resonance frequency, γ is the width of the resonance and q is the Fano-parameter, introduced by Ugo Fano in 1961 [136], which describe the asymmetry of the vibration. By this model, the asymmetry arises from the constructive and destructive interference of the two excitation pathways. Particularly, one excitation path is a broad spectral line or continuum, while the other one is a narrow or discrete resonance. In SEIRA, the plasmonic excitation correspond to the continuum (or broad resonance) and the molecular excitation corresponds to the discrete excitation (or narrow resonance). Depending of the phase of the electromagnetic interaction, which is described by the parameter q in the Fano-model [136], different kinds of line shapes can be observed: q_0 correspond to the anti-phase relation resulting in Lorentzian anti-absorption like peaks; asymmetric lines for intermediate detuning; without coupling, $|q| = \infty$, the Lorentzian line shape is recovered [103]. The asymmetry parameter q, adopted as a measure of the coupling is CEIRA, depends on the oscillator strength of the narrow vibrational excitation and on that of the plasmonic continuum at the same frequency, thus on the spectral tuning between the vibrational and the plasmonic extinction [103].

1.3.5 Effects of the spectral tuning

As we have seen, the enhanced vibrational signals arise as a consequence of the molecule-plasmon coupling, and as the antennas is detuned with respect to the vibrations, progressively decreasing vibrational signals with clearly asymmetric profiles are observed [103]. Thus, it is of paramount importance to match the plasmonic response (ω_{res}) and vibrational frequency (ω_{vib}): in fact, only at their best matching, or on in a close proximity of it, the strongest signal enhancement can be obtained [109], [122] (see Figure 1.21).

Moreover, as shown in Figure 1.21, it has been reported that when a target molecule is placed onto the surface of a nanoantenna, a red-shift of it ω_{res} occurs. However, for SEIRA and CEIRA applications this shift is of minor importance and in particular, if broad vibrational

absorptions are investigated, such as amides vibrations of proteins with a bandwidth of more than 100 cm^{-1}, this shift can be considered negligible. [137].

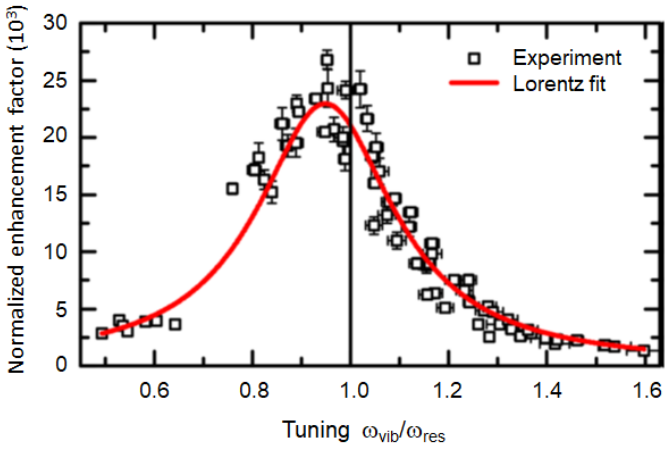

Figure 1.21: Exemplificative case: in the image normalized enhancement factor of the seven strongest modes of an organic molecule in the considered spectral fingerprint range versus the ratio of molecular and plasmonic resonance frequencies $\omega_{vib}/\omega_{res}$. In order to better understand the trend, a Lorentzian curve was fitted to the measured data (red curve). The vibrational signal enhancement peaks at $\omega_{vib}/\omega_{res} = 0.95$ revealing a red shift of the maximum plasmonic near-field intensity with respect to the far-field antenna resonance ($\omega_{vib}/\omega_{res} = 1$, black line). Adapted from ref [137]. Copyright 2015 The PCCP Owner Societies.

References

[1] J. M. Berg, J. L. Tymoczko, L. Stryer, J. M. Berg, J. L. Tymoczko, and L. Stryer, *Biochemistry*, 5th ed. W H Freeman, 2002.

[2] B. Alberts, A. Johnson, J. Lewis, M. Raff, K. Roberts, and P. Walter, *Molecular Biology of the Cell*, 4th ed. Garland Science, 2002.

[3] C. M. Dobson, "Protein folding and misfolding," *Nature*, vol. 426, no. 6968, pp. 884–890, Dec. 2003.

[4] H. M. Berman, "Creating a community resource for protein science," *Protein Sci. Publ. Protein Soc.*, vol. 21, no. 11, pp. 1587–1596, Nov. 2012.

[5] D. Eisenberg, "The discovery of the α-helix and β-sheet, the principal structural features of proteins," *Proc. Natl. Acad. Sci. U. S. A.*, vol. 100, no. 20, pp. 11207–11210, Sep. 2003.

[6] M. E. Karpen, P. L. D. Haseth, and K. E. Neet, "Differences in the amino acid distributions of 310-helices and α-helices," *Protein Sci.*, vol. 1, no. 10, pp. 1333–1342, Oct. 1992.

[7] R. S. Vieira-Pires and J. H. Morais-Cabral, "310 helices in channels and other membrane proteins," *J. Gen. Physiol.*, vol. 136, no. 6, pp. 585–592, Dec. 2010.

[8] L. V. Sequist, "First-Generation Epidermal Growth Factor Receptor Tyrosine Kinase Inhibitors in EGFR Mutation; Positive Non-small Cell Lung Cancer Patients," *J. Thorac. Oncol.*, vol. 3, no. 6, Supplement 2, pp. S143–S145, Jun. 2008.

[9] M. Biasini *et al.*, "SWISS-MODEL: modelling protein tertiary and quaternary structure using evolutionary information," *Nucleic Acids Res.*, vol. 42, no. W1, pp. W252–W258, Jul. 2014.

[10] G. E. Schulz and R. H. Schirmer, *Principles of Protein Structure*. Springer Science & Business Media, 2013.

[11] D. B. Wetlaufer, "Nucleation, Rapid Folding, and Globular Intrachain Regions in Proteins," *Proc. Natl. Acad. Sci.*, vol. 70, no. 3, pp. 697–701, Mar. 1973.

[12] S. K. Hanks and A. M. Quinn, "[2] Protein kinase catalytic domain sequence database: Identification of conserved features of primary structure and classification of family members," in *Methods in Enzymology*, vol. 200, Academic Press, 1991, pp. 38–62.

[13] S. K. Hanks and T. Hunter, "Protein kinases 6. The eukaryotic protein kinase superfamily: kinase (catalytic) domain structure and classification.," *FASEB J.*, vol. 9, no. 8, pp. 576–596, May 1995.

[14] M. Congreve, C. W. Murray, and T. L. Blundell, "Keynote review: Structural biology and drug discovery," *Drug Discov. Today*, vol. 10, no. 13, pp. 895–907, Jul. 2005.

[15] M. Duran-Frigola, R. Mosca, and P. Aloy, "Structural Systems Pharmacology: The Role of 3D Structures in Next-Generation Drug Development," *Chem. Biol.*, vol. 20, no. 5, pp. 674–684, May 2013.

[16] W. W. Yue, D. S. Froese, and P. E. Brennan, "The role of protein structural analysis in the next generation sequencing era," *Top. Curr. Chem.*, vol. 336, pp. 67–98, 2014.

[17] D. K. Brown and Ö. Tastan Bishop, "The role of structural bioinformatics in drug discovery via computational SNP analysis – a proposed protocol for analyzing variation at the protein level," *Glob. Heart*, vol. 12, no. 2, pp. 151–161, Jun. 2017.

[18] J. Drenth, *Principles of Protein X-Ray Crystallography*. Springer Science & Business Media, 2007.

[19] I. Schlichting, "X-ray crystallography of protein-ligand interactions," *Methods Mol. Biol. Clifton NJ*, vol. 305, pp. 155–166, 2005.

[20] A. P. Turnbull and P. Emsley, "Studying protein-ligand interactions using X-ray crystallography," *Methods Mol. Biol. Clifton NJ*, vol. 1008, pp. 457–477, 2013.

[21] A. J. Edwards and D. Reid, "Introduction to NMR of proteins," *Curr. Protoc. Protein Sci.*, vol. Chapter 17, p. Unit 17.5, May 2001.

[22] F. Castellani, B. van Rossum, A. Diehl, M. Schubert, K. Rehbein, and H. Oschkinat, "Structure of a protein determined by solid-state magic-angle-spinning NMR spectroscopy," *Nature*, vol. 420, no. 6911, pp. 98–102, Nov. 2002.

[23] V. Tugarinov and L. E. Kay, “Methyl Groups as Probes of Structure and Dynamics in NMR Studies of High-Molecular-Weight Proteins,” *ChemBioChem*, vol. 6, no. 9, pp. 1567–1577, Sep. 2005.
[24] “PDB Statistics: Growth of Structures from X-ray Crystallography Experiments Released per Year.” [Online]. Available: https://www.rcsb.org/stats/growth/xray. [Accessed: 30-Nov-2018].
[25] C. Haas and J. Drenth, “Understanding protein crystallization on the basis of the phase diagram,” *J. Cryst. Growth*, vol. 196, no. 2, pp. 388–394, Jan. 1999.
[26] K. V. Dunlop, R. T. Irvin, and B. Hazes, “Pros and cons of cryocrystallography: should we also collect a room-temperature data set?,” *Acta Crystallogr. D Biol. Crystallogr.*, vol. 61, no. Pt 1, pp. 80–87, Jan. 2005.
[27] “X-ray Crystallography Platform - Creative Biostructure.” [Online]. Available: https://www.creative-biostructure.com/x-ray-crystallography-platform_60.htm. [Accessed: 14-Nov-2018].
[28] A. L. Carvalho, J. Trincão, and M. J. Romão, “X-ray crystallography in drug discovery,” *Methods Mol. Biol. Clifton NJ*, vol. 572, pp. 31–56, 2009.
[29] K. R. Acharya and M. D. Lloyd, “The advantages and limitations of protein crystal structures,” *Trends Pharmacol. Sci.*, vol. 26, no. 1, pp. 10–14, Jan. 2005.
[30] N. E. Chayen and E. Saridakis, “Protein crystallization: from purified protein to diffraction-quality crystal,” *Nat. Methods*, vol. 5, no. 2, pp. 147–153, Feb. 2008.
[31] P. J. Loll, “Membrane proteins, detergents and crystals: what is the state of the art?,” *Acta Crystallogr. Sect. F Struct. Biol. Commun.*, vol. 70, no. Pt 12, pp. 1576–1583, Nov. 2014.
[32] I. Ubarretxena-Belandia and D. L. Stokes, “Present and future of membrane protein structure determination by electron crystallography,” *Adv. Protein Chem. Struct. Biol.*, vol. 81, pp. 33–60, 2010.
[33] M. A. Williams, “Protein–Ligand Interactions: Fundamentals,” in *Protein-Ligand Interactions: Methods and Applications*, M. A. Williams and T. Daviter, Eds. Totowa, NJ: Humana Press, 2013, pp. 3–34.
[34] M. Egli, “Diffraction Techniques in Structural Biology,” *Curr. Protoc. Nucleic Acid Chem. Ed. Serge Beaucage Al*, vol. CHAPTER 7, p. Unit-7.13, Jun. 2010.
[35] S. P. Meisburger, W. C. Thomas, M. B. Watkins, and N. Ando, “X-ray Scattering Studies of Protein Structural Dynamics,” *Chem. Rev.*, vol. 117, no. 12, pp. 7615–7672, Jun. 2017.
[36] A. G. Kikhney and D. I. Svergun, “A practical guide to small angle X-ray scattering (SAXS) of flexible and intrinsically disordered proteins,” *FEBS Lett.*, vol. 589, no. 19, Part A, pp. 2570–2577, Sep. 2015.
[37] W. Bras, S. Koizumi, and N. J. Terrill, “Beyond simple small-angle X-ray scattering: developments in online complementary techniques and sample environments,” *IUCrJ*, vol. 1, no. Pt 6, pp. 478–491, Sep. 2014.
[38] M. Carroni and H. R. Saibil, “Cryo electron microscopy to determine the structure of macromolecular complexes,” *Methods San Diego Calif*, vol. 95, pp. 78–85, Feb. 2016.
[39] E. Nogales and S. H. W. Scheres, “Cryo-EM: A Unique Tool for the Visualization of Macromolecular Complexity,” *Mol. Cell*, vol. 58, no. 4, pp. 677–689, May 2015.
[40] K. Murata and M. Wolf, “Cryo-electron microscopy for structural analysis of dynamic biological macromolecules,” *Biochim. Biophys. Acta BBA - Gen. Subj.*, vol. 1862, no. 2, pp. 324–334, Feb. 2018.
[41] Y. Liu, S. Gonen, T. Gonen, and T. O. Yeates, “Near-atomic cryo-EM imaging of a small protein displayed on a designed scaffolding system,” *Proc. Natl. Acad. Sci.*, vol. 115, no. 13, pp. 3362–3367, Mar. 2018.
[42] A. Merk *et al.*, “Breaking Cryo-EM Resolution Barriers to Facilitate Drug Discovery,” *Cell*, vol. 165, no. 7, pp. 1698–1707, Jun. 2016.
[43] Y.-H. Chen, J. T. Yang, and H. M. Martinez, “Determination of the secondary structures of proteins by circular dichroism and optical rotatory dispersion,” *Biochemistry (Mosc.)*, vol. 11, no. 22, pp. 4120–4131, Oct. 1972.
[44] S. Khrapunov, “CD spectroscopy has intrinsic limitations for protein secondary structure analysis,” *Anal. Biochem.*, vol. 389, no. 2, pp. 174–176, Jun. 2009.

[45] N. Sreerama and R. W. Woody, “Estimation of Protein Secondary Structure from Circular Dichroism Spectra: Comparison of CONTIN, SELCON, and CDSSTR Methods with an Expanded Reference Set,” *Anal. Biochem.*, vol. 287, no. 2, pp. 252–260, Dec. 2000.
[46] S. G. Kazarian and K. L. A. Chan, “ATR-FTIR spectroscopic imaging: recent advances and applications to biological systems,” *The Analyst*, vol. 138, no. 7, pp. 1940–1951, Apr. 2013.
[47] M. Jackson and H. H. Mantsch, “The use and misuse of FTIR spectroscopy in the determination of protein structure,” *Crit. Rev. Biochem. Mol. Biol.*, vol. 30, no. 2, pp. 95–120, 1995.
[48] D. M. Byler and H. Susi, “Examination of the secondary structure of proteins by deconvolved FTIR spectra,” *Biopolymers*, vol. 25, no. 3, pp. 469–487, Mar. 1986.
[49] A. Barth, “Infrared spectroscopy of proteins,” *Biochim. Biophys. Acta BBA - Bioenerg.*, vol. 1767, no. 9, pp. 1073–1101, Sep. 2007.
[50] N. Sreerama and R. W. Woody, “Estimation of protein secondary structure from circular dichroism spectra: comparison of CONTIN, SELCON, and CDSSTR methods with an expanded reference set,” *Anal. Biochem.*, vol. 287, no. 2, pp. 252–260, Dec. 2000.
[51] N. J. Greenfield, “Determination of the folding of proteins as a function of denaturants, osmolytes or ligands using circular dichroism,” *Nat. Protoc.*, vol. 1, no. 6, pp. 2733–2741, 2006.
[52] B. Stuart, “Infrared Spectroscopy,” in *Kirk-Othmer Encyclopedia of Chemical Technology*, American Cancer Society, 2005.
[53] A. Dong, P. Huang, and W. S. Caughey, “Protein secondary structures in water from second-derivative amide I infrared spectra,” *Biochemistry (Mosc.)*, vol. 29, no. 13, pp. 3303–3308, Apr. 1990.
[54] “Chapter 1. Introduction – the discovery and use of infrared.” [Online]. Available: http://elte.prompt.hu/sites/default/files/tananyagok/InfraredAstronomy/ch01.html. [Accessed: 30-Nov-2018].
[55] T. E. Madey and J. T. Y. Jr, Eds., *Vibrational Spectroscopy of Molecules on Surfaces*. Springer US, 1987.
[56] C. M. Caves, K. S. Thorne, R. W. P. Drever, V. D. Sandberg, and M. Zimmermann, “On the measurement of a weak classical force coupled to a quantum-mechanical oscillator. I. Issues of principle,” *Rev. Mod. Phys.*, vol. 52, pp. 341–392, Apr. 1980.
[57] E. B. Wilson, J. C. Decius, and P. Cross, *Molecular Vibrations: The Theory of Infrared and Raman Vibrational Spectra*, Revised edizione. New York: Dover Pubns, 1980.
[58] “ChemTube3D.” [Online]. Available: http://www.chemtube3d.com/. [Accessed: 16-Nov-2018].
[59] “Water absorption spectrum.” [Online]. Available: http://www1.lsbu.ac.uk/water/water_vibrational_spectrum.html. [Accessed: 14-Nov-2018].
[60] C. Berthomieu and R. Hienerwadel, “Fourier transform infrared (FTIR) spectroscopy,” *Photosynth. Res.*, vol. 101, no. 2–3, pp. 157–170, Sep. 2009.
[61] L. Vaccari, G. Birarda, L. Businaro, S. Pacor, and G. Grenci, “Infrared Microspectroscopy of Live Cells in Microfluidic Devices (MD-IRMS): Toward a Powerful Label-Free Cell-Based Assay,” *Anal. Chem.*, vol. 84, no. 11, pp. 4768–4775, Jun. 2012.
[62] E. Mitri, G. Birarda, L. Vaccari, S. Kenig, M. Tormen, and G. Grenci, “SU-8 bonding protocol for the fabrication of microfluidic devices dedicated to FTIR microspectroscopy of live cells,” *Lab. Chip*, vol. 14, no. 1, pp. 210–218, Jan. 2014.
[63] P. Zucchiatti *et al.*, “Contribution of Ribonucleic Acid (RNA) to the Fourier Transform Infrared (FTIR) Spectrum of Eukaryotic Cells,” *Anal. Chem.*, vol. 88, no. 24, pp. 12090–12098, Dec. 2016.
[64] K. A. Oberg and A. L. Fink, “A new attenuated total reflectance Fourier transform infrared spectroscopy method for the study of proteins in solution,” *Anal. Biochem.*, vol. 256, no. 1, pp. 92–106, Feb. 1998.
[65] E. Goormaghtigh, V. Raussens, and J. M. Ruysschaert, “Attenuated total reflection infrared spectroscopy of proteins and lipids in biological membranes,” *Biochim. Biophys. Acta*, vol. 1422, no. 2, pp. 105–185, Jul. 1999.

[66] P. Rigler, W.-P. Ulrich, and H. Vogel, “Controlled Immobilization of Membrane Proteins to Surfaces for Fourier Transform Infrared Investigations,” *Langmuir*, vol. 20, no. 19, pp. 7901–7903, Sep. 2004.
[67] S. Lupi *et al.*, “Performance of SISSI, the infrared beamline of the ELETTRA storage ring,” *JOSA B*, vol. 24, no. 4, pp. 959–964, Apr. 2007.
[68] B. R. Wood, “The importance of hydration and DNA conformation in interpreting infrared spectra of cells and tissues,” *Chem. Soc. Rev.*, vol. 45, no. 7, pp. 1980–1998, Apr. 2016.
[69] L. G. Benning, V. R. Phoenix, N. Yee, and M. J. Tobin, “Molecular characterization of cyanobacterial silicification using synchrotron infrared micro-spectroscopy11Associate editor: J. P. Amend,” *Geochim. Cosmochim. Acta*, vol. 68, no. 4, pp. 729–741, Feb. 2004.
[70] E. Benedetti, E. Bramanti, F. Papineschi, I. Rossi, and E. Benedetti, “Determination of the Relative Amount of Nucleic Acids and Proteins in Leukemic and Normal Lymphocytes by Means of Fourier Transform Infrared Microspectroscopy,” *Appl. Spectrosc.*, vol. 51, no. 6, pp. 792–797, Jun. 1997.
[71] M. J. Tobin *et al.*, “FTIR spectroscopy of single live cells in aqueous media by synchrotron IR microscopy using microfabricated sample holders,” *Vib. Spectrosc.*, vol. 53, no. 1, pp. 34–38, May 2010.
[72] M. J. Tobin *et al.*, “Infrared microscopy of epithelial cancer cells in whole tissues and in tissue culture, using synchrotron radiation,” *Faraday Discuss.*, vol. 126, pp. 27-39-92, 2004.
[73] P. Lasch and D. Naumann, “Spatial resolution in infrared microspectroscopic imaging of tissues,” *Biochim. Biophys. Acta BBA - Biomembr.*, vol. 1758, no. 7, pp. 814–829, Jul. 2006.
[74] E. Njoroge, S. R. Alty, M. R. Gani, and M. Alkatib, “Classification of Cervical Cancer Cells using FTIR Data,” in *2006 International Conference of the IEEE Engineering in Medicine and Biology Society*, 2006, pp. 5338–5341.
[75] S. Kenig, M. B. D. Alonso, M. M. Mueller, and T. T. Lah, “Glioblastoma and endothelial cells cross-talk, mediated by SDF-1, enhances tumour invasion and endothelial proliferation by increasing expression of cathepsins B, S, and MMP-9,” *Cancer Lett.*, vol. 289, no. 1, pp. 53–61, Mar. 2010.
[76] E. Mitri *et al.*, “A new light on Alkaptonuria: A Fourier-transform infrared microscopy (FTIRM) and low energy X-ray fluorescence (LEXRF) microscopy correlative study on a rare disease,” *Biochim. Biophys. Acta BBA - Gen. Subj.*, vol. 1861, no. 5, Part A, pp. 1000–1008, May 2017.
[77] D. E. Bedolla *et al.*, “Determination of cell cycle phases in live B16 melanoma cells using IRMS,” *Analyst*, vol. 138, no. 14, pp. 4015–4021, Jun. 2013.
[78] E. Mitri *et al.*, “Time-resolved FT-IR microspectroscopy of protein aggregation induced by heat-shock in live cells,” *Anal. Chem.*, vol. 87, no. 7, pp. 3670–3677, Apr. 2015.
[79] P. I. Haris and F. Severcan, “FTIR spectroscopic characterization of protein structure in aqueous and non-aqueous media,” *J. Mol. Catal. B Enzym.*, vol. 7, no. 1, pp. 207–221, Sep. 1999.
[80] R. Bhargava, S.-Q. Wang, and J. L. Koenig, “FTIR Imaging Studies of a New Two-Step Process To Produce Polymer Dispersed Liquid Crystals,” *Macromolecules*, vol. 32, no. 8, pp. 2748–2760, Apr. 1999.
[81] J. Kong and S. Yu, “Fourier Transform Infrared Spectroscopic Analysis of Protein Secondary Structures,” *Acta Biochim. Biophys. Sin.*, vol. 39, no. 8, pp. 549–559, Aug. 2007.
[82] J. N. Herron, W. Jiskoot, and D. J. A. Crommelin, *Physical Methods to Characterize Pharmaceutical Proteins*. Springer Science & Business Media, 2013.
[83] H. Susi, S. N. Timasheff, and L. Stevens, “Infrared Spectra and Protein Conformations in Aqueous Solutions I. THE AMIDE I BAND IN H2O AND D2O SOLUTIONS,” *J. Biol. Chem.*, vol. 242, no. 23, pp. 5460–5466, Dec. 1967.
[84] S. Choi and G. Birarda, “Protein Mixture Segregation at Coffee-Ring: Real-Time Imaging of Protein Ring Precipitation by FTIR Spectromicroscopy,” *J. Phys. Chem. B*, vol. 121, no. 30, pp. 7359–7365, Aug. 2017.
[85] S. Krimm and J. Bandekar, “Vibrational spectroscopy and conformation of peptides, polypeptides, and proteins,” *Adv. Protein Chem.*, vol. 38, pp. 181–364, 1986.

[86] S. Cai and B. R. Singh, “A Distinct Utility of the Amide III Infrared Band for Secondary Structure Estimation of Aqueous Protein Solutions Using Partial Least Squares Methods,” *Biochemistry (Mosc.)*, vol. 43, no. 9, pp. 2541–2549, Mar. 2004.
[87] A. Barth and C. Zscherp, “What vibrations tell about proteins,” *Q. Rev. Biophys.*, vol. 35, no. 4, pp. 369–430, Nov. 2002.
[88] Z. Movasaghi, S. Rehman, and D. I. ur Rehman, “Fourier Transform Infrared (FTIR) Spectroscopy of Biological Tissues,” *Appl. Spectrosc. Rev.*, vol. 43, no. 2, pp. 134–179, Feb. 2008.
[89] F. Piccirilli, G. Schirò, V. Vetri, S. Lupi, A. Perucchi, and V. Militello, “Decoding vibrational states of Concanavalin A amyloid fibrils,” *Biophys. Chem.*, vol. 199, pp. 17–24, Apr. 2015.
[90] W. Kabsch and C. Sander, “Dictionary of protein secondary structure: pattern recognition of hydrogen-bonded and geometrical features,” *Biopolymers*, vol. 22, no. 12, pp. 2577–2637, Dec. 1983.
[91] M. R. Whitbeck, “Second Derivative Infrared Spectroscopy,” *Appl. Spectrosc.*, vol. 35, no. 1, pp. 93–95, Jan. 1981.
[92] E. Knözinger, “P. R. Griffiths, J. A. de Haseth: Fourier Transform Infrared Spectroscopy, Vol. 83 aus der Reihe: Chemical Analysis—A Series of Monographs of Analytical Chemistry and Its Applications, John Wiley + Sons, Chichester, New York, Brisbane, Toronto, Singapore 1986. 656 Seiten, Preis: £ 76.75.,” *Berichte Bunsenges. Für Phys. Chem.*, vol. 90, no. 12, pp. 1240–1241, Dec. 1986.
[93] J. Steinier, Y. Termonia, and J. Deltour, “Smoothing and differentiation of data by simplified least square procedure” *Anal. Chem.*, vol. 44, no 11, pp 1906–1909. Sep. 1972.
[94] W. K. Surewicz, H. H. Mantsch, and D. Chapman, “Determination of protein secondary structure by Fourier transform infrared spectroscopy: a critical assessment,” *Biochemistry (Mosc.)*, vol. 32, no. 2, pp. 389–394, Jan. 1993.
[95] W.-J. Yang, P. R. Griffiths, D. M. Byler, and H. Susi, “Protein Conformation by Infrared Spectroscopy: Resolution Enhancement by Fourier Self-Deconvolution,” *Appl. Spectrosc.*, vol. 39, no. 2, pp. 282–287, Mar. 1985.
[96] F. Piccirilli, S. Mangialardo, P. Postorino, S. Lupi, and A. Perucchi, “FTIR analysis of the high pressure response of native insulin assemblies,” *J. Mol. Struct.*, vol. 1050, pp. 159–165, Oct. 2013.
[97] P. Pinkerneil, J. Güldenhaupt, K. Gerwert, and C. Kötting, “Surface-Attached Polyhistidine-Tag Proteins Characterized by FTIR Difference Spectroscopy,” *ChemPhysChem*, vol. 13, no. 11, pp. 2649–2653, Aug. 2012.
[98] J. Güldenhaupt *et al.*, “Ligand-Induced Conformational Changes in HSP90 Monitored Time Resolved and Label Free—Towards a Conformational Activity Screening for Drug Discovery,” *Angew. Chem. Int. Ed.*, vol. 57, no. 31, pp. 9955–9960, Jul. 2018.
[99] Y. Liu and L. Ojamäe, “Fingerprints in IR OH vibrational spectra of H2O clusters from different H-bond conformations by means of quantum-chemical computations,” *J. Mol. Model.*, vol. 20, no. 6, p. 2281, Jun. 2014.
[100] K. Rahmelow and W. Hubner, “Infrared Spectroscopy in Aqueous Solution: Difficulties and Accuracy of Water Subtraction,” *Appl. Spectrosc.*, vol. 51, no. 2, pp. 160–170, Feb. 1997.
[101] R. Adato and H. Altug, “In-situ ultra-sensitive infrared absorption spectroscopy of biomolecule interactions in real time with plasmonic nanoantennas,” *Nat. Commun.*, vol. 4, p. 2154, Jul. 2013.
[102] D. L. Allara, “Handbook of Infrared Spectroscopy of Ultrathin Films By Valeri P. Tolstoy (St. Petersburg State University, Russia), Irina V. Chernyshova (St. Petersburg State Polytechnical University, Russia), and Valeri A. Skryshevsky (Kyiv National Taras Shevchenko University, Ukraine). John Wiley and Sons, Inc.: Hoboken. 2003. xxvi + 710 pp. $250.00. ISBN 0-471-35404-X.,” *J. Am. Chem. Soc.*, vol. 126, no. 47, pp. 15633–15634, Dec. 2004.
[103] F. Neubrech, C. Huck, K. Weber, A. Pucci, and H. Giessen, “Surface-Enhanced Infrared Spectroscopy Using Resonant Nanoantennas,” *Chem. Rev.*, vol. 117, no. 7, pp. 5110–5145, Apr. 2017.
[104] Y. Huang, X. Duan, and C. M. Lieber, “Nanowires for Integrated Multicolor Nanophotonics,” *Small*, vol. 1, no. 1, pp. 142–147, Jan. 2005.

[105] M. I. Stockman, "Nanoplasmonics: past, present, and glimpse into future," *Opt. Express*, vol. 19, no. 22, pp. 22029–22106, Oct. 2011.
[106] S. R. Gallagher, "Quantification of DNA and RNA with absorption and fluorescence spectroscopy," *Curr. Protoc. Cell Biol.*, vol. Appendix 3, p. Appendix 3D, May 2001.
[107] B. R. Masters, "Principles of Fluorescence Spectroscopy, Third Edition," *J. Biomed. Opt.*, vol. 13, no. 2, p. 29901, Mar. 2008.
[108] K. Kneipp *et al.*, "Single Molecule Detection Using Surface-Enhanced Raman Scattering (SERS)," *Phys. Rev. Lett.*, vol. 78, no. 9, pp. 1667–1670, Mar. 1997.
[109] R. Adato, S. Aksu, and H. Altug, "Engineering mid-infrared nanoantennas for surface enhanced infrared absorption spectroscopy," *Mater. Today*, vol. 18, no. 8, pp. 436–446, Oct. 2015.
[110] W. L. Barnes, A. Dereux, and T. W. Ebbesen, "Surface plasmon subwavelength optics," *Nature*, vol. 424, no. 6950, pp. 824–830, Aug. 2003.
[111] V. F. Gili *et al.*, "Metal–dielectric hybrid nanoantennas for efficient frequency conversion at the anapole mode," *Beilstein J. Nanotechnol.*, vol. 9, pp. 2306–2314, Aug. 2018.
[112] M. S. Tame, K. R. McEnery, Ş. K. Özdemir, J. Lee, S. A. Maier, and M. S. Kim, "Quantum plasmonics," *Nat. Phys.*, vol. 9, no. 6, pp. 329–340, Jun. 2013.
[113] S. A. Maier, *Plasmonics: Fundamentals and Applications*. Springer Science & Business Media, 2007.
[114] F.-P. Schmidt, H. Ditlbacher, U. Hohenester, A. Hohenau, F. Hofer, and J. R. Krenn, "Universal dispersion of surface plasmons in flat nanostructures," *Nat. Commun.*, vol. 5, p. 3604, Apr. 2014.
[115] R. F. Aroca, D. J. Ross, and C. Domingo, "Surface-Enhanced Infrared Spectroscopy," *Appl. Spectrosc.*, vol. 58, no. 11, p. 324A–338A, Nov. 2004.
[116] S. Nie and S. R. Emory, "Probing Single Molecules and Single Nanoparticles by Surface-Enhanced Raman Scattering," *Science*, vol. 275, no. 5303, pp. 1102–1106, Feb. 1997.
[117] H. D. Wanzenböck, B. Mizaikoff, N. Weissenbacher, and R. Kellner, "Multiple internal reflection in surface enhanced infrared absorption spectroscopy (SEIRA) and its significance for various analyte groups," *J. Mol. Struct.*, vol. 410–411, pp. 535–538, Jun. 1997.
[118] K. Ataka and J. Heberle, "Biochemical applications of surface-enhanced infrared absorption spectroscopy," *Anal. Bioanal. Chem.*, vol. 388, no. 1, pp. 47–54, May 2007.
[119] A. Hartstein, J. R. Kirtley, and J. C. Tsang, "Enhancement of the Infrared Absorption from Molecular Monolayers with Thin Metal Overlayers," *Phys. Rev. Lett.*, vol. 45, no. 3, pp. 201–204, Jul. 1980.
[120] D. Enders and A. Pucci, "Surface enhanced infrared absorption of octadecanethiol on wet-chemically prepared Au nanoparticle films," *Appl. Phys. Lett.*, vol. 88, no. 18, p. 184104, May 2006.
[121] H. Wang, J. Kundu, and N. J. Halas, "Plasmonic nanoshell arrays combine surface-enhanced vibrational spectroscopies on a single substrate," *Angew. Chem. Int. Ed Engl.*, vol. 46, no. 47, pp. 9040–9044, 2007.
[122] F. Neubrech, A. Pucci, T. W. Cornelius, S. Karim, A. García-Etxarri, and J. Aizpurua, "Resonant Plasmonic and Vibrational Coupling in a Tailored Nanoantenna for Infrared Detection," *Phys. Rev. Lett.*, vol. 101, no. 15, p. 157403, Oct. 2008.
[123] R. Adato *et al.*, "Ultra-sensitive vibrational spectroscopy of protein monolayers with plasmonic nanoantenna arrays," *Proc. Natl. Acad. Sci. U. S. A.*, vol. 106, no. 46, pp. 19227–19232, Nov. 2009.
[124] K. Ataka, S. T. Stripp, and J. Heberle, "Surface-enhanced infrared absorption spectroscopy (SEIRAS) to probe monolayers of membrane proteins," *Biochim. Biophys. Acta BBA - Biomembr.*, vol. 1828, no. 10, pp. 2283–2293, Oct. 2013.
[125] A. Schwaighofer *et al.*, "Time-Resolved Surface-Enhanced IR-Absorption Spectroscopy of Direct Electron Transfer to Cytochrome c Oxidase from R. sphaeroides," *Biophys. J.*, vol. 105, no. 12, pp. 2706–2713, Dec. 2013.
[126] M. Osawa, "Surface-Enhanced Infrared Absorption," in *Near-Field Optics and Surface Plasmon Polaritons*, S. Kawata, Ed. Berlin, Heidelberg: Springer Berlin Heidelberg, 2001, pp. 163–187.

[127] "Molecular Simulation of Drug-Receptor-Interactions | Zuse Institute Berlin (ZIB)." [Online]. Available: http://www.zib.de/de/projects/molecular-simulation-drug-receptor-interactions. [Accessed: 18-Oct-2018].
[128] Y. Nishikawa, T. Nagasawa, K. Fujiwara, and M. Osawa, "Silver island films for surface-enhanced infrared absorption spectroscopy: effect of island morphology on the absorption enhancement," *Vib. Spectrosc.*, vol. 6, no. 1, pp. 43–53, Oct. 1993.
[129] A. Fasasi, P. R. Griffiths, and L. Scudiero, "Surface-enhanced infrared absorption (SEIRA) of adsorbates on copper nanoparticles synthesized by galvanic displacement," *Appl. Spectrosc.*, vol. 65, no. 7, pp. 750–755, Jul. 2011.
[130] G. Fahsold, A. Pucci, and K.-H. Rieder, "Growth of Fe on MgO(001) studied by He-atom scattering," *Phys. Rev. B*, vol. 61, no. 12, pp. 8475–8483, Mar. 2000.
[131] A. Cerea *et al.*, "Modified three-dimensional nanoantennas for infrared hydrogen detection," *Microelectron. Eng.*, vol. 162, pp. 105–109, Aug. 2016.
[132] C. D'Andrea *et al.*, "Optical Nanoantennas for Multiband Surface-Enhanced Infrared and Raman Spectroscopy," *ACS Nano*, vol. 7, no. 4, pp. 3522–3531, Apr. 2013.
[133] C. Huck *et al.*, "Surface-Enhanced Infrared Spectroscopy Using Nanometer-Sized Gaps," *ACS Nano*, vol. 8, no. 5, pp. 4908–4914, May 2014.
[134] O. Limaj *et al.*, "Infrared Plasmonic Biosensor for Real-Time and Label-Free Monitoring of Lipid Membranes," *Nano Lett.*, vol. 16, no. 2, pp. 1502–1508, Feb. 2016.
[135] D. Dregely, F. Neubrech, H. Duan, R. Vogelgesang, and H. Giessen, "Vibrational near-field mapping of planar and buried three-dimensional plasmonic nanostructures," *Nat. Commun.*, vol. 4, p. 2237, 2013.
[136] U. Fano. "Effects of configuration interaction on intensities and phase shifts". *Physical Review*, vol 124, no 6, pp 1866–1878, Dec 1961.
[137] J. Vogt, C. Huck, F. Neubrech, A. Toma, D. Gerbert, and A. Pucci, "Impact of the plasmonic near- and far-field resonance-energy shift on the enhancement of infrared vibrational signals," *Phys. Chem. Chem. Phys.*, vol. 17, no. 33, pp. 21169–21175, Aug. 2015.

Chapter 2

Chapter 2 focuses on the fabrication and characterization of the devices used during the PhD for Collective Enhanced InfraRed Absorption (CEIRA) microscopy. The design of the nanoantennas array and the main fabrication steps for making such structures will be illustrated in sections 2.1 and 2.2. Section 2.3 will be dedicated to the morphological characterization of the devices, while the FTIR response will be presented in section 2.4. Lastly, in section 2.5, the characterization of chip response for protein analysis and the interpretation of their enhanced signals will be addressed.

Design and fabrication of CEIRA devices have been carried out in collaboration with the Plasmon Nanotechnology Research Laboratory at the Istituto Italiano di Tecnologia (IIT) of Genova. The morphological characterization was done at the Nanoinnovation Lab at Elettra-Sincrotrone Trieste. Finally, the device investigation in terms of resonance response and protein signal enhancement was done at Chemical-Life Science branch of the SISSI beamline at Elettra-Sincrotrone Trieste [1].

2.0 Introduction

The aim of my PhD project is to apply the CEIRA approach for conformational studies of proteins of biological and biomedical relevance in physiological environment, overcoming the sensitivity issue of conventional FTIR spectroscopy.

In order to cover the entire Mid-IR domain, we proposed to develop *ad-hoc* substrates suitable for CEIRA microscopy made by four adjacent arrays of nanoantenna having resonance curves centered respectively at: 2900 cm^{-1} (Region of Interest-L0, ROI-L0), 1600 cm^{-1} (ROI-L1), 1400 cm^{-1} (ROI-L2) and 1100 cm^{-1} (ROI-L3). As introduced in Chapter 1, CEIRA technique guarantees an enhancement of Mid-IR signal of several orders of magnitude exploiting the coupling of the resonances provided by ordered engineered nanoantenna arrays with the vibrational signals of target molecules [2]. The importance of the matching between resonance response of nanoantenna assemblies and the target vibrational fingerprint of molecules has been highlighted in section 1.3.4. In Figure 2.1 the overlap between the measured resonance responses of the four different arrays and the

signals of a model protein is presented. In particular, ROI-L0 is centered on the vibrational signals of aliphatic moieties, belonging mainly to amino acid side chains of proteins. On the contrary, ROI-L1 deals the main protein backbone signals: Amide I and Amide II bands, which are sensitive for the conformational details of secondary structure of proteins. ROI-L2 is designed for enhancing Amide III signals, while ROI-L3 refers to a spectral region relevant for phosphates and sugar moieties, which can be used to follow eventual post-translational modification such as protein phosphorylation and glycosylation.

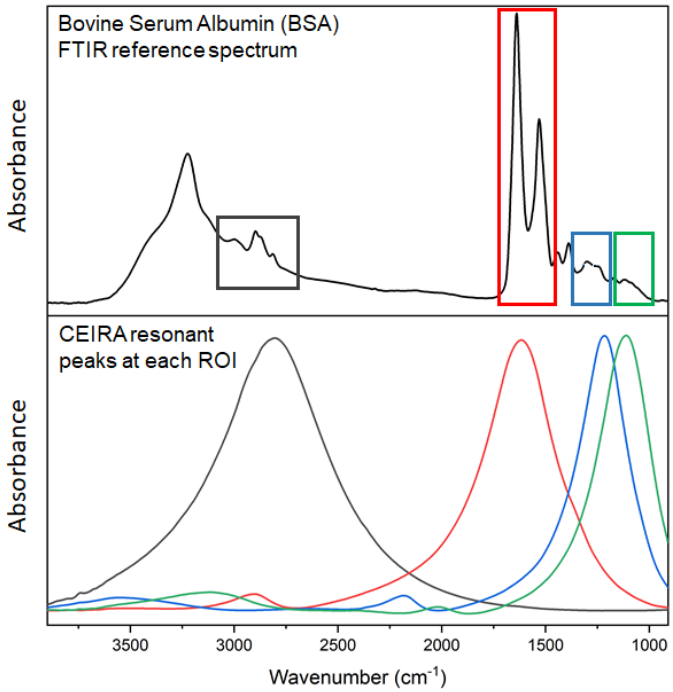

Figure 2.1: Ideal overlap between the antenna resonant response and protein signals in the different Regions of Interest (ROIs). Upper panel: the reference FTIR spectrum of a model protein (bovine serum albumin – BSA) is reported. Squares highlight the four ROIs. Lower panel: measured resonant response for each ROI. In the figure curves and boxes follow this color code: Black – ROI-L0; Red – ROI-L1; Blue – ROI-L2; Green – ROI-L3.

Since the most relevant vibrational signals of proteins for secondary structure analysis, are those of amide bands (ROI-L1), in the present thesis we focus our attention mainly on this region. As we have seen in the introduction, the combination of nanoantennas in ordered arrays give rise to a collective response of the entire nanoantenna assembly in the so-called CEIRA approach. Moreover, also the dipole effect between neighboring antennas separated by a nanometer inter-particle gap provides an increased near-field enhancement. Hence, we decided to build up plasmonic sensors which take advantage from both these phenomena: ordered assemblies of coupled antennas were engineered. Due to the ordered nature of our

arrays we defined our technique Collective Enhanced IR Absorption (CEIRA) microscopy. The basic layout of our devices comprehends the repetition of 4 arrays (50x50 μm^2 in size). The arrays were on purpose designed in order to have a small active area but, at the same time, to cover the entire Mid-IR spectral range. The proposed strategy has the following advantages:

1. It allows to reduce the device fabrication time increasing the nanostructure reliability, in terms of device response;
2. It allows to use a very low amount of protein material to obtain an enhanced signal. This aspect is of paramount importance for proteins produced at high-purity and very low amounts, i.e. the typical situation for protein of biomedical relevance.

In order to discriminate between the responses of the different arrays, the capability to focus at 50x50 μm^2 or better the IR light is mandatory, and therefore the need of FTIR microscopy, possibly coupled with InfraRed Synchrotron Radiation (IR-SR) to push further the lateral resolution. A schematic representation of the final plasmonic device is reported in Figure 2.2. In the text, the terms "CEIRA devices" and "plasmonic devices" are used as synonymous.

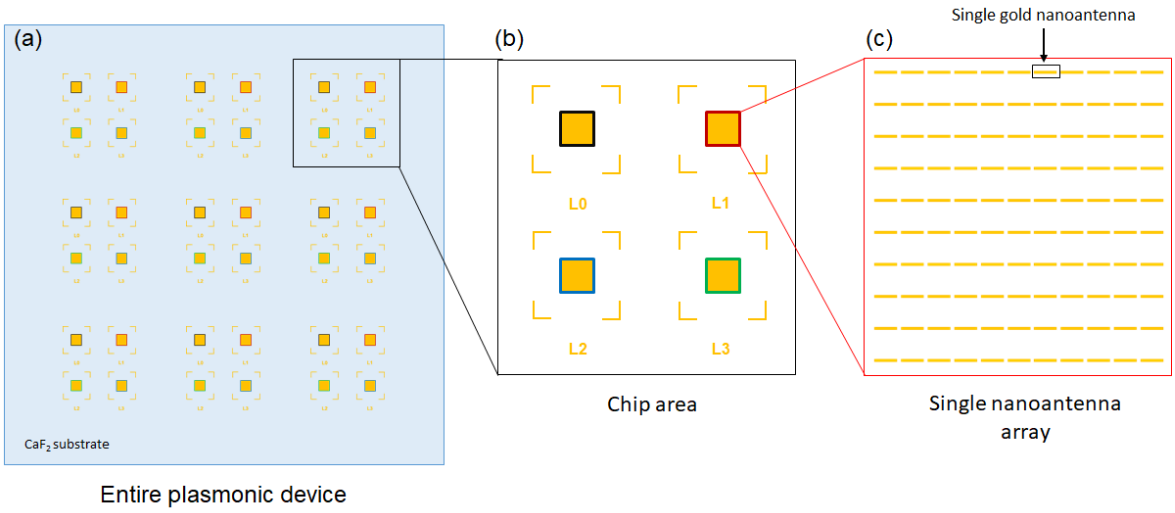

Figure 2.2: Scheme of the CEIRA device. (a) Entire Plasmonic device, composed by nine repeated chips onto the CaF_2 substrate window. (b) Single chip is made by 4 adjacent nanoantennas arrays. Each array has a resonant response in a specific ROI, accordingly to the label (from L0, to L3). (c) Single nanoantenna array. For more details on antenna disposition, which defines the array geometry, see section 2.1. A single gold nanoantenna is highlighted in the figure.

During the three years of the PhD activity, we developed two different array geometries. At first, a "conventional" geometry for the nanoantennas assemblies was adopted, where the

repeated arrays were composed by linear chains of nanorod shaped antennas spaced by a nanometer gap (see the scheme in the Figure 2.3 a-b). Hereafter we refer to these structures as *line-to-line* devices. Then, we developed a second geometry, which we called *cross*-geometry, where nanoantenna assemblies were composed by antennas disposed to form centrosymmetric crosses (see the scheme in Figure 2.3 c-d). Since *cross*-devices are an evolution of the basic design adopted for the first generation, the fabrication steps for both geometries are the same. The characterization of both devices is discussed in parallel in this chapter and principal advantages offered by second generation devices are highlighted in the section 2.4.

2.1 Structure design and numerical simulations

The nanoantennas design and the plasmonic material employed for their fabrication strongly influence the absorption cross-section of nanostructures and, consequently, their signal enhancement capabilities [2]. As stated, in order to promote the so-called "collective enhancement", nanoantennas were engineered in periodic arrays [3]. Furthermore, nanoantennas were separated by nanometer gap in order to take advantage from both near-field and far-field nanoantennas interactions (see section 1.3.3 "*Interaction between Resonant Nanoantennas: the Collective enhancement*"). Linear repetitions of nanorod shaped antennas are the model systems usually adopted for both resonant SEIRA and CEIRA experiments, due to their high performances and simplicity [4]–[6]. As for the material, gold is the most commonly chosen for such nanostructures, since it combines three important advantages: (i) its dielectric function follows nearly perfect Drude-type behavior in the IR regime with a low damping rate, leading to sharp plasmonic resonances [7]; (ii) its superior stability under different environmental and ambient conditions; (iii) its biocompatibility and the wide range of functionalization processes that have been set for gold surfaces [8], [9]. A material transparent both to both visible and IR light was needed as substrate for the chip fabrication. Moreover, since our final goal is to measure proteins in aqueous solution, an insoluble material was required as well. Among others, we chose calcium fluoride (CaF_2) since it represents a good compromise in terms of IR transparency (nearly 95% between 66666-1110 cm^{-1}, depending on the window thickness) and low solubility in water (0.00151g/100g H_2O at 20°C) [10].

A schematic view of a CEIRA nanoantenna array and the gap regions is presented in Figure 2.3. In the upper panels the *line-to-line* geometry is presented, while the *cross*-geometry is

reported in the lower panels. Figure 2.3 shows an infinite repetition of dipole antennas with nanometric height (h) and width (w) and micrometric length (L), coupled end-to-end through a fixed gap (G_x) along their long-axis direction. For *line-to-line* geometry arrays, G_y is the spacing along the antenna short-axis direction, while for *cross*-geometry arrays this parameter was defined as the sum of the distances between the repeated elements and the nanoantennas length (L).

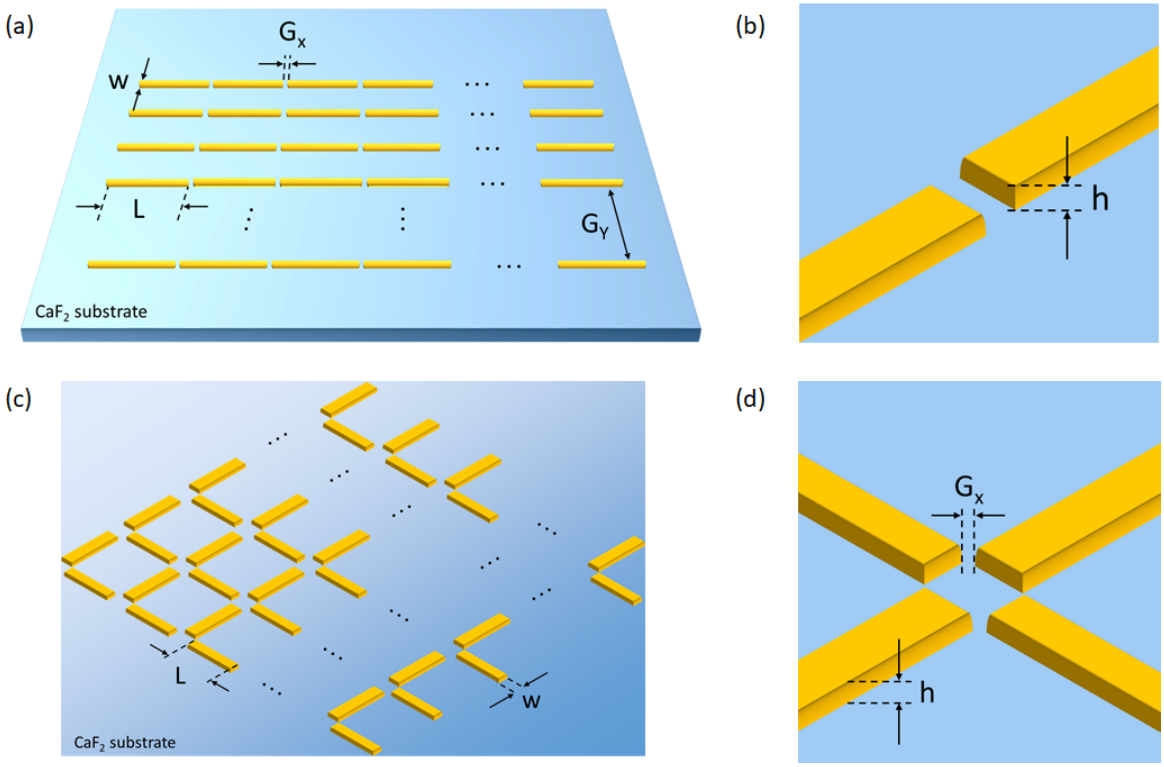

Figure 2.3: Plasmonic nanostructure and CEIRA array design **a)** Schematic view of antennas disposition in *line-to-line* devices on the CaF_2 windows. **b)** Sketch showing the details of the gap between two adjacent antennas. **c)** Schematic view of antennas disposition in *cross*-geometry devices on the CaF_2 windows. **b)** Sketch showing the details of the gap between antennas.

It is important to remember that the resonance tuning of the nanorod arrays is governed by the nanaontenna aspect ratio, in particular by varying the length of the antennas. In order to estimate the best geometrical parameters in terms of spectral matching between antenna resonances and different protein ROIs, EM simulations have been performed through rigorous coupled wave analysis (Synopsys' Optical Solutions, RSoft). The present approach has been specifically developed for providing the transmission/reflection/absorption spectra of periodic structures. For the purpose of the simulations, the refractive index of CaF_2 substrate was assumed constant at n_{sub} = 1.38 in the Mid-IR regime (wavelength between 3.1 to 9.1 µm), while the complex permittivity of gold was taken from literature [11].

We considered nanoantenna with the following geometrical parameters: for *line-to-line* G_x was considered about 50 nm, and G_y 5 µm, while for *cross*-geometry G_x was considered 25 nm. Antenna dimensions were kept constant in height and with (80nm) while the nanoantenna length was varied in the range between 2-10 µm to tune the array resonance peak position in the different ROIs. All the nanoantenna sharp edges were rounded with a radius of curvature of 20 nm, except the ones perpendicular to the substrate in correspondence of the gaps, where a curvature of 40 nm was introduced to better resemble the fabricated structures. The response of an infinite nanoantenna array was obtained by imposing appropriate periodic boundary conditions in the plane of the array.

For the input illumination conditions, we employed a linearly polarized plane wave. The plane of incidence was tilted by an angle $\varphi = 45°$ with respect to the nanoantenna long axis, while the angle of incidence in this plane was set to $\theta = 18°$ (in accordance with the mean incidence angle of the Cassegrain objective used in FTIR microscopy measurements). The polarization state of the wave was then set taking into account the input polarization (parallel to the nanoantenna long axis) impinging on the Cassegrain objective. A scheme representing the input illumination conditions is reported in Figure 2.4.

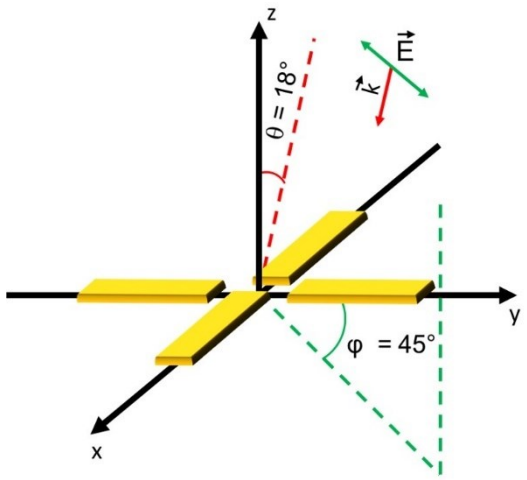

Figure 2.4: Schematic representation of the input illumination conditions employed in the simulations. The vector k points the direction of incidence light. The angle of incidence was set to $\theta = 18°$. The electric field of the linearly polarized plane wave (E) was tilted by an angle $\varphi = 45°$ with respect to the nanoantenna long axis.

We found that this type of illumination well reproduces the transmission response of the arrays. Parameters provided by simulations were optimized during the fabrication procedure.

2.2 Fabrication of CEIRA device

Arrays of rod-shaped gold nanoantennas were fabricated by electron beam lithography (EBL) on CaF_2 windows. EBL is a powerful tool for the fabrication of nanostructures, which cannot be realized by conventional photolithography techniques due to their nanometer dimensions. State of the art EBL systems can reach resolution down to sub-10 nm using an accelerated beam of electrons for patterning structures on substrates coated with an electron sensitive polymeric material, called resist. The electron beam, acting as a sort of pencil on the resist film, introduces changes on its molecular structure and solubility. After electron exposure, the resist is developed in a suitable solvent to selectively dissolve either the exposed or unexposed areas with respect to the resist characteristics (positive or negative). Afterward, the resist layer remaining onto the substrate can be used as a mask or a template for either transferring the pattern into the substrate or depositing metallic films onto it.

The fabrication of the devices employed in this thesis was carried out in a cleanroom environment (ISO6, or equivalently class 1000) to prevent sample contamination from environmental pollutants. A schematic representation of the various fabrication steps is provided in Figure 2.5, and briefly described in the following. As first step, 1x1 cm^2 CaF_2 (100) windows 1 mm thick were cleaned briefly with a plasma cleaner (100W, 100% O_2, 1min) and then in an ultrasonic bath of acetone and isopropyl alcohol (Figure 2.5a). A 160-nm thick poly(methyl methacrylate) (PMMA) layer was spin-coated at 1800 rpm on the CaF_2 windows. Subsequently, annealing was performed at 180 °C for 7 minutes (Figure 2.5b). In order to prevent charging effects during the electron beam exposure, a 10-nm thick Al layer was thermally evaporated on the PMMA surface (Figure 2.5c). Electron beam direct-writing of the nanoantenna arrays was carried out using an ultrahigh resolution Raith 150-Two e-beam lithography and imaging system, which enables nanostructures fabrication with a resolution down to few nanometers (Figure 2.5d). During exposure of the nanoantennas, a beam energy of 20 keV was employed. The exposure dose, on average 550 $\mu C/cm^2$, was adjusted with respect the EBL status. The Al conductive layer was then removed by a 1 M KOH solution, and consequently the exposed resist was developed in a solution of methyl isobutyl ketone (MIBK)/isopropanol (IPA) (1:3) for 30 seconds (Figure 2.5e). Electron beam evaporation in a high vacuum chamber (base pressure 10^{-7} mbar) was exploited to produce first a 5-nm adhesion layer of titanium and an 80-nm Au film, with a 0.3 Å/s deposition rate (Figure 2.5f). Finally, the unexposed resist was removed through a conventional lift-off process in hot acetone (Figure 2.5g). To improve the lift-off efficiency, sonication for 2 minutes at 40 kHz frequency was employed. The substrate was then rinsed out in IPA for 30

second and dried with a nitrogen stream. Patterned windows (Figure 2.5h) were finally O_2 plasma cleaned (power 100 W) for 10 minutes in order to remove organic residues.

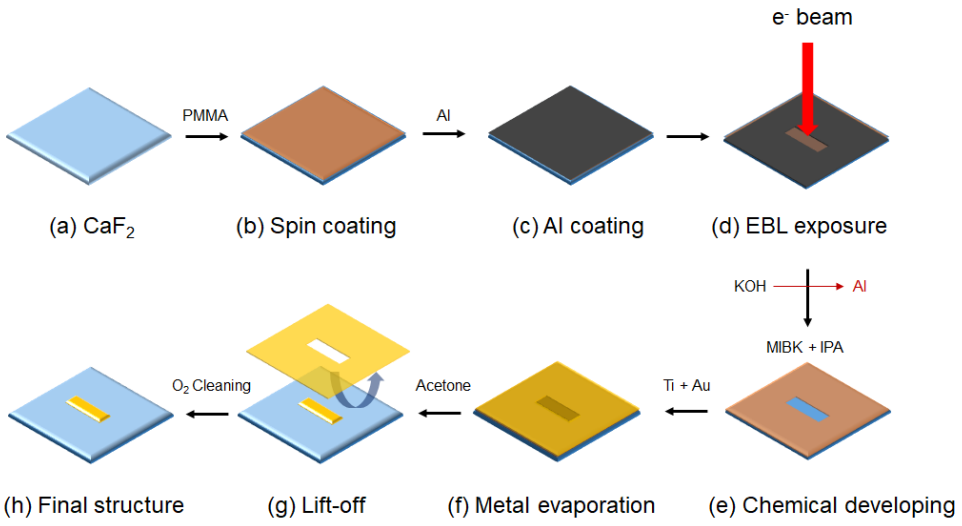

Figure 2.5: Schematic representation of the principal EBL fabrication steps.

2.3 CEIRA structures characterization

After the fabrication, the morphology of the samples was characterized by means of scanning electron microscopy (SEM) and atomic force microscopy (AFM) in order to verify the quality of the fabricated nanostructures and their final dimensions.

2.3.1 SEM imaging

A Helios NanoLab 600i scanning electron microscope (SEM) was used to acquire the images of nanoantenna arrays. In order to avoid the charging of the surface during the SEM acquisitions, a representative sample for each batch of fabrication was chosen and coated with a 10 nm layer of carbon. In Figure 2.6 the SEM images for both *line-to-line* (upper panels) and *cross* (lower panel) devices are reported. For both geometries, a zoom of their elementary cell, at the nanometer gap region, is presented. Width and length of antennas were evaluated, together with the gap dimensions. In general, dimensions of *line-to-line* antennas were slightly smaller than designed: they were 75 nm wide, with gaps of 52±3 nm. Instead, in *cross*-devices nanoantenna arrays showed dimensions consistent with the drawings. However, since the efficiency of EBL depends on the writing direction, the resulted crosses were not perfectly centrosymmetric. Nanoantennas disposed along the vertical axis of the arrays were slightly shifted with respect of the direction of the e^- beam exposure. Thus, nanogaps between antennas resulted smaller in one direction (22±3 nm) with respect to the other (27±3 nm).

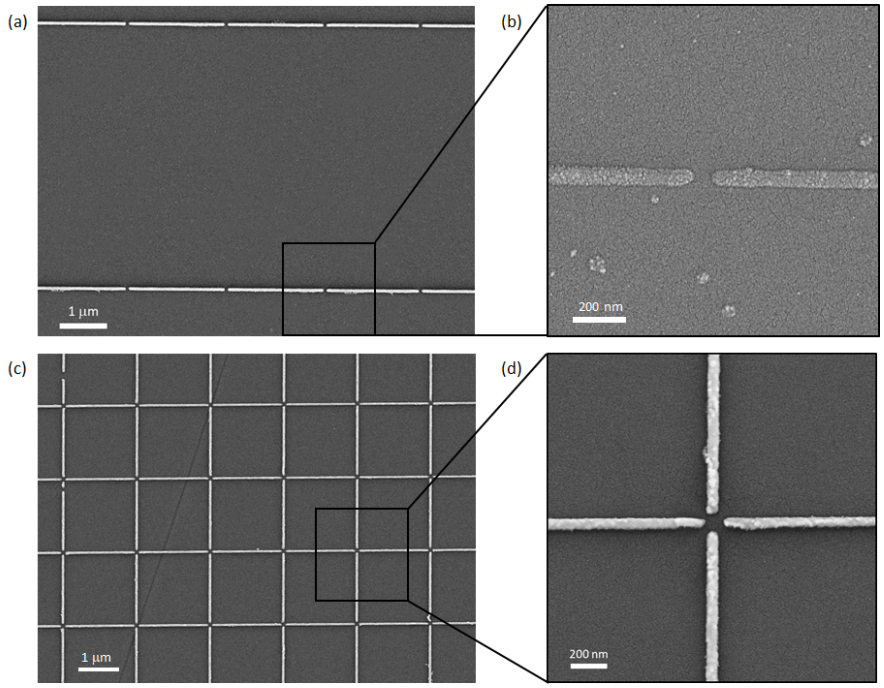

Figure 2.6: SEM images of the nanoantenna array arrangement of our CEIRA devices. (a) SEM image of chains of nanorod-shaped gold nanoantennas in *line-to-line* arrays for the L1 ROI. (b) Zoom showing the details of the gap between two coupled nanoantennas. (c) SEM overview of *cross*-geometry arrays for the L1 ROI. (d) Zoom of (c), showing the centrosymmetric crosses composed by 4 nanoantennas with the long axes perpendicular each other. As can be appreciated from the image, resulting crosses are not perfectly centrosymmetric.

As discussed before, for shifting the resonant response of the nanoantennas in the Mid-IR in order to match the four different ROIs, arrays containing antennas with different length were fabricated. Thus, also final antennas length was characterized for the four different ROIs. *Line-to-line* arrays presented antennas with the following lengths: ROI-L0, L=810nm; ROI-L1, L=1120nm; ROI-L2, L=1410nm, ROI-L3, L=1910nm. Arrays in *cross*-devices were made by antennas with a length: ROI-L0, L=1200nm; ROI-L1, L=2040nm; ROI-L2, L=2800nm, ROI-L3, L=3200nm.

2.3.2 AFM characterization

AFM images were acquired using the available microscope at the Nanoinnovation laboratory at Elettra (MFP-3D Stand Alone AFM from Asylum Research, Santa Barbara, CA). Measurements were carried out at room temperature working in dynamic tapping-mode. A commercially available silicon cantilevers (NSC19, Mikro-Masch, Poland, nominal spring constant 0.6 nN/nm) have been chosen for in air imaging. Cantilevers were used working at low oscillation amplitudes with half free amplitude set-point. High resolution images (512 x

512 pixels' frames) were acquired at 0.6÷1 lines/s scan speed. In Figure 2.7, the AFM topography of structures with both geometries is reported.

AFM results were complementary to those obtained by SEM acquisitions. Due to tip-effects, it was hard to obtain information onto the distance between nanoantennas in the nanoantenna gap region. However, such measurements allowed to make a validation of the height distribution of the fabricated structures. Antennas of *cross*-geometry arrays were more uniform in terms of dimensions than those of *line-to-line* array, which instead presented antennas with heights between 60-85 nm, while antennas in the *cross*-arrays were 80 nm high in average. Indeed, the poorer quality of *line-to-line* arrays could have been for sure improved if we kept working on this geometry. Nevertheless, we decided to focus on *cross*-geometry arrays almost at the beginning of the PhD activity, in order to be independent from light polarization, and on these devices we also concentrated the efforts for the optimization of the fabrication parameters. Overall, this accounts the higher quality of the second generation devices with respect to the first.

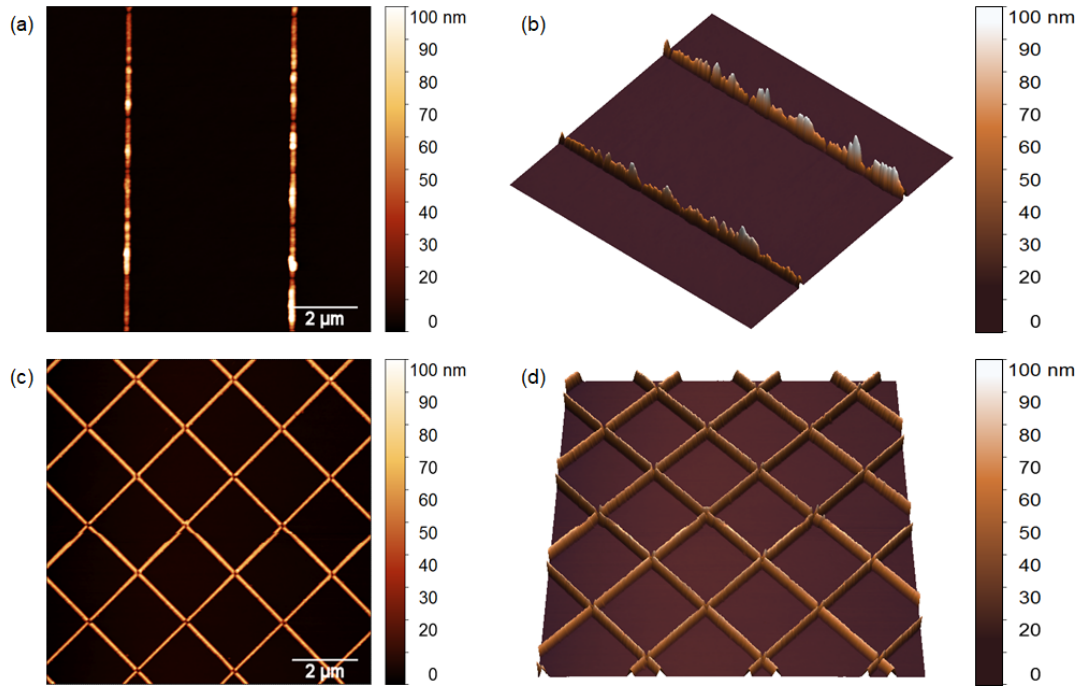

Figure 2.7: AFM images of fabricated structures. Upper panels: *Line-to-line* devices. Lower panels: *Cross*-devices. In panels (a) and (c) the topography of the sample is reported. In panel (b) and (d) the estimation of the nanoantennas' height is reported. AFM images were analyzed using Gwyddion, open-source modular program for scanning probe microscopy (SPM) data visualization and analysis.

As last remark, also AFM measurements highlighted that antennas disposed in the horizontal and vertical axes were fabricated with a different efficacy. While antennas along the horizontal axis appear uniform, those disposed along the vertical axis appear as a

combination of different strands printed perpendicular to the longitudinal axis of nanoantennas (data not shown). Indeed, EBL exposes the PMMA film by following a single writing direction, thus providing higher accuracy along its fast axis, i.e. horizontal structures, with respect to vertical ones.

2.4 FTIR characterization

CEIRA substrates were benchmarked in terms of resonance band position and electric field distribution by means of FTIR microscopy. Preliminary nano- FTIR spectroscopy experiments were also performed.

2.4.1 IR response

The resonant response of the arrays was evaluated by FTIR microscopy. Due to the dimensions of repeated arrays, 50x50 μm^2, a Vis/IR microscope was required. FTIR spectra were collected by a conventional thermal source (Globar) using a Bruker Hyperion 3000 Vis-IR microscope equipped with a mid-band HgCdTe detector having a 100 μm sensitive element, coupled with a Bruker Vertex 70v interferometer. Spectra were collected in transmission mode using 15x Schwarzschild condenser and objective and setting knife-edge apertures to 50x50 μm^2 in order to acquire the entire array of nanoantennas. Each spectrum was collected from 4000 to 800 cm^{-1} in double side, forward/backward acquisition mode with a scanner velocity of 40 kHz. For each spectrum 1024 scans were averaged with a spectral resolution of 4 cm^{-1}. Fourier transform was carried out with Mertz phase correction, Blackman-Harris 3-terms apodization function. Spectra were normalized with respect to a background signal taken on a 1 mm thick CaF_2 window.

It is widely documented that for linear nanoantennas the resonance response strongly depends on the polarization of light [4]. In fact, it is well established that the relative IR transmittance spectrum of individual metal nanowires show an intense resonance plasmonic feature only for parallel polarization (//-polarization) of the IR radiation with respect to the long antenna axis, while the nanoantenna response is suppressed for perpendicular polarization (⊥-polarization) of the incoming IR radiation [4], [12]. For the same reason, arrays composed by coupled antennas that are disposed all in the same direction require the use of //-polarized light for excite the resonance response. It is the case of *line-to-line* arrays which were used for our first CEIRA experiments. In order to characterize these structures, light polarization was set by using a Mid-IR polarizing filter before the sample. In Figure

2.8, the comparison of three acquisitions made on the same *line-to-line* array with both light polarization directions (//- blue spectrum and ⊥-red spectrum) and not-polarized light (black spectrum) are reported.

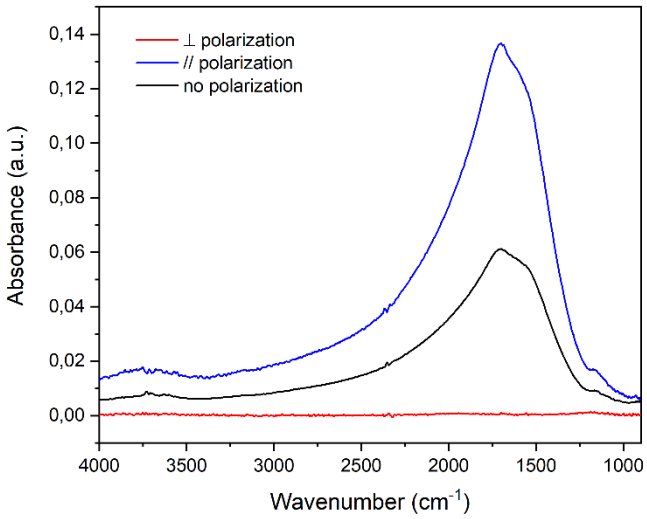

Figure 2.8: *Line-to-line* array response for ROI-L1 with respect the polarization status of input IR light. In the figure, the resonance response of L1 arrays is shown, as representative of the systems.

Results confirmed that the resonance response is completely suppressed for ⊥ light polarization. Moreover, the comparison of the substrate response with not-polarized light highlighted the needed of using //-polarization in order to collect the maximum resonance response for such structures. As mentioned before, our CEIRA substrates consist of adjacent areas containing arrays of nanoantennas with different length in order to enhance signals in the entire Mid-IR domain. The resonant response of the four arrays in a device chip (see Figure 2.2) can be seen in Figure 2.9.

Figure 2.9 clearly shows that the resonant curves are properly centered in the four ROIs and that resonant peaks are shaper and more intense at lower frequencies. Furthermore, for both L1 and L2 ROIs, the resonance curves seem to present at least two different contributions, the main of which is centered at the right position (1600 cm^{-1} for ROI-L1 and 1300 cm^{-1} for ROI-L2) and a second lower contribution red-shifted for L1 arrays and blue-shifted for L2 arrays. The resonant behavior of such nanoantenna assemblies may be due to some in-homogeneities in the fabrication, as seen by AFM. Nevertheless, the resonance intensities of the fabricated structures were consistent with those observed in literature for similar structures [5], [12].

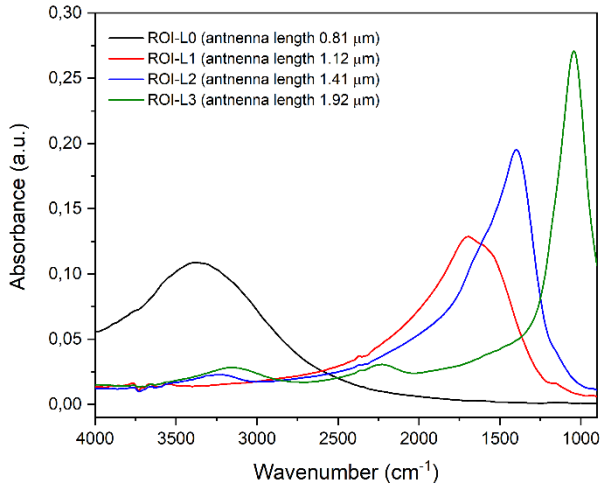

Figure 2.9: Response for the four different ROIs of one representative chip of a *line-to-line* device.

During the second year of the PhD, we developed the *cross*-design and their fabrication parameters. As already described, we designed arrays with antennas disposed in centrosymmetric crosses, with the two arms perpendicular each other, in order to take advantage from both the directions of light polarization and therefore work with not-polarized light. As a first step, the resonant response of such *cross*-devices was characterized by using polarized light. Results are reported in Figure 2.10.

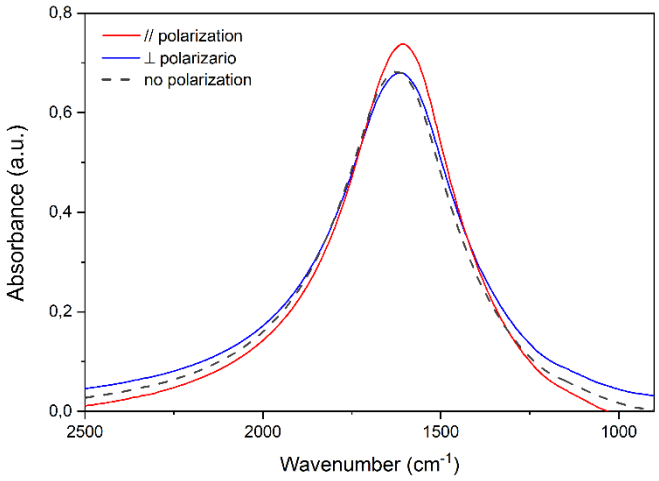

Figure 2.10: *Cross*-arrays resonance response with respect the polarization state of light in the ROI-L1.

Figure 2.10 clearly demonstrates that, with this array geometry, the plasmonic resonance of the structures is provided by both directions of light polarization. A small difference in the resonance can be appreciated between the two responses, probably related to the writing efficiency of EBL, already highlighted in sections 2.3.1 and 2.3.2. However, such differences

in resonance are negligible if compared to those of *line-to-line* arrays. Thus, *cross*-devices can be considered equally responsive to both parallel and perpendicular polarization. In Figure 2.10 the dashed line represents the resonance response of the same structure acquired with not-polarized light. As can be seen, it matches very well with former acquisitions. Results clearly show that the response of our *cross*-devices can be exploited without the need of polarizing the light.

The resonance response in the different ROIs was characterized also for *cross*-devices (see Figure 2.11).

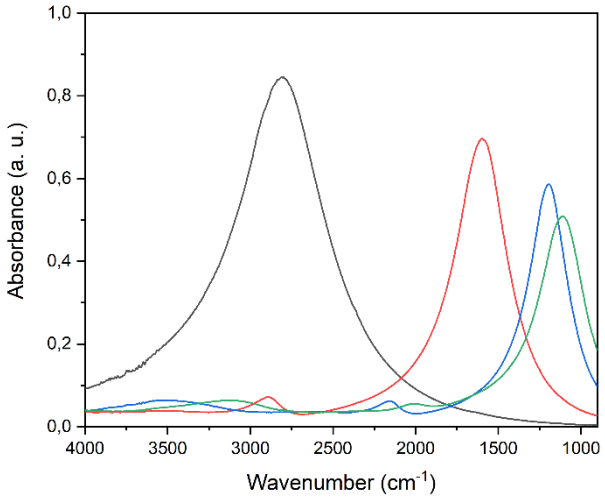

Figure 2.11: Average resonance response in the different ROIs of *cross*-devices.

As can be seen from the image, differently from *line-to-line* devices, the resonance curves decrease in intensity at lower frequencies. This happens as a consequence of the lower number of nanoantennas in the arrays for ROIs at lower wavenumbers. For reaching resonances that match such ROIs, nanoantenna length increases from 1.2 to 3.2 μm from L0 to L3 ROI. Since the density of these structures decreases in both vertical and horizontal direction by extending antenna length, while for *line-to-line* antennas the vertical spreading remains constant, the number of resonators for longer antennas in *cross*-devices dramatically decreases. As a consequence, the resonance curves appear weaker than those of shorter antennas.

Resonance response of ROI-L1 arrays of *line-to-line* and *cross*-devices is compared in Figure 2.12.

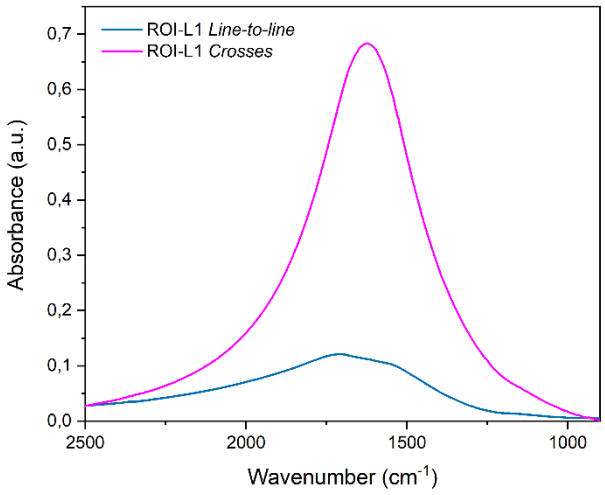

Figure 2.12: Comparison between the resonance response of *line-to-line* and *cross*-devices for ROI-L1 arrays. Measures have been acquired with perpendicular polarized light.

As can be seen, there is a substantial difference in terms of resonance intensity between the two structures. In particular, *line-to-line* structures promoted resonance curves with an intensity up to 0.13 a.u., while *cross*-geometry structures provide resonance curves up to 0.76 a.u. for the same ROI. Moreover, also the line shape of the resonance curve is different between the two systems. Indeed, the differences in terms of resonance intensity can be primarily justified considering the different geometry of the two devices. As reported in section 1.3.3, the dimensions of the nanoantennas gaps strongly influence the enhancement capabilities of the structures. As known by numerical calculations [6], [13], for nanoantennas separated by a nanometer size gaps the field intensity is enhanced of several orders of magnitude. Closer is the gap, higher is the resulting field intensity. Device of second generation were designed with a gap size smaller than those of *line-to-line* devices. Thus, the stronger response can be consequent also to this variation. Moreover, it has to be taken in account that *cross*-devices are made by a higher number of nanoantennas, due to the increased density in the vertical direction (5 µm for *line-to-line* devices, the nanoantenna length plus the nanometer gaps for *cross*-devices). Nevertheless, we cannot exclude that the different fabrication efficiency of the two structures, discussed in section 2.3, could influence the quality of the nanoantenna response as well.

Nevertheless, independently form the reasons why *cross*-devices offer better performances in respect to *line-to-line* ones, they do not require the use of polarized light, making such structures more versatile. The better performances of *cross*-devices resulted also in a better signal response for protein analyses. In Chapter 3.2, the great advantages for protein sensing

and conformational analyses offered by this second geometry will be discussed more in depth.

2.4.2 Nano-IR measurements: distribution of the electric field

The scope of nanoantennas is to convert the propagating radiation into enhanced optical fields localized on nanometric regions of their surface, the so-called "hot spots" [6], [13]. This effect arises from the resonant coupling of light with collective oscillations of conduction electrons of metal nanoparticles called localized surface plasmon resonances (LSPR). The enhancement of the EM field at this hot spot region enables a dramatic increase of the vibrational signal of molecules located therein. In literature, the different spatial distributions of the enhanced fields in the IR regime have been studied by finite difference time domain (FDTD) numerical calculations [13]. Such calculations defined that the hot spots in the IR at the fundamental resonance are localized at the antenna apexes.

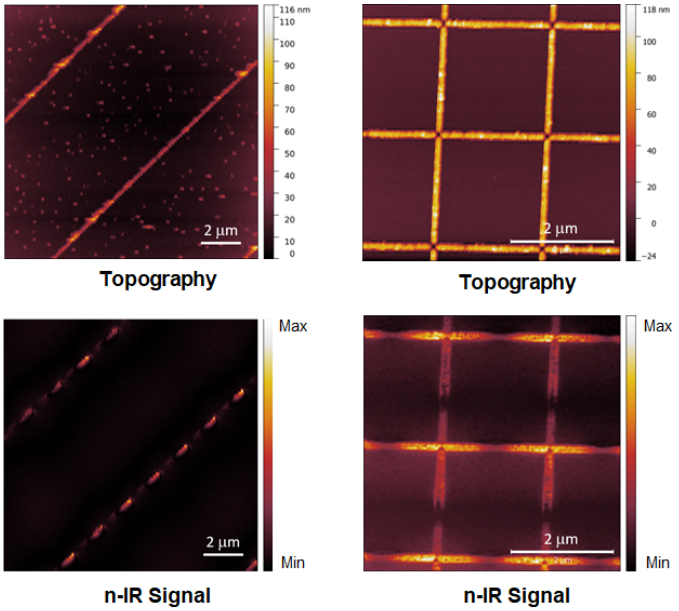

Figure 2.13: AFM-IR measurements. In the upper panel the morphological characterization of the samples is reported. In the lower panel, the EM field distribution is show using the third harmonic of the recorded optical signal.

An experimental way for mapping the EM-field distribution of the IR signal and highlighting the hot-spot position is to perform scattering-Scanning Near-field Infrared Microscopy (sSNIM) measurements [14]. To this aim we did a preliminary characterization of our device at two different facilities: *line-to-line* arrays were measured at Neaspec GmbH, Munich, using a Difference Frequency Generation (DFG) laser as a source, whereas *cross*-devices were measured at Advanced Light Source (ALS) in Berkeley, using infrared synchrotron radiation as a source. Results are reported in Figure 2.13. As can be seen, the maximum of

the mid-IR response for *line-to-line* samples is located at the nanoantennas-gaps. Instead, EM field of *cross*-devices is concentrated at the center of the crosses and follows both the nanoantennas direction. Despite this, intensity is higher for the horizontal axis.

The presented results do not aim to be a comprehensive characterization of the optical response of our devices. This is beyond the scope of this PhD thesis and it would require a continuous access to this type of instrumentation that we did not have during the PhD period. However, they are sufficiently informative on the proper plasmonic behavior of our devices.

2.5 Protein response

The results shown in section 2.4, clearly demonstrate the need to localize the target molecules in the hot-spot regions of nanoantennas arrays. All the steps for the immobilization of target proteins onto the gold nanoantennas will be discussed in the Chapter 3. In this section, we aim to present the general aspects for the interpretation of the data obtained from of our CEIRA measurements.

2.5.1 Matching between simulation and experimental data

As a first step, protein CEIRA response was simulated by numerical calculations (rigorous coupled wave analysis). We assumed to cover the CEIRA structures with a uniform layer of protein, namely BSA. Height of the layer was set to 3 nm and protein signals were simulated considering a resonator which was imposed to vibrate at the common positions of Amide I and Amide II bands. For the protein layer, we assumed that its retractation has a dispersion that can be modeled with a set of resonators with a Lorentzian linewidth at the vibrational absorption frequencies of the amide bands of BSA.

Results of the simulation on bare nanoantennas and fully protein-covered nanoantenna array (red curve) for ROI-L1 are reported in Figure 2.14a (dark gray curve), and they are compared with the corresponding experimental acquisitions in Figure 2.14b. As can be seen, experimental results match the simulations. Resonance curves are well coupled with the vibrational fingerprint of selected biomolecules. Protein fingerprints appear as peaks at the bottom of the overall resonance curves, as highlighted by black arrows in Figure 2.14. Both in simulations and real data, a red shift of ~0.12 μm (about 30 cm^{-1}) of the resonant curve is detectable as a consequence of the change of the refractive index at the array surface [15]. Such shift was detected for all the acquired ROIs, and can be considered as a marker of the protein linking to the surface of the CEIRA chip.

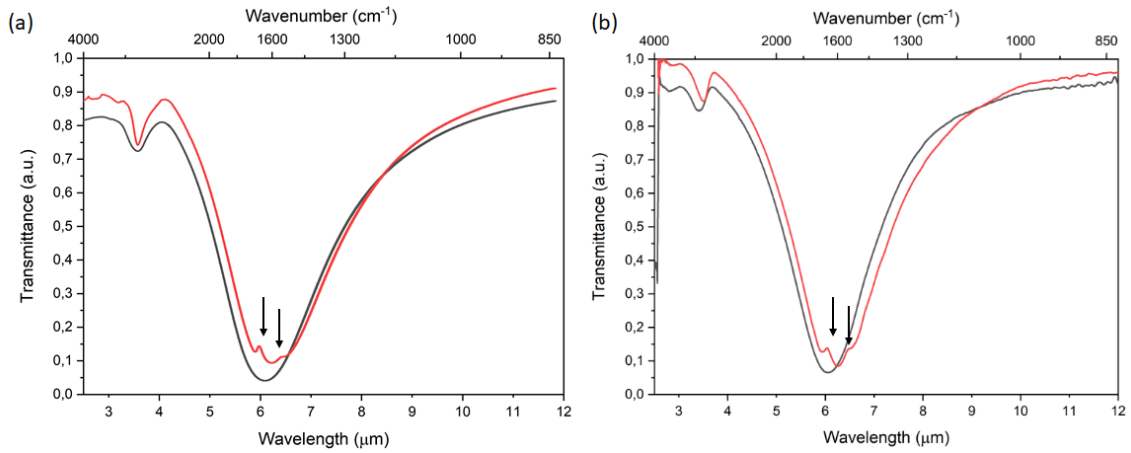

Figure 2.14: (a) FTDT simulation of the resonance response of bare antennas (dark gray spectrum) and a protein covered array (red spectrum) in the ROI-L1. (b) Corresponding experimental acquisitions for a bare array (dark gray spectrum) and the same array covered with a target protein BSA (red spectrum). In order to facilitate the matching between FTDT calculations, in this case we reported the acquisitions as transmittance spectra as a function of the light wavelength. Extra-transmittance protein peaks are highlighted by the black arrows.

2.5.2 Protein signal in the entire Mid-IR domain

As presented, the purpose of the PhD thesis was to develop a sensor which allowed a signal enhancement in different region of interest within the Mid-IR regime. Both *line-to-line* and *cross*-devices were made by chips containing adjacent arrays with nanoantennas with different length, and the resonance tuning was demonstrated.

In Figure 2.15 the resonance responses of the structures functionalized with proteins are reported in the four ROIs. As can be seen, despite antennas provided resonance tuned with respect the different vibrational features of proteins in the diverse ROIs, enhanced protein signals were detectable only in certain regions.

In particular, the detection of the Amide I and II bands in ROI-L1 is quite straightforward, since these are the strongest protein signals. In Figure 2.15 is possible to appreciate that also enhanced signals of aliphatic moieties are visible for ROI-L0. Instead for those arrays tuned for the Amide III (ROI-L2) and the ROI-L3 it was very hard to detect any protein feature. This can be because, except few cases, such as for some collagen types [16], protein Amide III is much weaker than Amide I and II bands. Despite the plasmonic enhancement, such weak signals are very hard to extract from resonance curves. Finally, as mentioned before, ROI-L3 can be used for monitoring the phosphorylation and glycosylation status of target proteins. However, the model proteins employed in our studies, as well as the ones of biomedical relevance, do not present such post translational modifications (PTMs).

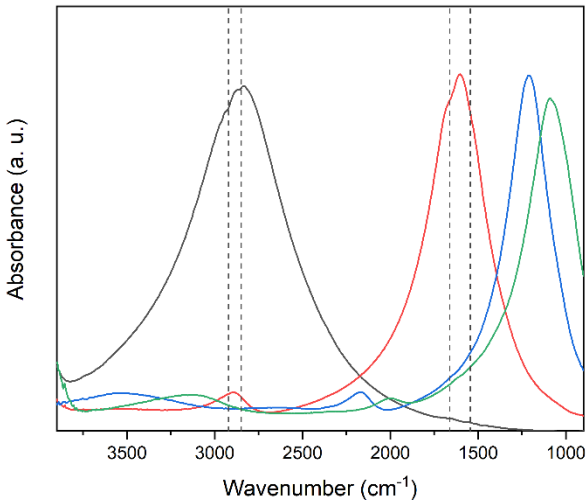

Figure 2.15: Representative CEIRA response in the different ROIs of protein-coated arrays in dried conditions. Dashed lines highlight the vibrational features of proteins visible in ROI-L0 and L1.

Despite the limitations here reported, structures tuned in ROIs L2 and L3 may be used for detecting other bimolecular species, such as nucleic acids and sugars. Concerning ROI-L0, in literature it has been reported that nanostructures that provide resonances in this region can be used for probing the phospholipid membrane formation and accumulation onto the CEIRA structures [17]. In the present thesis, structures tuned in the ROI-L0 were used for follow the formation of thiol Self Assembled Monolayers (SAMs) formation (see Chapter 3.1).

In conclusion, as expected, the most responsive region for protein sensing and analysis is the ROI-L1, which contains the protein spectral band more tightly related to protein conformation: the Amide I band.

2.5.3 Protein signal extraction

As we have seen, protein signals appear as troughs in the resonance curves. In order to extract protein signals, CEIRA resonance curves of functionalized devices were subtracted by the antenna resonance contribution, as reported in Figure 2.16.

It is important to notice that, it is not possible to use as reference signal that of bare antennas. As we have shown in section 2.5.1, the interaction between the protein and the nanoantennas arrays promote a variation in their resonance response, both in terms of resonance positions and bandwidth. Thus, resonance curves were fitted by using a polynomial spline as shown in Figure 2.16. The direct subtraction of the spline generates negative "bands" of the protein-like signals, that can be plotted in a more common representation by multiplying by "-1". In

Figure 2.16, the reconstructed absorbance spectrum of the protein is reported (lower dotted spectrum).

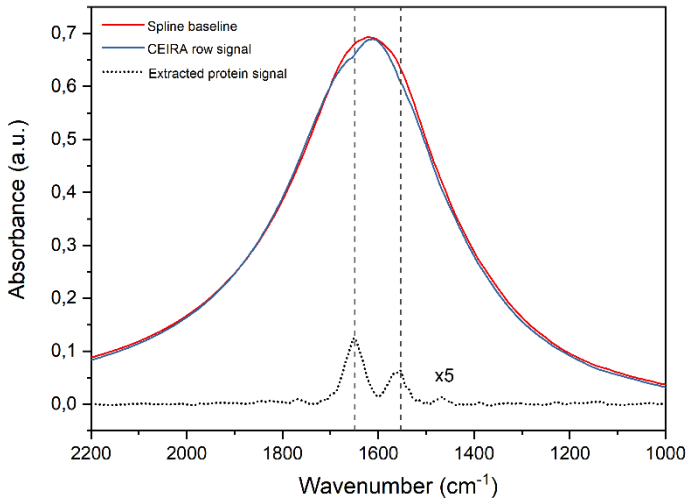

Figure 2.16: Extraction of protein nanolayer vibrational fingerprint: CEIRA resonance raw signals of functionalized devices were baseline corrected by fitting a polynomial spline curve. The reverse of the resulting difference spectra shows the enhanced fingerprint of protein layers.

2.5.4 Plasma cleaning: how to reuse the structures

Due to the cost and the time-consuming production of the devices, we defined a protocol for re-using them for multiple experiments, exploiting the durability of the gold nanostructures.

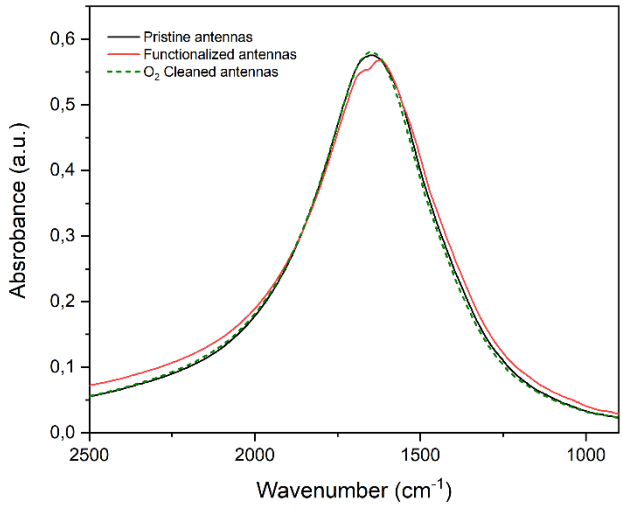

Figure 2.17: Characterization of the cleaned nanoantennas arrays: CEIRA resonance pristine antenna (black line) and a protein functionalized one (red line) are matched against a spectrum acquired on the very same antenna array after O_2 plasma cleaning (dashed green line). Figure demonstrates that it is possible to clean our chips after the functionalization procedure without modifying their resonance response.

The protocol relies on the removal of the organic part from the gold surface. To this aim, plasma cleaning using oxygen gas 100% at 100 W was applied for 15 minutes for burning the organic matter on top of the gold antennas. A subsequent response characterization of the structures confirmed that such cleaning both completely removed the organic material on the surface of the devices and did not modified their resonant response (Figure 2.17). Thus, our fabricated structures not only are suitable for one experiment, but can be re-used up to 5 experiments.

References

[1] S. Lupi *et al.*, "Performance of SISSI, the infrared beamline of the ELETTRA storage ring," *JOSA B*, vol. 24, no. 4, pp. 959–964, Apr. 2007.

[2] F. Neubrech, C. Huck, K. Weber, A. Pucci, and H. Giessen, "Surface-Enhanced Infrared Spectroscopy Using Resonant Nanoantennas," *Chem. Rev.*, vol. 117, no. 7, pp. 5110–5145, Apr. 2017.

[3] R. Adato, S. Aksu, and H. Altug, "Engineering mid-infrared nanoantennas for surface enhanced infrared absorption spectroscopy," *Mater. Today*, vol. 18, no. 8, pp. 436–446, Oct. 2015.

[4] F. Neubrech, A. Pucci, T. W. Cornelius, S. Karim, A. García-Etxarri, and J. Aizpurua, "Resonant Plasmonic and Vibrational Coupling in a Tailored Nanoantenna for Infrared Detection," *Phys. Rev. Lett.*, vol. 101, no. 15, p. 157403, Oct. 2008.

[5] R. Adato *et al.*, "Ultra-sensitive vibrational spectroscopy of protein monolayers with plasmonic nanoantenna arrays," *Proc. Natl. Acad. Sci. U. S. A.*, vol. 106, no. 46, pp. 19227–19232, Nov. 2009.

[6] C. Huck *et al.*, "Surface-Enhanced Infrared Spectroscopy Using Nanometer-Sized Gaps," *ACS Nano*, vol. 8, no. 5, pp. 4908–4914, May 2014.

[7] J. Trollmann and A. Pucci, "Infrared Dielectric Function of Gold Films in Relation to Their Morphology," *J. Phys. Chem. C*, vol. 118, no. 27, pp. 15011–15018, Jul. 2014.

[8] E. Pensa *et al.*, "The Chemistry of the Sulfur–Gold Interface: In Search of a Unified Model," *Acc. Chem. Res.*, vol. 45, no. 8, pp. 1183–1192, Aug. 2012.

[9] D. Enders and A. Pucci, "Surface enhanced infrared absorption of octadecanethiol on wet-chemically prepared Au nanoparticle films," *Appl. Phys. Lett.*, vol. 88, no. 18, p. 184104, May 2006.

[10] "Database of Optical Materials." [Online]. Available: https://www.crystran.co.uk/optical-materials. [Accessed: 26-Oct-2018].

[11] R. L. Olmon *et al.*, "Optical dielectric function of gold," *Phys. Rev. B*, vol. 86, no. 23, p. 235147, Dec. 2012.

[12] F. Neubrech *et al.*, "Infrared Optical Properties of Nanoantenna Dimers with Photochemically Narrowed Gaps in the 5 nm Regime," *ACS Nano*, vol. 6, no. 8, pp. 7326–7332, Aug. 2012.

[13] C. D'Andrea *et al.*, "Optical Nanoantennas for Multiband Surface-Enhanced Infrared and Raman Spectroscopy," *ACS Nano*, vol. 7, no. 4, pp. 3522–3531, Apr. 2013.

[14] P. Alonso-González *et al.*, "Experimental verification of the spectral shift between near- and far-field peak intensities of plasmonic infrared nanoantennas," *Phys. Rev. Lett.*, vol. 110, no. 20, p. 203902, May 2013.

[15] R. Adato and H. Altug, "*In-situ* ultra-sensitive infrared absorption spectroscopy of biomolecule interactions in real time with plasmonic nanoantennas," *Nat. Commun.*, vol. 4, p. 2154, Jul. 2013.

[16] M. Gąsior-Głogowska, M. Komorowska, J. Hanuza, M. Ptak, and M. Kobielarz, "Structural alteration of collagen fibres--spectroscopic and mechanical studies," *Acta Bioeng. Biomech.*, vol. 12, no. 4, pp. 55–62, 2010.

[17] O. Limaj *et al.*, "Infrared Plasmonic Biosensor for Real-Time and Label-Free Monitoring of Lipid Membranes," *Nano Lett.*, vol. 16, no. 2, pp. 1502–1508, Feb. 2016.

Chapter 3

In Chapter 2, the fabrication and optical characterizations of CEIRA devices have been presented. Chapter 3 focuses on the use of CEIRA devices for protein secondary structure analysis, both in dried and aqueous conditions. In this chapter, the results of several experiments performed exploiting CEIRA microscopy for protein studies will be presented, with the aim to prove the capability of the technique to act as an advanced tool for secondary structure analysis of protein monolayers in physiological conditions.

Our first goal will be to verify the detection capabilities of the fabricated plasmonic devices, pushing the sensitivity limit of FTIR technique down to protein monolayer. In the section 3.1 a critical presentation of the different strategies tested for protein monolayer formation will be reported, while section 3.2 will focus on the exploitation of Synchrotron Radiation Infrared (SR-IR) source for improving device sensitivity enhancing the signal to noise spectral ratio. Then, we will concentrate on the investigation of protein secondary structure by using CEIRA microscopy. Specifically, in section 3.3, we will highlight the capabilities of CEIRA microscopy to discern between two different model proteins, namely Bovine Serum Albumin (BSA) and Concanavalin A (ConA), in dried conditions. The conformational studies of the same proteins in physiological environment will be detailed in section 3.4. Finally, in section 3.5 will be presented the characterization of a protein of biomedical interest, the kinase domain of Epidermal Growth Factor Receptor (EGFR), in dried, hydrated and dynamic conditions.

For each section, a short *Introduction* to the topic is given, followed by a *Materials and Methods* subsection which sums up the experimental procedures. It follows a detailed discussion of the results (*Results and Discussion)* and their relevance is summarized in the *Conclusions*.

3.1 Controlling protein monolayer formation

Detection capabilities of CEIRA devices have been extensively demonstrated in literature with different samples and materials [1]–[3], reporting a molecular sensitivity up to attomoles [4]. The sensing capability was obtained by coupling the vibrational fingerprint of selected molecules with the strong resonant enhancement given by plasmonic nanoantenna

assemblies. As reported in Chapter 2, our CEIRA devices were fabricated following the same scheme. In section 2.4, we have demonstrated that the maximum signal enhancement offered by our CEIRA devices is concentrated at the gap between adjacent nanoantennas (see section 2.4.2 "*Nano-IR measurements: distribution of the electric field*"). Therefore, we need to anchor the analyte as close as possible to the antenna hot-spots. A good strategy is to deposit the analytes in ultrathin films onto the surface of the plasmonic device. Ultrathin films, also called nanolayers, are layers of molecules that spread from sub-monolayers to several monolayers [5], and play a fundamental role in different applications, ranging from biomaterial, biomedical or bioanalytical applications to nano-biotechnological devices and sensors [6]–[9]. Due to their importance in many fields of research and multitude of applications, different strategies for nanolayer formation have been reported in literature [10]–[13], ranging from drop casting, that exploits the *coffee ring* effect [14], to physisorption onto the substrate surface [15], to molecular immobilization via chemical modification of the surfaces [16].

In order to find the best deposition method for our purposes, we selected three among the most common strategies for protein anchoring onto the surface of our plasmonic devices: (i) film deposition by drop casting, (ii) protein physisorption and (iii) protein monolayer covalent bonding via amine coupling. Bovine Serum Albumin (BSA), a protein often employed as reference or as concentration standard in lab experiments, was selected as a model protein due to its availability and high stability in different environmental conditions. Moreover, this protein is also suitable for our purposes due to its secondary structure: BSA is mainly α helix folded, and it gives a well recognizable spectrum in the Amide I-II range. All the measurements reported in section 3.1 were performed in dried conditions, which is the easiest measurement condition for validating the obtained results. The discussion will highlight pros and cons of each investigated anchoring method.

Materials and methods

Gold nanoantennas Arrays fabrication

Ad-hoc substrates suitable for CEIRA microscopy were fabricated as described in the section 2.2 "*Fabrication of CEIRA arrays*". In these experiments, *line-to-line* arrays were used. The substrates contained regions made by tuned gold nanoantenna that allow Mid-IR signal enhancement into two Regions of Interest (ROIs): the Amide ROI, centered at 1600 cm^{-1} (ROI – L1, λ=6,2 μm) and the Lipid ROI, centered at 2900 cm^{-1} (ROI – L0, λ=3,45 μm). Final structures were linear individual nanoantennas 80 nm high and 75 nm wide. The length

of nanoantennas was 1,8 μm and 1,4 μm for ROI-L1 and L0 respectively. The gap between adjacent nanoantennas was 50 nm (G_x), and the structure was repeated in order to form a lineup 50 μm long. The lineup was repeated 10 times with a fix step of 5 μm (G_y) in the vertical direction, in order to obtain 50x50 μm^2 active nanoantenna array. See Figure 2.3 a-b in section 2.1 for a graphical representation of the nanoantenna array.

Proteins solutions preparation

Lyophilized Bovine Serum Albumin (BSA – Sigma Aldrich) was chosen as model protein. 10 μM BSA stock solution in water was prepared. Two aliquots were used: (i) one aliquot was diluted in ultrapure water at 0,1 μM final concentration and used for protein physisorption and drop casting assays; (ii) the second aliquot was equilibrated with sodium acetate solution 10 μM, pH 4.5, with a protein final concentration of 0,1 μM, and used for the amine coupling procedure.

Protocols for protein thin layer formation on gold nanoantennas

Three protocols for the covering of the nanoantennas with BSA thin film were explored:

(i) Thin film formation by drop casting taking advantage from the *coffee ring* effect: 0.4 μL of BSA solution 0.1 μM were dropped directly in correspondence of nanoantenna chips, letting the drop to dry for about 1 hour. Dried drops were rinsed with ultrapure water in order to remove the excess of non-absorbed protein. Samples were further dried under gentle nitrogen stream. Protein deposition caused the formation of a protein *coffee ring* (~250 nm) outside the nanoantenna array and of a thiner layer (6±2 nm) of proteins on the nanoantenna active device area. Thickness of both the external *coffee ring* and internal thin layer were evaluated by AFM measures, shaving the residual protein layer onto the CaF_2.

(ii) Protein physisorption: Before protein anchoring, the gold surface was activated by oxygen plasma 100% (100W for 2 minutes): plasma treatment eliminates biological contaminants, promoting the formation of volatile carbon compounds, and makes the gold surface more hydrophilic, introducing chemical functional groups (carbonyl, carboxyl, hydroxyl) at the gold surface. CEIRA devices were then immersed overnight (ON) in a 0,1 μM solution of BSA at 4°C in order to promote the physical adsorption of the protein onto the gold surface of the nanoantennas. After protein adsorption, the arrays were smoothly rinsed with ultrapure water and dried under gentle nitrogen stream.

(iii) Amine coupling immobilization procedure: Gold nanoantennas of the CEIRA devices were functionalized taking advantage from Au-thiol chemistry [17]. After nanofabrication, the devices were cleaned by O_2 plasma 100% (100W for 2 minutes), in order to activate the gold surface of the nanoarrays, as in the former case. Substrates were then incubated ON in a 300 μM ethanol solution of 6-MercaptoHexanoic Acid (MHA) at 4°C in order to generate a Self-Assembled Monolayer (SAM) onto the gold surface of nanoantennas. CEIRA substrates were rinsed with 95% ethanol and dried under a nitrogen stream. The SAM of MHA was functionalized by 1-ethyl-3-(3-dimethylaminopropyl) carbodiimide hydrochloride (EDC) and N-hydroxysuccinimide (NHS) reaction: samples were incubated in a solution of EDC-NHS 200μM:50μM 1:1 v/v for 7 minutes at room temperature (RT). EDC-NHS solution was then removed and substrates were covered with the activated protein solutions 0,1 μM in sodium acetate, pH 4.5 for 30 min. The excess of unreacted proteins was then removed by rising the substrate with ultrapure water and samples were dried under gentle nitrogen stream.

FTIR microscopy measures

FTIR measurements were done at the Chemical and Life Sciences branch of SISSI beamline at Elettra Sincrotrone Trieste – ITALY [18].

Signal response of the individual arrays was evaluated using FTIR microscopy. FTIR spectra were collected by using conventional globar source. Spectra were acquired by using a Bruker Hyperion 3000 Vis-IR microscope equipped with a mid-band HgCdTe detector having a 100 μm sensitive element, coupled with a Bruker Vertex 70v interferometer. Spectra were collected in transmission mode using 15x Schwarzschild condenser and objective and setting knife-edge apertures to 50x50 μm^2 in order to match the dimension of each nanoantenna array. In order to control light polarization, a Mid-IR polarizer was placed in the light path, before the sample. Each spectrum was collected from 4000 to 800 cm^{-1} in double side, forward/backward acquisition mode with a scanner velocity of 40 kHz. 1024 scans were averaged with a spectral resolution of 4 cm^{-1}. Fourier transform was carried out with Mertz phase correction, Blackman-Harris 3-terms apodization function. Spectra were rationed against a background acquired onto a clean CaF_2 window 1 mm thick.

CEIRA data analysis

Raw spectra were corrected for carbon dioxide and water vapor using OPUS 7.5 routines (BrukerOptics GmbH). Enhanced protein signals have been extracted compensating raw

resonance spectra by adopting the polynomial spline correction method described in section 2.5.3.

Results and Discussion

The efficiency of the three protocols exploited for thin BSA layer formation, (i) thin film formation by drop casting taking advantage from the *coffee ring* effect, (ii) protein physisorption and (iii) amine coupling immobilization procedure, has been primarily verified by the analysis of the correspondence between the BSA CEIRA spectra and the BSA reference spectrum.

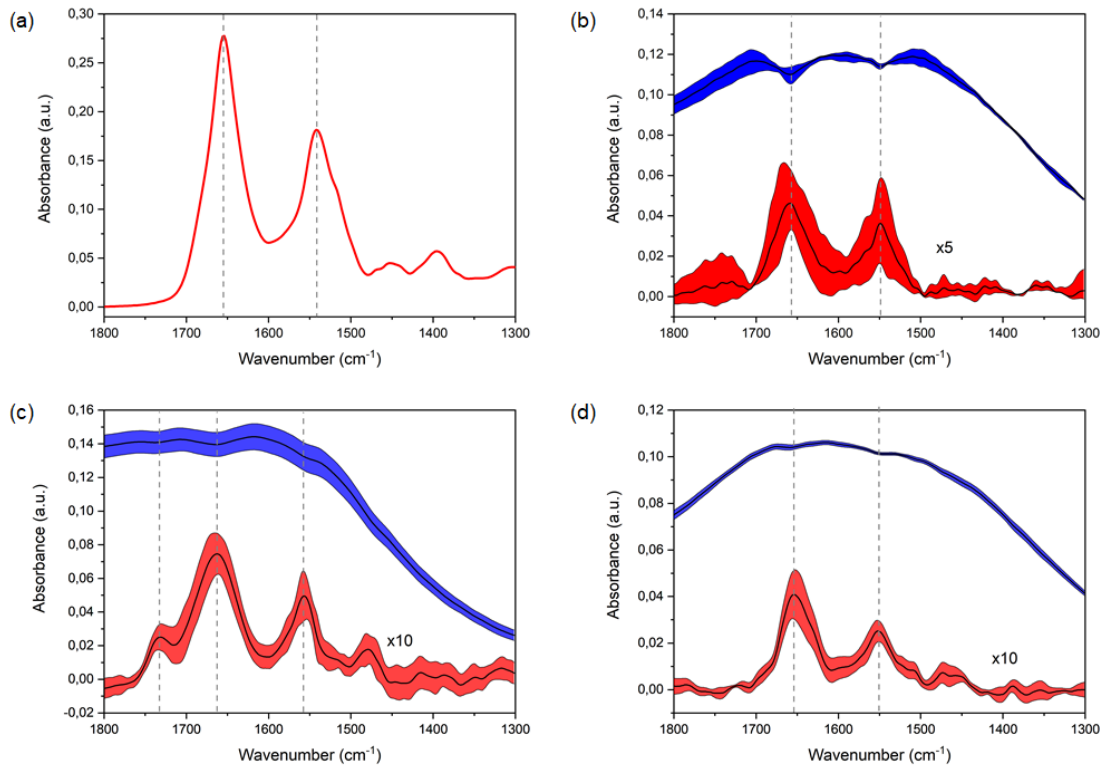

Figure 3.1: Main results of CEIRA acquisitions in the L1-ROI on BSA anchored with different strategies and comparison with BSA ATR-FTIR reference spectrum. (a) BSA reference spectrum acquired by ATR-FTIR spectroscopy; (b) CEIRA acquisitions of BSA deposited by drop casting (c) CEIRA acquisitions of physisorbed BSA; (d) CEIRA acquisition of BSA protein monolayers obtained by amine coupling procedure. Blue curves represent raw data while red curves the extracted absorption spectra. Shadowed areas represent the standard deviation of different acquisitions, while the black lines represent the average spectra.

In Figure 3.1a, the FTIR reference spectrum of BSA protein is presented in the Amide I-II spectral region. The reference spectrum has been collected with ATR-FTIR spectroscopy using the PIKE Miracle single-reflection accessory, mounting a Ge IRE. The final spectrum has been corrected for intensity of the absorption bands using the "Extended ATR

correction" algorithm of OPUS 7.5 software, setting the refractive index of BSA constant at 1.45 in the acquisition range [19]. As already mentioned, BSA is a globular protein, mainly α helical folded. The shape of Amide I and Amide II bands reflects this condition: Amide I band is centered at ~1656 cm^{-1} while Amide II is peaked at ~1542 cm^{-1}, as highlighted by the dashed vertical lines in Figure 3.1a. More details on BSA structure will be provided in section 3.3. In Figure 3.1, panels b-d show in the upper part the average raw CEIRA spectra for each investigated anchoring technique in the ROI L1 (black curve); the blue shadowed area represents the standard deviation among the repeated measurements done on nanoarrays of different chips of the same plasmonic device. For each panel, the lower curve is the average of the extracted absorption spectra of the protein obtained by using the procedure described in section "2.5.3 *Protein signal extraction*"; the red shadowed area shows the standard deviation of the different acquisitions. The vibrational fingerprint of protein layers can be detected for each investigated immobilization approach: it can be seen as negative peaks on the resonance curve of L1 arrays of CEIRA devices, and better appreciated upon extraction of the absorption spectra. More details on the three protocols are provided in the following sections.

(i) Drop casting and *coffee ring* effect

A simple, low-cost, versatile, and potentially scalable casting method for thin layer formation is to drop a diluted solution onto a substrate in a process called drop casting [20]. It usually results in the spreading of a liquid solution and the formation of a solid thin film after solvent evaporation. In the case of a protein solution, as the drop dries, most of the material is dragged at the edge of it and the protein precipitates forming the so-called *coffee ring* [14]. Namely, when the drop starts shrinking due to the evaporation, it generates a well-defined film of degrading thickness towards the center of the drop, until all water is evaporated.

For our experiment, BSA thin films were obtained by letting drops of a diluted protein solution to dry on top of the arrays of the CEIRA devices. It is a matter of fact that, due to the phenomenon exploited for layer formation, the resulting film is non-uniform; in our experiments the volume of the drop was tuned in order to minimize the BSA layer thickness on top of the nanoantennas. In order to measure the thickness of this resulting layer in correspondence of nanoantenna arrays, nano-grafting AFM measures were performed on functionalized samples. The external protein coffee ring resulted to be ~250 nm high, while the internal thin layer wasn't uniform and its height was ~6±2 nm. FTIR BSA transmission

spectra on non-patterned regions of the device were also collected, both in correspondences of the thick external ring and for the thinner layers outside nanoantenna arrays. The results are reported in Figure 3.2. From the figure, amide bands can be clearly distinguished for the spectrum collected on the thicker ring (average absorbance $2x10^{-2}$ a.u), while they are 10 times less intense for the thinner ones. The spectra acquired on bare CaF_2 inside the coffee ring are very noisy, making impossible to discern any conformational protein detail. Therefore, due to this specific deposition protocol, protein signals outside antenna arrays in the coffee ring were barely detectable with conventional FTIR spectroscopy, but not negligible (see inset in Figure 3.2). In order to properly compensate the far-field contribution to BSA spectrum, CEIRA spectra shown in Figure 3.1b were normalized with a background taken onto the CaF_2 outside antenna arrays but within the coffee ring. CEIRA signals and FTIR conventional acquisitions taken outside the arrays, namely on BSA thin film having the same thickness (~6±2 nm on average), are compared in Figure 3.3. From the figure, it can be appreciated the strong signal enhancement effect and the increasing of signal to noise ratio (S/N) provided by CEIRA measurements.

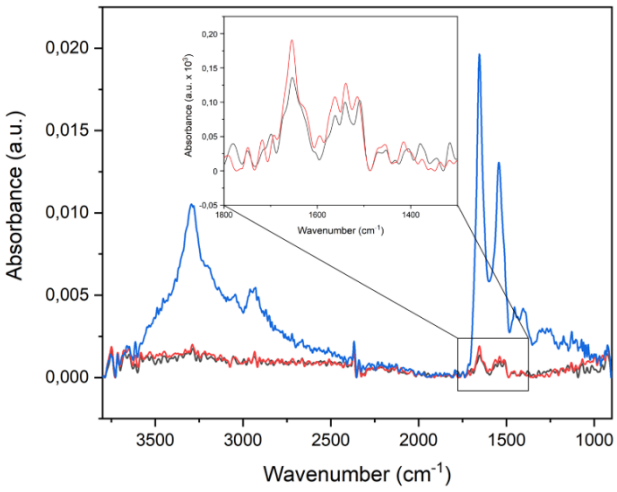

Figure 3.2: FTIR spectrum of BSA drop casted onto CEIRA devices, outside nanoantenna arrays. Blue spectrum was acquired onto the external coffee ring of BSA. Red and gray spectra were acquired on the thinner protein film on bare CaF_2 substrate within the coffee ring. In the inset, thin-film protein signals are highlighted.

Enhanced signals were at least one order of magnitude more intense than those in the coffee ring outside the plasmonic areas of the device. From the BSA CEIRA spectrum, it is possible to recognize the characteristic spectral features of BSA at ~1656 cm^{-1} (Amide I) and ~1542 cm^{-1} (Amide II), consistent with those of BSA reference spectrum reported in Figure 3.1a. However, the inhomogeneity of the generated film, demonstrated by the large standard

deviation of the measurements repeated over different arrays (see shadowed red area in Figure 3.1b) and confirmed by AFM data, underlines the necessity to better control the protein layer homogeneity and thickness. To this aim, the protein physisorption method was tested.

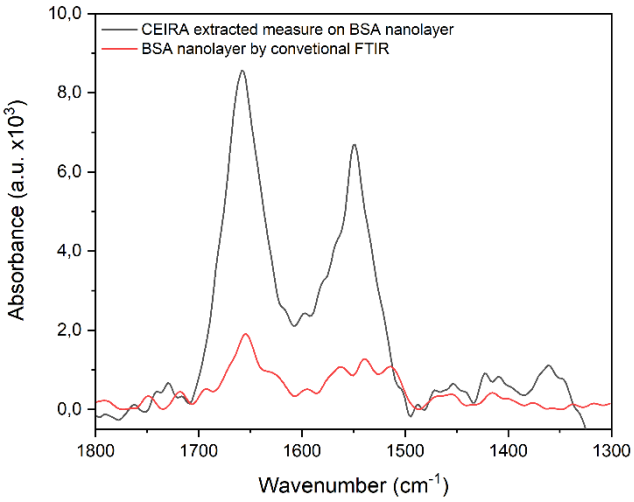

Figure 3.3: Comparison between CEIRA spectrum and conventional FTIR transmission spectrum on a BSA film of comparable thickness (~6±2 nm). Average spectrum of six CEIRA measurements on different chips is reported.

(ii) Protein Physisorption

Physisorption, which is also called physical adsorption, is a non-covalent functionalization procedure, where the binding of samples to the surface is mostly driven by weak adhesive van der Waals forces [15].

In order to promote this process, CEIRA sensor was incubated ON in a 0,1 µM BSA solution. This led to the formation of an uncontrolled nanolayer of BSA: by method construction, proteins are attached both onto the surface of gold nanoantennas and CaF_2 substrate, with thicknesses that can differ region by region. However, conventional FTIR measurements on regions outside the plasmonic active areas did not show any detectable protein signals (data not shown). In Figure 3.1c are reported both CEIRA resonant response (blue curve) and extracted BSA signals (red curve) in the protein's amide L1 ROI. Average Amide I peak maximum is about 8.6×10^{-3} a.u. for physisorbed BSA protein vs 1.2×10^{-2} a.u. for drop casted thinner BSA layer. Therefore, we can speculate that the protein layer obtained by physisorption is definitely thinner than ~6 nm. It is noteworthy to mention than for a such thin layer is impossible to establish its exact thickness by AFM since, even if a non-homogenous monolayer of protein was present, its thickness is in the order of the average

roughness of the CaF_2 substrate (about ±2nm). Nevertheless, from the standard deviation of the spectra reported in Figure 3.1c, it is also possible to assert that this method does not guarantee the formation of a uniform layer.

By matching the spectra of Figure 3.1c with the BSA reference spectra, some variations in the spectral shape of physisorbed BSA can be noticed. For CEIRA BSA acquisitions, while the two main peaks of Amide I and Amide II bands area centered in the same position of the reference (at ~1655 and ~15454 cm^{-1}), an extra peak appears centered at about 1735 cm^{-1}. This contribution can be attributed to C=O stretching vibrations [21]. Actually, such feature does not belong to native protein spectrum. It is known from the literature that the binding of a protein to hydrophobic surfaces may induce protein unfolding and aggregation, resulting in the exposure of the hydrophobic groups [22]. Nevertheless, in our spectra the positions of the Amides bands do not support this hypothesis. The shape of Amide I band excludes also the presence of protein aggregates, since typically, when it occurs, spectra present a significant contribution at ~1630–1615 cm^{-1} [23], which instead cannot be seen in our CEIRA acquisitions. As a consequence, the only reasonable hypothesis for the contribution at 1735 cm^{-1} is the residual presence of PMMA resist. If it derives from an under-development of the antennas or to an incomplete lift-off it is hard to establish, but for sure the plasma cleaning was not sufficient to remove it completely. However, due to the isolated case, and considered the unsatisfactory homogeneity of the layer thickness, we did not investigate further. Indeed, since both deposition methods did not provide a uniform protein layer, we decided to direct our attention to anchoring strategies that relies on the chemical link of the protein to the gold antennas.

(iii) Amine coupling: protein monolayer formation

In order to finely control protein immobilization, different protocols based on chemisorption have been developed [10], [16]. These functionalization methods usually require the formation of a self-assembled monolayer (SAM) onto the surface of interest, in order to modify its chemical properties. SAMs are formed by immersing a substrate into a solution of the active molecules that will constitute the SAM. The driving forces for the spontaneous formation of molecular assembly include the chemical bonding of such molecules onto the surface and the intermolecular interactions between them. The selective interaction of thiols with gold surfaces, which leads to the formation of thiols SAMs, is the most popular way for generating molecular films on gold, due the resulting oxide-free, clean, flat surfaces. Upon formation, thiol SAMs can be further modified both in the gas phase and in liquid media

under ambient conditions [24]. One of the benefits of working with gold nanoantennas is the possibility to take advantage from this so-called gold-thiol chemistry [17]. Furthermore, due to this specificity, we can guaranty the linking of proteins only onto the gold surface of SAM modified nanoantennas. In order to bind target proteins onto the surface of our CEIRA device, we chose thiols with a carboxylic head group, which were induced to react with the amine group, more likely at the *N*-terminal of proteins, via amine coupling procedure [25]. In Figure 3.4 a scheme of the SAM formation is presented.

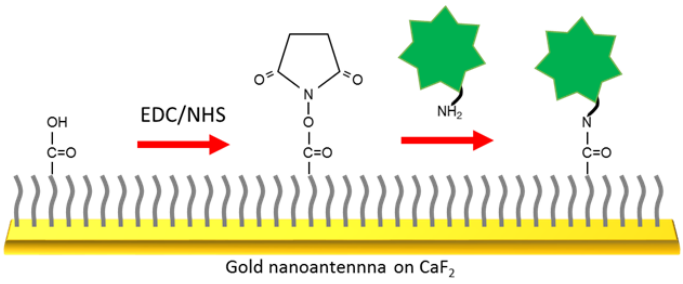

Figure 3.4: Scheme of amine coupling steps. After the formation of a MHA SAM onto the gold surface of nanoantennas, the COOH head group of thiols was activated by EDC/NHS (See methods). Finally, BSA protein was added to solution and induced to react with the EDC/NHS activated SAM. The reaction promoted the linking of the target protein via its NH_2 exposed moieties.

CEIRA substrates were firstly incubated with a solution of 6-Mercaptohexanoic acid (MHA) in order to promote the SAM formation via Au-S bonding. After incubation, samples were rinsed with ethanol and directly measured in order to verify SAM formation.

Specifically, the formation of the thiol monolayer was verified by sensing the methyl and methylene moieties vibrational signals at 2920 and 2854 cm^{-1} with arrays in the L0-ROI (resonant peak centered at ~2900 cm^{-1}, see Figure 3.5a). After the SAM formation, the protein was anchored to the thiol terminal COOH groups via amine coupling procedure, upon sodium acetate activation at pH 4 (protein solution 0.1 μM). CEIRA measurements have been performed with arrays in the L1 – ROI (resonant peak centered at ~1600 cm^{-1}, see Figure 3.5b). As can be seen in the Figure 3.5b, standard deviation of measures is very low, suggesting that proteins bounding onto the gold surface of our structures are organized in a uniform layer. Maximum intensity of protein signals was even less than that of physisorbed one ($5.3x10^{-3}$ vs $8.6x10^{-3}$ a.u.), confirming that this protocol guaranties the linking of the thinner protein layer onto our surface. Since the SAM does not form on bare CaF_2, we can assume that a monolayer of protein is present only in correspondence of gold structures. Moreover, the formation of such a uniform film of protein seems to well preserve BSA conformation. BSA-anchored spectra are dominated by the main protein bands Amide I and

II, centered at 1656 and 1542 cm^{-1}, which are highlighted by the dashed gray lines in Figure 3.4b.

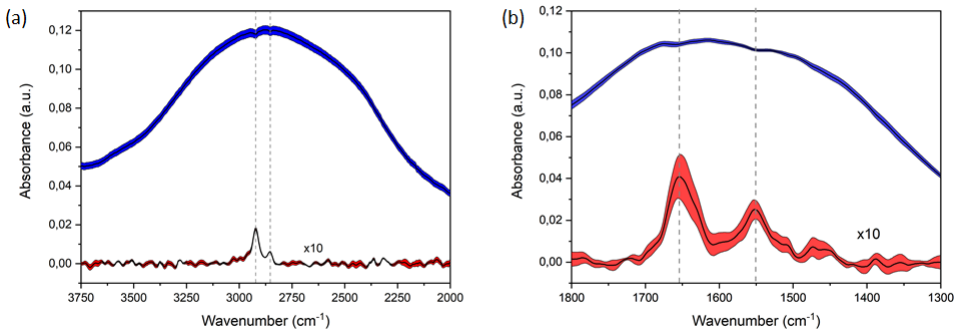

Figure 3.5: CEIRA acquisitions of BSA chemisorbed via amine coupling. Panel (a): CEIRA results in the ROI-L0 before BSA amine coupling. Panel (b): CEIRA results in the ROI-L1. Enhanced signals are highlighted by the vertical dashed lines. Shadowed areas represent the standard deviation while black lines represent the average spectrum for spectra collected on 9 arrays of three different CEIRA devices.

As a side experiment, we applied the very same protocol for protein binding onto a nanostructured surface, obtained evaporating a thin gold layer onto 100 nm deep silicon ridges , which did not provide plasmon resonances. These samples were measured both with conventional FTIR reflectance measurements and by grazing angle FTIR measurements. In both cases, we could not detect the presence of protein monolayers attached onto the surface (data not shown). Infrared scattering-type Scanning Near Field Optical Microscope (IR s-SNOM) measures were also done on the grating, with the NeaSNOM microscope at Neaspec GmbH, Munich, using a DFG laser as a source. These IR measurements with nanometric lateral resolution (tip radius about 20 nm) confirmed the formation of BSA monolayers on these flat gold systems, as can be appreciated by the Amide peaks visible in the spectra plotted in in Figure 3.6.

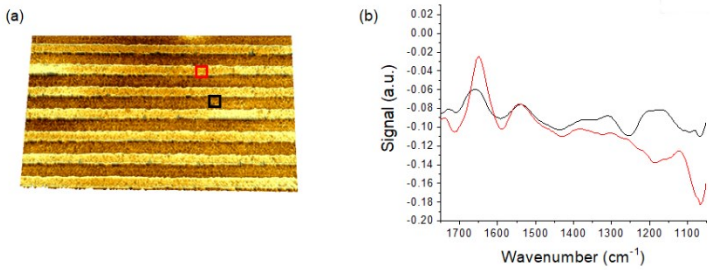

Figure 3.6: IR s-SNOM measures on gold gratings functionalized with BSA via gold-thiol chemistry and amine coupling procedure. Panel (a): AFM topology of BSA-functionalized substrate, showing a gold grating made by ridges 100 nm high and wide. Black and red squares represent the chosen points for the measurements with IR s-SNOM. Panel (b): two representative IR s-SNOM measures in the selected spots.

Conclusions

With this series of experiments, we provided the evidence that CEIRA devices for Mid-IR enhancement of the protein signals offer the chance to probe protein monolayers. Specifically, we verified that thin film formation by both drop-casting exploiting the *coffee ring* effect and physisorption do not guarantee the formation of a controlled film, generally thicker that a single monolayer and rather inhomogeneous in height. Conversely, thanks to the amine coupling procedure, it was possible to guarantee the formation of a uniform monolayer of proteins. It has to mention at this stage that only for very controlled monolayers it is possible to estimate the magnitude of signal enhancement provided by CEIRA devices, while many paper in literature provide this estimation relying of non-fully controlled deposition approaches, as already highlighted in section 2.0. For our devices functionalized with BSA via amine coupling, if we consider a theoretical homogenous coverage of the gold surface of nanoantennas with BSA protein, it is possible to roughly estimate the number of molecules probed with a single CEIRA measurement. In particular, the dimension of the spot for each acquisition with globar source was 50x50 μm^2 and it includes all the nanoantenna sensitive area. Considering nanoantenna dimension (1,8 μm long for L1-ROI and 75 nm wide) and array geometry (10 rows spread in the vertical direction by 5 μm), each L1 array contains 276 nanoantennas, covering a surface area of approximately 37 μm^2. Being the dimensions of BSA protein about 140 Å × 40 Å × 40 Å [26], assuming the complete coverage of the gold surface of the device, we probed at most $6{,}7*10^5$ proteins at each measurement. Moreover, since that the plasmonic effect is enhanced at the nanoantenna apices (hot-spots), the effective number of measured molecules is even lower. The rough estimation here presented suggests that the sensing capability of our CEIRA devices reaches attomolar sensitivity (10^{-18} mol) in dried conditions. These data are in accordance with those reported in literature for similar plasmonic substrates [4].

3.2 Pushing forward the FTIR limit of detection

Despite their significant sensitivity, the signal to noise ratio (S/N) of BSA CEIRA spectra previously shown was not enough for making finer conformational studies via, for example, second derivative analysis. For improving the quality of spectra, two different strategies can be adopted: (i) the increase of the number of hot spots in the ordered array; (ii) the increase of the signal exploiting a brighter IR source, such as IR-SR. During my PhD activity, both

strategies were pursued. In the section 3.2.1 we demonstrate how to collect spectra with good S/N by coupling the vibrational fingerprint of molecules with stronger resonance responses obtained by *cross*-designed plasmonic architectures. Additionally, IR-SR has been used instead of conventional globar source in order to further increase the detection limit, as described in section 3.2.2. Section 3.3 will focus on the conformational analysis of two different model proteins, exploiting the strategies described in the present section.

3.2.1 Maximizing the CEIRA response: a new substrate geometry

As seen in section 2.4.1, the *cross*-geometry guarantees a signal enhancement of at least one order of magnitude higher than that of linear nanoantennas, with the capability to provide the same resonance response for both direction of polarization. The experiments presented in the following aimed to verify if the generation of such intense resonances by *cross*-geometry substrates resulted also in an increased quality of the enhanced protein signals. To this purpose, we functionalized both *line-to-line* and *cross*-geometry substrates with BSA by using the amine-coupling procedure previously introduced. CEIRA protein spectra obtained with the two different structures were matched in order to highlight the spectral advantages offered by this second generation layout.

Materials and methods

Gold nanoantenna arrays fabrication

CEIRA devices with *line-to-line* and *cross*-geometries were fabricated as reported in Section 2.2. Substrates contained regions made by specific gold nanoantennas that allow Mid-IR signal enhancement in the amide ROI, centered at 1600 cm^{-1} (ROI – L1, λ=6,2 μm). For *line-to-line* devices, final structures were linear nanoantennas 80 nm high and 75 nm wide, 1,8 μm long, spaced by 50 nm in the x direction (G_x), in order to form a 50 μm lineup, and by 5 μm in the y direction (G_y). *Cross* devices were made by adjacent 50x50 μm^2 arrays with nanoantennas disposed to form centrosymmetric crosses. Gold nanoantennas of these arrays were 80 nm high and wide, with a length of 2,04 μm. The gap between the antennas corners (G_x) was of 25 nm. See Figures 2.3 c-d for more details on the geometries.

Plasmonic devices have been functionalized with BSA by following the protocol detailed in 3.1, "*Amine coupling immobilization procedure*". They have been measured by FTIR microscopy and data analyzed following the same experimental parameters reported in 3.1 "*FTIR microscopy measures*" and "*CEIRA data analysis*".

Results and discussion

The comparison between the resonance responses of the *line-to-line* and *cross*-geometry devices has been already discussed in Chapter 2. In Figure 3.7, both resonance response and extracted protein spectra of BSA-functionalized devices are reported. Definitely, *cross*-devices provide BSA resonance spectra more intense than those of *line-to-line*, as can be appreciated by comparing Figure 3.7a and b. The reasons of this behavior have been already highlighted in 2.4.1.

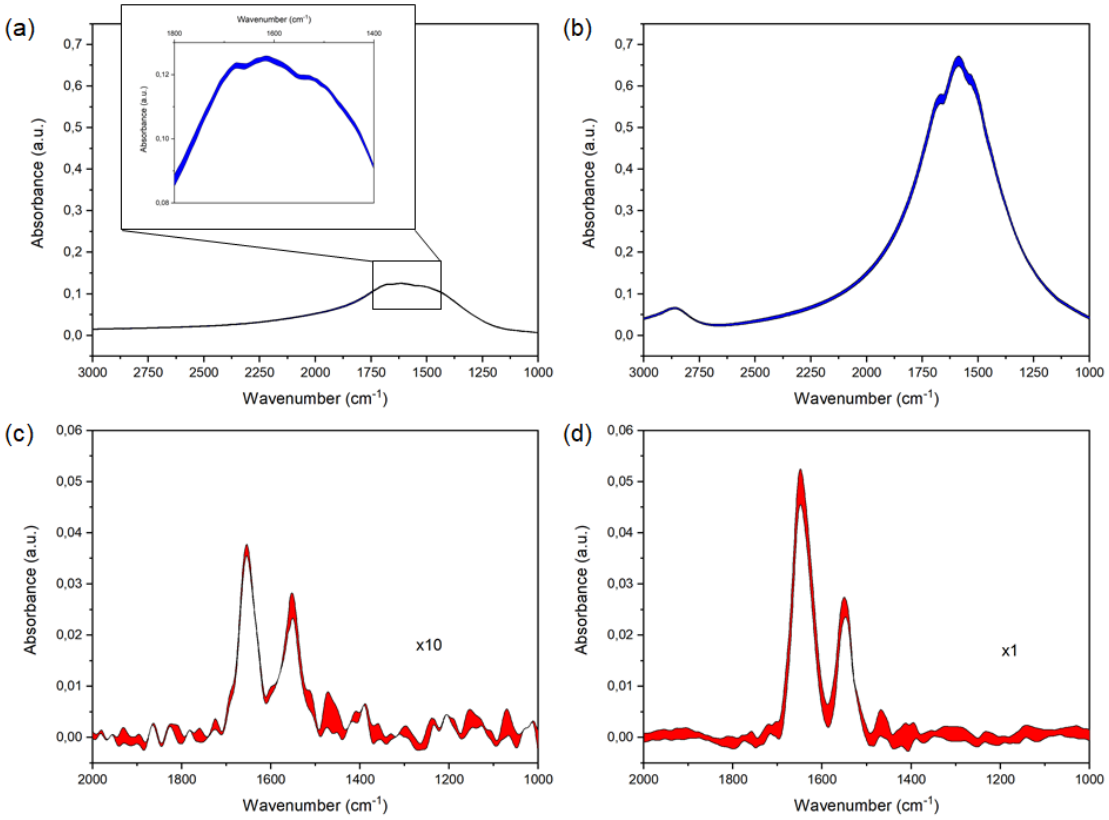

Figure 3.7: Comparison between CEIRA response of *line-to-line* and *cross*-devices functionalized with BSA protein by amine coupling. (a) Resonance response of *line-to-line* functionalized device. In the inset, protein enhanced features are highlighted. (b) Resonance response of *cross*- device functionalized with BSA. (c) Extracted protein spectrum from *line-to-line* devices. Signals have been amplified 10 times. (d) Extracted protein spectra from *cross*-devices. Shadowed areas represent the standard deviation of spectra collected on 9 arrays of three individual *line-to-line* (a,c) and *cross*-geometry (b,d) CEIRA devices.

By comparing the two extracted protein spectra (Figure 3.7 c-d), it is possible to appreciate the difference in BSA signal intensity: *cross*-devices extracted signals (intensity of Amide I band: 5.4×10^{-2} a.u.) are at least 10 times more intense than those obtained from *line-to-line* devices (intensity of Amide I band: 3.9×10^{-3} a.u.). In addition, signals of protein extracted

from *cross*-devices emerge better from the noise and the shape of the bands is better preserved.

As matter of fact, the increased intensity of protein signals obtained by using *cross*-devices comes as a result of the higher intensity of the resonance curve provided by of these structures. Moreover, we have to consider that *cross*-devices are made by a higher number of gold plasmonic elements per arrays. As a consequence, the gold surface onto protein can attach is larger than that of *line-to-line* devices. Thus, the increased signal intensity is also related to an increased number of probed molecules.

Conclusions

With the presented experiments we demonstrated that CEIRA *cross*-devices improve the detection of protein monolayers as a consequence of a higher CEIRA resonance response. The capability of this technique to provide spectra with high quality opens the possibility to obtain reliable and detailed conformational information on the sample under investigation.

3.2.2 CEIRA on proteins using Synchrotron Radiation

Our goal is to apply the CEIRA technique to spectroscopically detect and characterize proteins at concentrations relevant for biological applications, using an amount of sample as small as possible. Thanks to the *cross*-geometry, we increased the spectral quality of extracted protein signals. In order to further increase the efficiency of the technique, it was proposed to couple plasmonic devices with brighter IR light sources, such as synchrotron [27] or quantum cascade lasers [28]. In particular, the IR synchrotron radiation (IR-SR) is about two to three order of magnitude more brilliant than conventional globar sources [18], enabling infrared chemical imaging at the diffraction limit. In literature, first trials in this direction were made by Neubrech *et al.* on single nanoantennas, producing signal enhancement up to few attograms of target model molecules [1].

In order to prove the behavior and the advantages offered by IR-SR for our systems, the very same BSA functionalized *cross*-geometry devices presented in the previous section were characterized with the IR-SR and the response compared with the one collected in the previous experiments with the conventional globar source.

Materials and methods

Plasmonic devices with *cross*-geometry have been fabricated as reported in 3.2.1. They have been functionalized with BSA by following the protocol detailed in 3.1 "*Amine coupling immobilization procedure*".

FTIR measures and data analyses

C*ross*-geometry devices were measured by using IR-SR at the Chemical and Life Sciences branch of SISSI beamline at Elettra, with the same instrumentation reported in 3.1 "*FTIR microscopy measures*". Spectra were collected in transmission mode using 15x Schwarzschild condenser and objective and setting knife-edge apertures to 10x10 μm^2, collecting maps with a raster scan of 7x7 points. For each spectrum, collected from 4000 to 800 cm^{-1} in double side, forward/backward acquisition mode with a scanner velocity of 120 kHz (SR), 1024 scans were averaged with a spectral resolution of 4 cm^{-1}. Fourier transform was carried out with Mertz phase correction, Blackman-Harris 3-terms apodization function. Spectra were rationed against a background taken on a 1 mm thick CaF_2 window. Raw CEIRA spectra have treated corrected as reported in 3.1 "*CEIRA data analysis*".

Results and discussion

While conventional globar sources can be considered not polarized, IR-SR is characterized by a well-defined polarization [29]. As proven in section 2.4.1 *"IR response"*, *cross*-geometry devices are substantially independent from the direction of polarization of light. However, being IR-SR polarized, in order to obtain a homogenous response of the two arms of the cross, we first selected the polarization state that guarantees the maximum signal, and then we adjusted the orientation of the nanoantenna array at 45° with respect to it. To this aim, the device was mounted onto a goniometer. A 2D map of the substrate representing the distribution of the IR signal enhancement was obtained by raster scanning the sample with a spot light of 10 μm^2 and is reported in Figure 3.8 (left panel); a 3D plot of the same distribution is shown in Figure 3.8, right panel. As can be seen, the best resonance performance of such substrates is located at the center of the arrays (dark red area in Figure 3.8). The capability to focus the light beam in this region of the array allows selectively measuring at the highest signal enhancement.

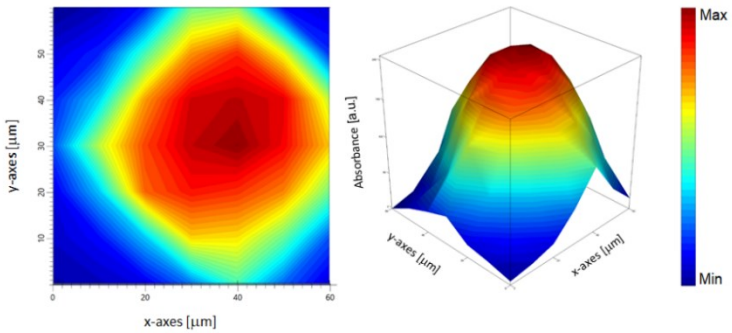

Figure 3.8: 2D and 3D maps of the intensity of the signal provided by the CEIRA *cross*-geometry devices. Maps are the result of a 7x7 raster scan of the sample by using IR-SR source with a spot of 10x10 µm^2. Distribution of the intensity was obtained integrating the resonant peak between 1800-1400 cm^{-1} (L1-ROI).

In Figure 3.9 the resonance responses and the extracted protein signals obtained by measuring with both IR-SR and globar sources onto the very same nanoantenna array functionalized with BSA are reported. The resonance response obtained by using IR-SR is higher than using conventional globar source. As in the former case, higher is the resonance, more intense are extracted protein signals, and protein spectra acquired with IR-SR at 10x10 µm^2 resulted more intense than those obtained with globar source at 50x50 µm^2.

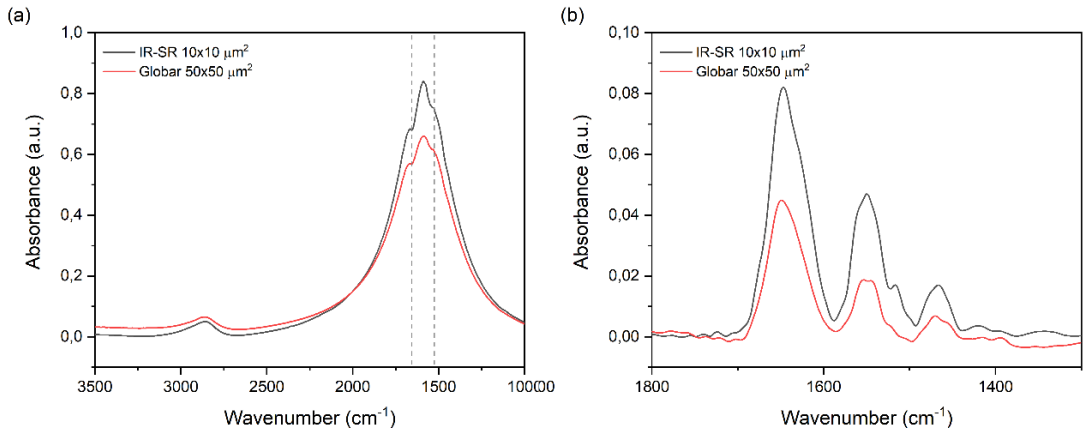

Figure 3.9: Comparison between BSA CEIRA spectra acquired with conventional sources (red spectrum) and IR-SR (black spectrum) for *cross*-geometry devices. In panel (a), the resonance response is compared. In panel (b) a comparison of protein extracted signals is reported.

Despite a gain of only about a factor of 2 of IR-SR with respect to conventional source can be appreciated from Figure 3.9, it should be noticed that the measured sample areas differ for a factor of 25 (2500 µm^2 versus 100 µm^2), so that within the area illuminated with synchrotron beam (10x10 µm^2) there are almost 25 times less antennas than those sampled with the globar source (50x50 µm^2). Therefore, the global IR-SR advantage approaches the value of 10^2, conventionally achieved in operative conditions for comparable lateral

resolutions. Nevertheless, by considering the average dimension of the BSA molecule and the number of nanoantennas in the 10x10 μm^2 sampling region, we can estimate the number of molecules probed by IR-SR, that result to be limited to few thousands.

Conclusions

In conclusion, the combination of high performant *cross*-geometry plasmonic structures with IR brilliant sources provides a significant and consistent improvement in both plasmonic response and enhanced signals. The possibility to probe smaller areas would allow reducing the single nanoantenna array size and consequently to fabricate denser chips that could be functionalized with the same volume of analytes, limiting the sample waste. Therefore, we can envisage that the exploitation of more brilliant sources, such as IR-SR but also broad-band Mid-IR lasers, could push this technology for high-throughput screening approaches.

3.3 Conformational studies on model proteins in dried conditions

The development of sensitive tools which enable a strong field enhancement in the Mid-IR regime offers not only the possibility to detect the presence of ultra-low quantities of biomolecules, but also the unique chance to obtain their conformational details [30].

In the following section, we demonstrate the capabilities of our plasmonic devices to allow conformational studies on protein monolayers. As model systems, we chose two well defined proteins with distinct secondary structure conformation: Bovine Serum Albumin (BSA), which is a globular protein mainly alpha helix folded [31], and Concanavalin A (ConA) which is a lectin obtained from jack bean, the secondary structure of which is characterized mainly by β sheet strands [32]. In order to verify the reliability of CEIRA protein vibrational bands, reference spectra of the same proteins were acquired with FTIR-ATR spectroscopy. ATR and CEIRA results were matched both in absorbance and second derivative, in order to verify the capability of CEIRA microscopy to recognize and distinguish proteins on the base of their conformation.

Material and methods

Cross-geometry devices were used for the present experiments. They have been fabricated accordingly with the general procedure reported in Section 2.2 and with the geometric parameters reported in 3.2.1, "*Gold nanoantenna arrays fabrication*".

Model proteins preparation and device functionalization.

Lyophilized reference proteins Bovine Serum Albumin (BSA – Sigma Aldrich) and Concanavalin A (ConA – Sigma Aldrich) were used. Mother solutions of both proteins 10 μM were prepared at pH 7 in ultrapure water. For ATR-FTIR measurements, mother solutions were used, while diluted protein solutions (100 nM) in Sodium-Acetate 100 mM at pH 4.5 were prepared for amine coupling procedure. Plasmonic devices were functionalized with BSA and ConA via amine coupling, following the protocol detailed in 3.1, "Amine coupling immobilization procedure"

FTIR measures

Concentrated protein water solutions (10 μM) were measured by ATR-FTIR spectroscopy using the MIRacle™ single Reflection ATR accessory (PIKE Technologies) with Germanium (Ge) Internal Reflective Element (IRE). Spectra were acquired using the Vertex 70 interferometer (Bruker Corporation) equipped with a deuterated triglycine sulfate (DTGS) detector. FTIR-ATR spectra were collected from 3900 to 650 cm^{-1} in double side, forward/backward acquisition mode with a scanner velocity of 5 kHz. For each spectrum, 128 scans were averaged with a spectral resolution of 4 cm^{-1}. Fourier transform was carried out with Mertz phase correction, Blackman-Harris 3-terms apodization function. Background was collected on clean Ge IRE. 10μL of BSA and ConA solutions were dropped onto the IRE crystal, and ATR-FTIR spectra continuously acquired up to complete drying. SR-CEIRA transmission measurements were acquired as reported in section 3.2.2 "*FTIR measurements and data analysis"*.

Data analysis

ATR-FTIR spectra were corrected for position and intensity of the absorption bands using the "Extended ATR correction" algorithm of OPUS 7.5 software, setting the refractive index of BSA at 1.45 [19]. Raw CEIRA spectra were treated as described in section 3.1.1. Upon spectra extraction, vector normalized second derivative were computed (Savitzky-Golay filter, 17 smoothing points) in the region of interest by OPUS7.5 routines.

Results and discussion

IR-SR CEIRA measures were performed in dried conditions onto different devices functionalized with BSA and ConA. Enhanced spectra on BSA and ConA monolayers in the amide I-II region are reported in Figure 3.10.

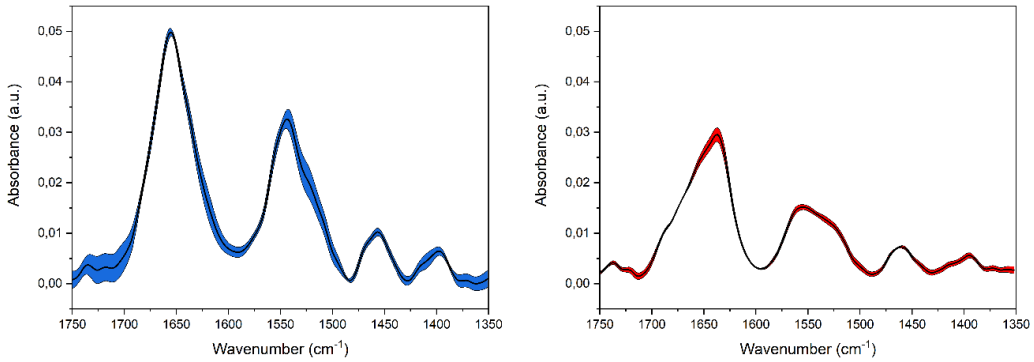

Figure 3.10: IR-SR CEIRA spectra of BSA and ConA protein monolayers, acquired using *cross*-geometry devices. Black line represents the average spectra of acquisitions on individual arrays of two different CEIRA devices, containing 9 arrays each. Shadowed areas represent the standard deviation of the measurements.

Amide bands of proteins are well detectable and repeatability of the measures is confirmed by the modest standard deviation appreciable in Figure 3.10. As aforementioned, BSA is mainly α-helix folded, and its Amide I band is peaked at 1656 cm^{-1}, position which is usually associated to the C=O stretching of the peptide backbone for α helix folded motifs [33]. Instead, ConA presents an Amide I band peaked at 1633 cm^{-1}, position usually associated to the native β-sheet components of proteins [34]. Due to the good quality of the data, it is quite straightforward to distinguish between the two proteins by the shape of their Amide I bands. The results of CEIRA measurements are shown in Figure 3.11 and compared with reference acquired on bulk material by single reflection ATR-FTIR. The good match between the CEIRA and ATR-FTIR spectra confirms the capabilities of our CEIRA devices to discriminate between differently folded proteins in dry conditions.

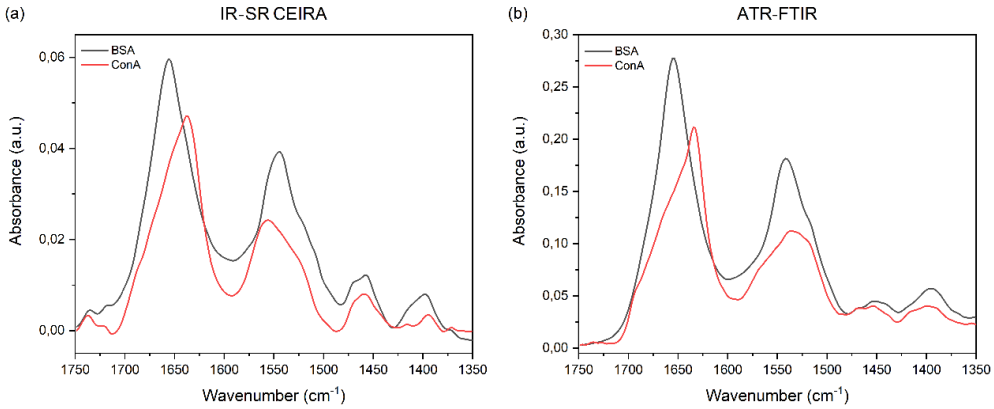

Figure 3.11: Comparison between IR-SR CEIRA and ATR-FTIR spectra of BSA and ConA. (a) IR-SR CEIRA enhanced spectra of proteins monolayer. Black curve BSA, Red curve ConA. (b) ATR-FTIR spectra of thick protein layers dried onto the Ge IRE. Black curve BSA, Red Curve ConA.

Furthermore, for better resolving the components corresponding to each conformational motif, second derivative analysis of CEIRA spectra was performed. Results are presented in Figure 3.12.

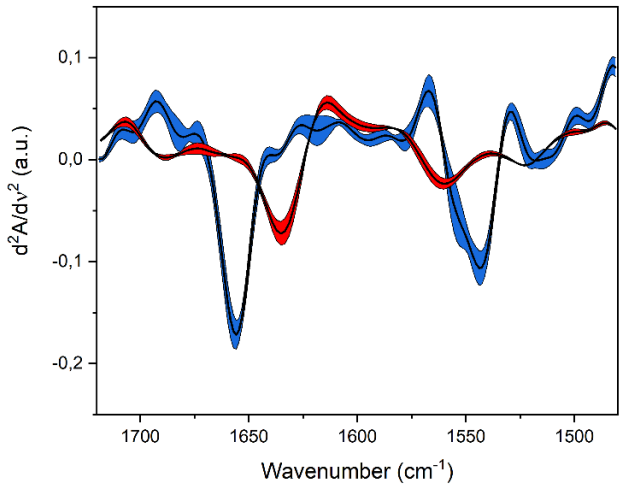

Figure 3.12: Second derivative IR-SR CEIRA spectra of BSA (blue spectra) and ConA (red spectra) protein monolayers.

Second derivative spectra in the Amide I range (1700-1600 cm^{-1}) are characterized by a single strong peak, centered at 1656 and 1633 cm^{-1} for BSA and ConA respectively. ConA second derivative spectrum shows an additional weak shoulder around 1690 cm^{-1}, which is associated with loops and β-turns [35]. BSA instead, presents a relevant component at 1686 cm^{-1} that can be associate to the abundance of loops and turns in the protein secondary structure [36]. The Amide II range (1595-1480 cm^{-1}) of the two proteins appears very different. BSA is characterized by a strong peak 1542 cm^{-1}, while ConA appears broader with two features peaked at 1556 and 1517 cm^{-1}. In particular, this second peak is also diagnostic for the abundancy of tyrosine (Tyr) side chains in proteins [37]. In the case of ConA, Tyr contributes for about the 4% of the primary sequence of the protein. Such contribution is visible also in the BSA second derivative spectrum, where the abundancy in the primary structure of Tyr resides is also of about 4%.

Conclusions

With the series of experiments reported in sections 3.1 to 3.3, we demonstrate that CEIRA is a repeatable and reliable method for performing protein conformational studies on dried systems. Results clearly indicated that CEIRA allows retrieving conformational details of protein monolayers with a spectral quality comparable to that of ATR-FTIR technique on bulk material. The geometry of the nanoarrays and the quality of the fabrication process are

key factors for achieving a significant signal enhancement that can be further improved by exploiting high brilliant IR sources such as IR-SR. Whether the formation of a uniform and ultrathin protein layer is guaranteed, as in the case of nanoantenna functionalization by amide-coupling, CEIRA sensitivity up to few thousands of molecules can be estimated.

3.4 Facing the water absorption issue

Although dried proteins are far from their physiological environments, this kind of measurements is generally accepted [36], [38]. However, in order to depict information about protein in their active form or follow their dynamics under different conditions, performing measurements in an aqueous environment, close to the physiological one, is mandatory [39]. Unfortunately, FTIR transmission measurements in aqueous solution are very challenging. Even though water is an essential component of most biological processes, it represents the main obstacle in FTIR spectroscopy [40], since it is a strong IR absorber and, specifically, in the case of protein measurements, its bending absorption peak (maximum at about 1640 cm^{-1}) overwhelms the Amide I and, partially, the Amide II band. A more exhaustive explanation of this limitation, as well as on the sensitivity limit of the technique, has been given in the section 1.2.5 "*The water absorption barrier*". Despite these constraints, during the PhD thesis, CEIRA devices have been tested also in classical transmission geometry, creating a static liquid cell, closed without using any spacer in order to restrict as much as possible the optical path. Despite the efforts, the water far-field signal was so intense to overwhelm the plasmonic response (data not shown). In order to further minimize the water contribution to transmission CEIRA spectra, we tested an approach based on the use of graphene as ultra-thin confining layer. This strategy was part also of a larger project called RENEWALS CERIC-ERIC. We recently demonstrated that environmental liquid cells made by using single layer graphene sheets can successfully act as conformable water-confining membranes for FTIR analysis of hydrated biological specimens [41]. However, at the actual stage of development of the technique, we did not obtain a good device sealing. Since, these tests were done at the end of my PhD and they require much more optimization no results will be shown in this thesis.

ATR sampling partially addresses the issue of the water layer confinement, since a fixed path length is determined by the evanescent field penetration [42]. Nonetheless, it cannot achieve an adequate sensitivity without any additional strategy for signal amplification. In 2013, Altung and co-workers developed a strategy for overcoming the water issue for

resonant SEIRA measurements, proposing a method which takes advantage from both the strong field enhancement of resonant antennas and an ATR-like geometry of acquisition [43]. They called this method "Plasmonic Internal Reflectance – PIR". In their scheme, the internal reflection was provided not by the classical total internal reflection, but directly by the patterned plasmonic surface (See Figure 3.14).

By taking advantage from this particular geometry of acquisition, we performed a series of measurements for demonstrating that our CEIRA devices can enable conformational studies on protein monolayers in water environment. In the following, the PIR concept is presented and the capability to extract secondary structural details of the model proteins is demonstrated.

Materials and methods

CEIRA substrates

In these experiments, devices made by *cross*-geometry arrays have been used. They have been fabricated accordingly to the general scheme described in section 2.1 "*Fabrication of CEIRA devices*". For these experiments, in order to ensure good reflectance performances, we enlarged the width of the nanoantennas up to 100 nm, while keeping G_x fix at 25 nm. Final structures were 2.04 μm long, 80 nm high and 100 nm wide.

The Plasmonic Internal Reflection (PIR) concept

A schematic view of the Plasmonic Internal Reflection (PIR) concept is reported in Figure 3.14. In the measurement scheme proposed in the figure, the plasmonic elements can be considered as an Internal Reflection Element (IRE). The near-field enhancement induced by the nanoantenna arrays enables the so-called PIR effect, providing the chance to probe samples at the nanoantenna surface via CEIRA microscopy limiting the penetration depth of the IR light. In this scheme, the field enhancement provided by plasmonic substrates decays exponentially into the surrounding background in few tens of nanometers [4], [43]. The possibility to confine the biological sample in the proximity of the plasmonic hot spot at the nanoantennas gaps enables selective local probing of the biological samples, minimizing the interference of water solution, similarly to ATR-FTIR spectroscopy.

PIR-CEIRA measurements were acquired with the same instrumentation used in former experiments. PIR FTIR spectra were collected in reflection geometry by using conventional globar source, using a 15x Schwarzschild objective and setting knife-edge apertures to 50x50 μm^2. Each spectrum was collected from 4000 to 800 cm^{-1} in double side, forward/backward

acquisition mode with a scanner velocity of 40 kHz. 1024 scans were averaged with a spectral resolution of 4 cm^{-1}. Fourier transform was carried out with Mertz phase correction, Blackman-Harris 3-terms apodization function. Spectra were rationed against a background taken on flat gold mirror.

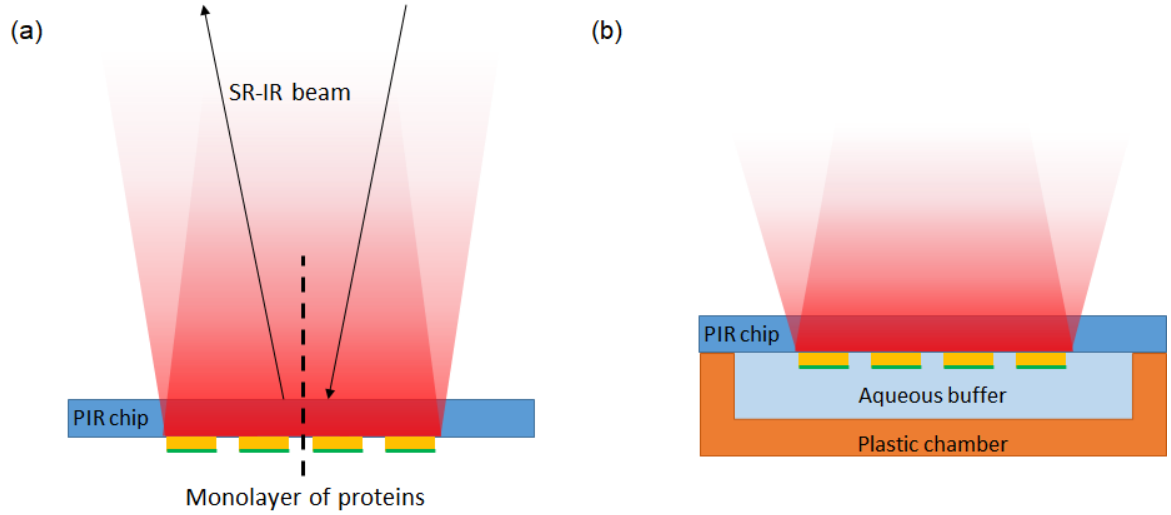

Figure 3.14: (a) Scheme of the Plasmonic Internal Reflection (PIR) geometry of acquisition. The IR beam is focused onto the back of the nanoantenna assemblies, passing through the IR-transparent material (CaF_2). Adapted from [43]. (b) Ideal aqueous chamber for PIR measurements. Images are not in scale.

Protein preparation

Plasmonic devices were functionalized with BSA and ConA via amine coupling, following the protocol detailed in 3.1, "*Amine coupling immobilization procedure*". For the purpose of the liquid measurements, protein monolayers have been equilibrated in a 25 mM HEPES (4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid) buffer with added 2mM DTT (1,4-Dithiothreitol), NaCl 0.2 M and 5% of Glycerol (hereafter HEPES buffer) in order to better resembling the physiological environment.

Fabrication of the liquid devices

CEIRA substrates were integrated into liquid devices used in static mode, which keep the sample fully hydrated during the measurement. Specifically, plastic liquid chambers were fabricated by using 3D printing, according to the design shown in Figure 3.14b. In Figure 3.15a the assembly of our static chamber is reported. Prototype device were fabricated with German RepRap X350 Prp 3D printer in acrylonitrile butadiene styrene (ABS), whereas final devices were obtained using a Form2 by Formlabs, and realized in a proprietary UV curable photopolymer resin (Formlabs).

Final static devices were made by two parts. The lower part, the design of which is reported in Figure 3.15b, contains a well which is filled with either water or HEPES solution, and a slot to hold the PIR device. The upper part is a plastic lid that presents a hole which allows the IR light to pass and to be focused onto the nanoantenna array (highlighted with the red arrow in Figure 3.15c). A soft gasket separates the plastic lid from the back the plasmonic substrate. A picture of the resulting device is reported in Figure 3.15 (c-d).

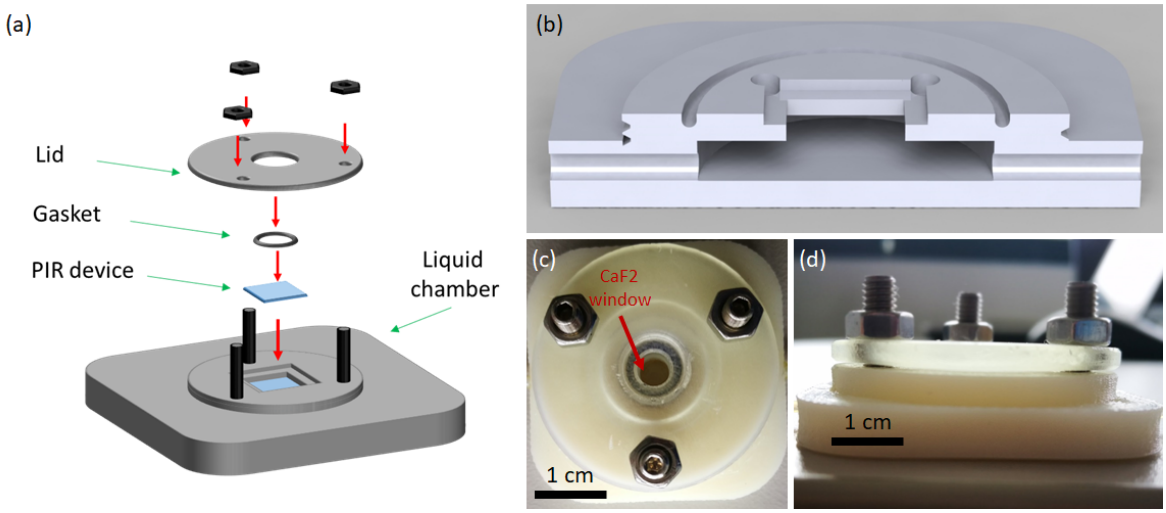

Figure 3.15: Static liquid cell for PIR measurements in water environment. (a) Steps for assembling the static chamber: (i) the PIR device is placed in the slot present onto the top of the lower part of the device, which contains the chamber filled with the final buffer; (ii) the PIR chip is pressed by a gasket and (iii) the lid is placed on top of it. (b) Design of the lower part of the cell: this element contains a well with the solution for equilibrating proteins and it has been designed for accommodate the PIR chip. (b-c) Pictures of the assembled liquid device. The hole which allows the illumination of the sample is highlighted by the red arrow. (b) Top view. (c) Side view.

PIR-CEIRA data analyses

PIR-CEIRA raw spectra were corrected for carbon dioxide and water vapor using OPUS 7.5 routines (BrukerOptics GmbH). Second derivative analyses were made directly on compensated raw spectra. Vector normalized second derivative were computed (Savitzky-Golayfilter, 17 smoothing points) in the region of interest by OPUS7.5 routines.

Results and discussion

The plasmonic behavior of CEIRA devices was characterized in aqueous conditions by integrating the CEIRA substrates into the in-house fabricated static liquid cells filled with both ultrapure water and HEPES buffer solution, more suitable for protein measurements. In

Figure 3.16 are reported the average PIR CEIRA spectra of the two solutions, compared with the resonant response of bare CEIRA antennas in dry condition.

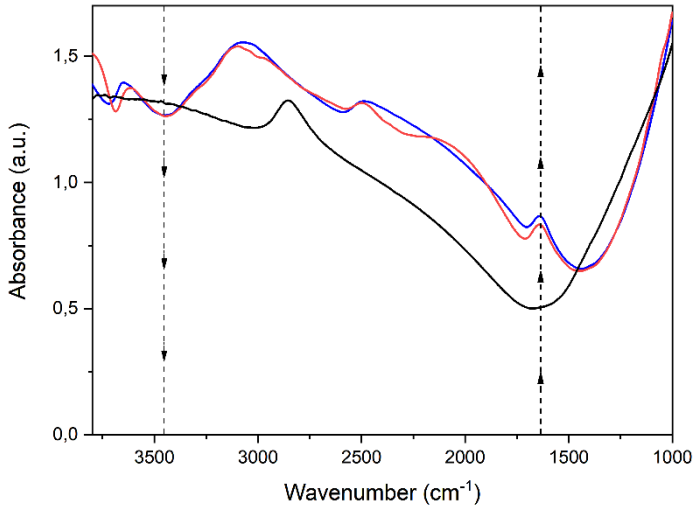

Figure 3.16: PIR CEIRA average spectra of water (blue spectra) and HEPES buffer (red spectra) are matched with the bare-CEIRA response of the system (black spectra), acquired in reflection-PIR mode in dried conditions. Dashed lines highlight water signatures. Direction of the arrows highlights the enhanced H-O-H bending signals (H-O-H bending 1640 cm^{-1} – arrows with positive direction) and the out-of-resonance O-H stretching (at about 3450 cm^{-1} – arrows with negative direction).

In order to understand the trends shown in Figure 3.16, we should remember that PIR signals are collected in reflection mode, while in the former experiments (sections from 3.1 to 3.3) signals were collectd in transmission geometry. It is important to notice that more light is backword reflected when the nanoantennas arrays are excited at the resonace frequency. This leads to decreased transmitted signal at resonance, while increased in case of measurements in reflection geometry. In this thesis, it has been decided to always plot the spectra in absorbance instead of transmittance or reflectance (see section 1.2.3), thus in our plots, curves collected in transmission and reflection geometry show an opposite trend: if in the former transmission experiments, resonant curves appeared as positive peaks and protein enhanced features as deeps in the resonance, in the case of PIR measurements resonant peaks appear as depressions at the resonance and the resonant peaks are positive (see Figure 3.16). In hydrated conditions, the resonant response of plasmonic devices resulted red-shifted of about 100 cm^{-1} as a result of the increased refractive index of the surrounding medium [44]. Red shift of the resonance curve upon nanoantennas interaction with surrounding material have been described in section 1.3.4, stating that CEIRA applications such shift is of minor importance if broad vibrational absorptions are investigated, such as those of amides bands and can be considered negligible [3]. Resonance spectra were characterized by the

contribution of bending of water at about 1640 cm^{-1}. It is important to notice that, due to the geometry of acquisition, signals of water are relatively low, since they originate only by the thin layer within the evanescent field of PIR enhancement [43]. It has been reported that the penetration depth of such evanescent field is about 100 nm, and it decays exponentially with the distance from the gold nanoantenna surface [30]. HEPES buffer and MilliQ water showed similar resonant spectra: HEPES-buffer did not show any strong spectral contribution due to its composition, suggesting that the concentration of its constituents within the enhancement field of antenna is too low for being detected. HEPES buffer was therefore selected for protein measurements.

Once characterized the system in aqueous environment, model proteins were measured in HEPES buffer. After protein anchoring by amine-coupling protocol, the devices were rinsed with the buffer solution in order to avoid the complete drying of the chip surface and then immediately integrated in the liquid chambers. Results of BSA and ConA PIR-CEIRA measurements in buffer are reported in Figure 3.17. As can be appreciated from the figure, Amide II band of both BSA and ConA are detectable also in raw spectra, thanks to their proximity to the "hot spots" of the plasmonic structures, as defined in section 2.4.2. A strong and well-defined signal in the region of Amide I is also clearly detectable. However, the water contribution to this signal cannot be neglected *a priori*. Therefore, at a first glance, protein CEIRA spectra in buffer should present also water features.

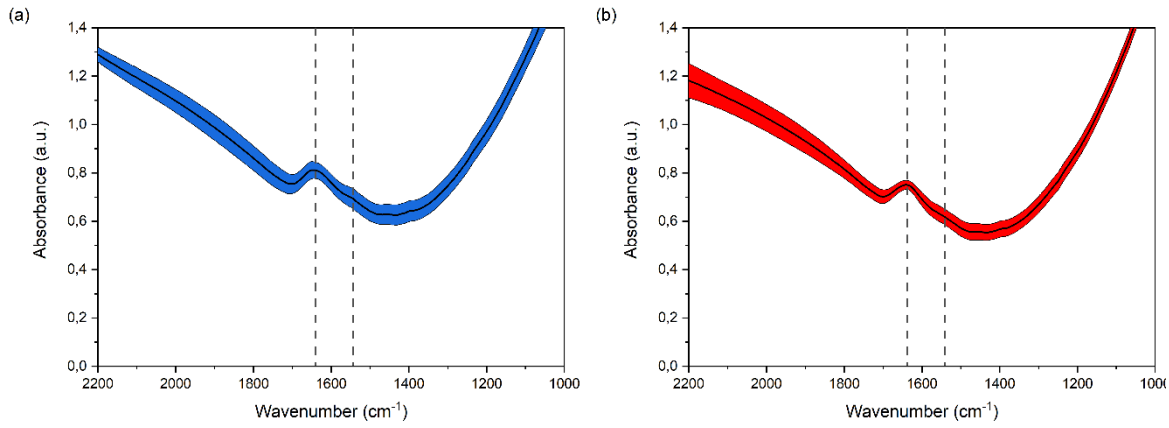

Figure 3.17: PIR-CEIRA acquisitions of BSA (a) and ConA (b) in HEPES buffer. Raw spectra are presented. Shadowed area represents the standard deviation of different acquisitions onto the nine arrays ROI-L1 arrays of two different devices. Black line represents the average spectrum. Gray dashed lines highlight protein amide bands signatures.

In order to understand the actual interference of water in the interpretation of the protein spectra, second derivative analyses have been performed on raw data, as reported in Figure

3.18. The main positions of Amide I and II bands in second derivative spectra are consistent with those observed for dried samples (see Figure 3.12). Specifically, BSA Amide I is peaked at 1654 cm^{-1}, instead for ConA peaks at 1633 cm^{-1}. It is interesting to notice that, for both BSA and ConA there are some spectral variation upon hydration involving smaller contributions in the Amide I range. In particular, for ConA while dried spectra presented a shoulder at about 1690 cm^{-1} relevant mainly for β-turns, in hydrated ones a second contribution at 1672 cm^{-1} is visible, which is considered relevant for native antiparallel β sheet components. This spectral change upon protein hydration, suggests that buffer allows the correct folding of the protein, which probably is hindered in dried conditions. These evidences further underline the importance of working with proteins in the right buffer conditions for enabling their correct folding. Moreover, results point out that water signals do not interfere with protein conformational interpretation: water bending mode contribution, at about 1640 cm^{-1}, which usually dominates protein spectra in solution, cannot be recognized. We can speculate that the outcomes of these experiments are related to the higher ε of proteins with respect to water and their high relative concentration in the near field at the antenna hot-spots.

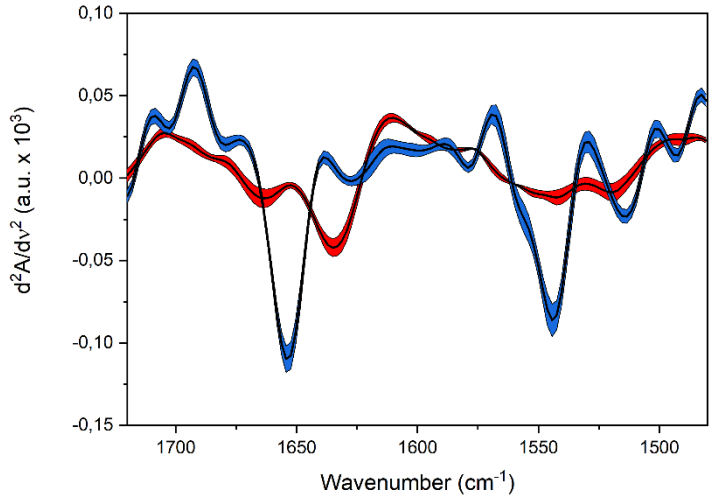

Figure 3.18: Second derivative spectra of raw PIR-CEIRA acquisitions. BSA – blue spectra: ConA – red spectra.

Conclusions

The results here reported demonstrate that, by exploiting the great advantages offered by the combination of the sensing capability of our CEIRA devices with the PIR approach, it is possible to retrieve information on the secondary structure of protein monolayers in aqueous environment without suffering water signals interference. In a more general prospective, the

results presented up to here demonstrate that by exploiting the PIR-CEIRA approach, it is possible to characterize in a fast and label-free way different conformational states of protein monolayers, with very high sensitivity. This evidence opens the possibility to use CEIRA devices for reliable detection of conformational state of proteins in different environmental conditions.

3.5 CEIRA measurements on protein of biomedical interest: Epidermal Growth Factor Receptor

Measurements on model proteins demonstrate that our plasmonic devices allow the detection of conformational states of protein monolayers in a label-free and non-destructive manner, even in aqueous environment. However, moving from model proteins to samples of biomedical interest introduces different challenges, mainly related to protein stability [45]–[47]. In this section, the main results obtained in the application of PIR-CEIRA microscopy for the characterization of the kinase domain of the Epidermal Growth Factor Receptor (EGFR) are presented. EGFR is a membrane protein made by an extracellular domain, the receptor of extracellular ligands, and an intracellular domain, the kinase domain (KD). This protein is involved in the development and progression of different forms of cancer [48], in particular lung cancer [49]. Lung cancer is one of the leading causes of cancer death worldwide [50], [51]. Among the different lung cancer forms, non-small-cell lung cancer (NSCLC) constitutes about 85% of cases of lung adenocarcinoma histological type [52]. Protein mutations that lead to EGFR overexpression or its over activation are the most common driver mutation in NSCLC. EGFR-directed tyrosine-kinase inhibitors (TKIs) against EGFR-mutated NSCLC have been developed and investigated extensively in the past 2 decades [53]–[56] and first- and second-generations of TKIs guaranteed a dramatic success in clinics [57]. Despite this, the development by tumor cells of a dominant resistance mechanism upon different punctual mutations has limited the long-term efficacy of these targeted therapies [58]. One on the most common EGFR mutation is the threonine–methionine substitution at position 790 of EGFR (T790M mutant). Specifically, threonine 790 is the gatekeeper residue in EGFR, because its key location at the entrance to a hydrophobic pocket in the back of the ATP binding cleft makes it an important determinant of inhibitor specificity in protein kinases [56]. The T790M alteration remains the only validated second-site resistance mutation for EGFR, asking for focused research to understand and overcome EGFR T790M-mediated resistance [59].

In the past years, the whole tridimensional structure of the kinase domain of EGFR in its wild type, *wt*, and mutated, *mut*, forms have been deeply investigated by means of X-ray crystallography [60]. In particular, 3D structure of EGFR is characterized by a bi-lobate-fold, which is typical of protein kinase domains [61], [62]. A picture of the XRD structure of EGFR APO is reported in Figure 3.19.

The figure clearly shows that the two lobes are quite different both in dimensions and conformation. The NH_2-terminal lobe (N-lobe) is smaller than the COOH terminal lobe (C-lobe), and it is composed by mostly β-sheet motifs and one α-helix, called αC, which, in the active state of EGFR kinase, is adjacent to the ATP-binding site in the so called "αC-in" conformation [61]. Instead, the larger C-lobe is mostly α-helical folded, with an "activation loop" that is highly conformationally variable, depending on the activation status of the protein [55]. The ATP-binding site is located between the two lobes, which are separated by a cleft similar to those in which ATP, ATP analogues, and ATP-competitive inhibitors bind [49]. The secondary structure obtained by hydrogen bond estimation algorithm (DSSP - Dictionary of Secondary Structures in Proteins) [64] of the kinase domain of APO/EGFR is the following: 36% helical structure (corresponding to 16 helices, involving 121 residues), 15% beta sheet strands (corresponding to 18 strands involving 52 residues) [61].

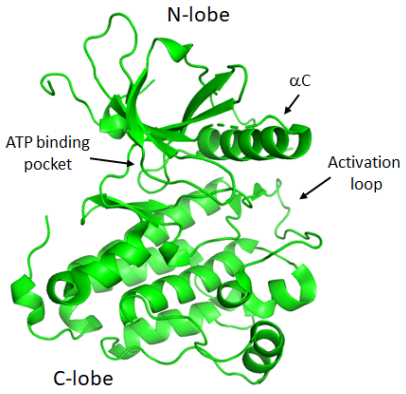

Figure 3.19: XRD reconstruction of the 3D structure of the unbounded (APO) kinase domain of EGFR in its wild type form (2GS2). Image adapted from [63].

Is important to notice that, X-ray structures collected in Protein Data Bank (PDB) of T790M mutated form show minimal secondary structure difference with respect to the wild type. By looking to its DSSP predictions reported in PDB, EGFR T790M estimated secondary structure is 32% helical (12 helices; 105 residues) and 15% beta sheet (18 strands; 51 residues) [58]. Due to its importance in medical research, we focused our attentions on this

case study. The kinase domain (KD) of EGFR in both wild-type, *wt,* and T790M mutated, *mut,* versions were produced at the Protein Production Facility at Elettra by recombinant technology. From here on, with EGFR we refer only to its kinase domain. It is therefore clear that the possibility to perform conformational studies onto such samples in a fast and not material consuming way can increase our capabilities of understanding the bases of the molecular activity upon protein mutations.

Three different experiments are here presented. First, the main challenges in working with conformational sensitive proteins are highlighted, discussing an experiment aimed to discriminate the conformational differences between EGFR *wt* and *mut* (section 3.5.1). Second, the optimization of a new strategy for immobilizing sensitive proteins in an ordered way onto gold nanoantennas without compromising their native conformation is presented (section 3.5.2). Finally, the third section is dedicated to dynamic measurements on the EGFR upon ligand-protein interaction (section 3.5.3).

3.5.1 Wild-type and mutated form of EGRF: conformational differences

With the aim to prove the sensitivity limit of the CEIRA approach through the detection of minimal conformational variations, EGFR *wt* and T790M *mut* have been anchored onto two distinct CEIRA devices with *cross*-geometry with the conventional amine-coupling procedure. The results of the measurements, performed both in dry and hydrated conditions, were compared in order to highlight if it was possible to discern the subtle spectral difference related to conformational changes induced by the punctual mutation.

Materials and methods

Protein production

DNA encoding for the kinase domain residues 696–1022 of the human EGFR wild-type was inserted in pFB-LIC- Bse and pFB-6HZB transfer vectors (kindly provided by O. Gileadi, SGC-Oxford). The T790M mutant was obtained by Q5 site directed mutagenesis (NEB) and sequence verified. Recombinant baculovirus was generated using the Bac2Bac system (Invitrogen) by transfecting Sf9 insect cell line, a clonal isolate derived from the parental Spodoptera frugiperda cell line IPLB-Sf-21-AE, seeding cells in 6-well plates (1.5 x 10^6 cells/well) with FuGENE HT (Promega). The high titer virus was used for protein expression in Sf9 cells grown in ESF 921™ (Expression Systems) at 27°C and collected at 72 hours' post-infection. For protein purification, cells were doubly homogenized (Emulsiflex – Avestin) in lysis buffer (50 mM Tris pH 8.0, 500 mM NaCl, 10% glycerol, 10 mM imidazole,

2 mM TCEP) supplemented with 10 µg/ml DNaseI, 1 mM $MgCl_2$ and protease inhibitor cocktail (complete EDTA-Free, Roche).

Proteins were purified by Ni^{2+}-NTA affinity chromatography (Qiagen) followed by TEV cleavage (performed in ON dialysis at 4°C against 25 mM Tris pH 8.0, 0.2 M NaCl, 5% glycerol, 5 mM DTT), ion exchange chromatography on Fractogel® EMD TMAE (M) column (Merck) and size-exclusion chromatography on Superdex 75 (10/300 GL, GE Healthcare) in 25 mM (4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid (HEPES) pH 8, 250 mM NaCl, 5% glycerol, 2 mM TCEP. Stock protein samples were flash frozen at the final concentration of 0.4 –1 mg/ml and stored in aliquots at -80°C until further use. Typical yield of purified EGFR KD constructs was of 0.6 - 1.8 mg/ml x 10^9 Sf9 cells.

Experimental setup and data analyses

The experimental setup was the same of the former experiments (section 3.3). Spectra were acquired by PIR-CEIRA microscopy using the same acquisition parameters of former experiments (section 3.4). Both dry and hydrated samples were measured. Static liquid cells were used for measures in HEPES buffer solution. Raw spectra of dried acquisitions were corrected as reported in section 3.1, while PIR-CEIRA acquisitions were corrected and treated as reported in section 3.4.

Results and discussion

Preliminary observation based on the experimental results of CEIRA on EGFR *wt* and 790M *mut* proteins, for both dry and hydrated conditions, is that the EDC/NHS functionalization protocol did not work well for EGFR. For two on three arrays, CEIRA spectra exhibited very low signals, hard to identify, hinting to a low efficiency and repeatability of the amine-coupling functionalization protocol for these proteins.

For the arrays that gave detectable protein signals, according with the data analysis approach followed in the previously reported experiments, spectral differences between EGFR *wt* and T790M *mut* proteins were evaluated by means of second derivative analysis. First, dried measures were considered (see Figure 3.20). Second derivatives show clear differences in the Amide I region for *wt* and *mut* forms. Amide I of both proteins displays a prominent contribution peaked at 1654 cm^{-1}, which is diagnostic for the abundancy of α helix structural components. In addition, a weaker peak can be seen at 1685 cm^{-1}, which can be assigned to the presence of turns and loops (usually between 1670-1695 cm^{-1}) and signals from antiparallel β sheet motifs (usually between 1670-1685 cm^{-1}) [36]. Amide I of EGFR *wt* spectrum is characterized by a second strong contribution at 1632 cm^{-1}, which we tentatively

assigned to native β sheet motifs. This contribution is almost suppressed in EGFR T790M spectra. Amide II band, known to be less sensitive than Amide I to protein secondary structure [64], has very similar shape in both cases: it is characterized by a preeminent peak at 1545 cm^{-1} which is characteristic of α helix motifs [37] and a second contribution at 1517 cm^{-1}, which can be assigned to Tyrosine residues [36].

The huge spectroscopic variations highlighted by CEIRA spectra between *wt* and T790M *mut* forms of EGFR is clearly not consistent with X-ray crystallography outcomes, and therefore cannot be justified by the punctual mutation. Since the 3D structure of EGFR protein and its mutants is very sensitive to environmental conditions, in a first instance we associated the discrepancy to the dried conditions used for spectra acquisition, a state far from the physiological one that could have affected the mutated form more than the wild-type one.

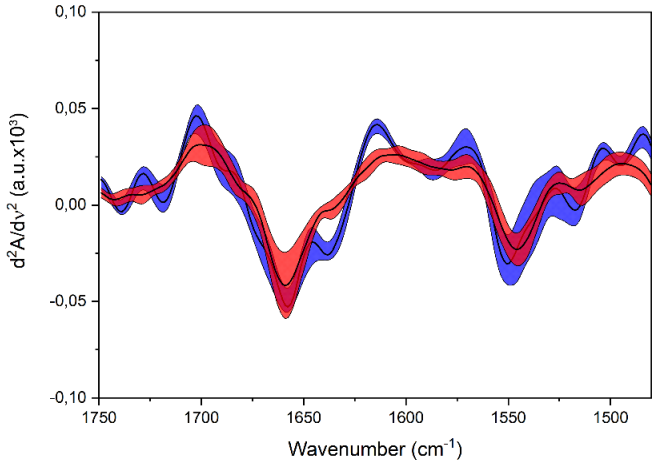

Figure 3.20: Second derivative spectra in the amide L1 ROI of nine PIR-CEIRA acquisition in dry conditions of EGFR *wt* (blue spectra) and EGFR T790M (red spectra). Shadowed area represents the standard deviation calculated onto nine acquisitions, while the black line represents the average spectrum.

In order to verify this hypothesis, EGFR *wt* and T790M *mut* proteins were measured in HEPES buffer. The experimental results are reported in Figure 3.21. As can be seen in the figure, the resulting spectra differ from those obtained in dried conditions. EGFR *wt* spectrum is characterized by an intense contribution at 1629 cm^{-1}, and by a second narrowed peak centered at the 1654 cm^{-1}. Essentially, the two contributions identified in the dry conditions for EGFR *wt* are both present, but in different proportions. The spectrum of T790M *mut* protein is characterized by a single wide peak, that seems to convolve the

contribution already highlighted for hydrated EGFR *wt*. The broadening is also reflected by the spectral shape of EGFR *wt* Amide II band.

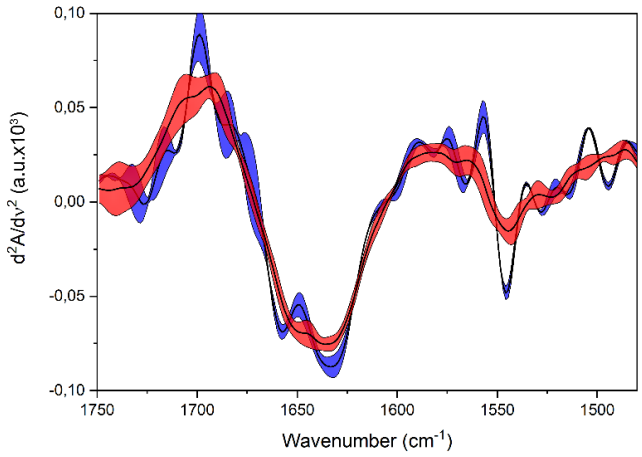

Figure 3.21: Second derivative spectra in the amide ROI L1 of different PIR-CEIRA acquisition in HEPES buffer of EGFR *wt* (red spectra) and EGFR T790M (blue spectra). Shadowed area represents the standard deviation calculated onto 10 different acquisitions, while the black line represents the average spectrum.

Despite the noise in the spectra, the data presented in Figures 3.20 and 3.21 demonstrate that proteins were most likely misfolded both measurement conditions. Furthermore, Amide I of both *wt* and *mut* forms is broadened. This may be consequent to a higher contribution of water than to what observed with ConA and BSA in section 3.4. This is possibly due to the less efficiency of the anchoring protocol, as already stated at the beginning of this section. Both EGFR proteins employed for the measures, *wt* and *mut*, were produced by recombinant technology. Such proteins are very sensitive to buffer exchange and they could undergo to misfolding and precipitation upon aggregation in solution. The amine coupling employed for protein anchoring requires a passage for protein activation in a sodium acetate buffer at pH 4. Such condition is probably too harsh for this kind of proteins, inducing aggregation before anchoring. This hypothesis allows interpreting the experimental results. Proteins aggregates do not deposit homogenously onto the nanoantennas, determining the large variability among the measurements, the very low signals for the large majority of the measured arrays and a larger contribution of water, that interfere with the interpretation of Amide I bands. In this light, the spectral broad and intense contribution detected at about 1630 cm^{-1} should be attributed not only to native β sheet motifs but also to protein aggregation [65].

Conclusions

The presented results show that reliable conformational measurements of both EGFR *wt* and *mut* were mostly hindered by the adopted immobilization procedure, too harsh for efficiently anchoring EGFR. If these conditions are still acceptable for model proteins, they become prohibitive for proteins that require specific environmental conditions for maintaining a stable conformation, such as most of the proteins of biomedical interest. Thus, in order to make reliable protein measures for such complex systems, it is mandatory to find a functionalization protocol which guaranties the right protein folding and, possibly, a selective orientation of the proteins onto the gold surface of nanoantennas.

3.5.2 A new protocol for immobilizing protein of biomedical interest: the NTA/His-Tag recognition

As previously highlighted, our CEIRA devices have been developed for the analysis of proteins of biomedical interest, often produced by recombinant technology. Such proteins are usually made with a terminal "tag" linked at a specific position in the protein construct, which is used for selectively separate the expressed protein during the purification passages [66]. The tags are usually located at the N- or the C-term of a protein, in a region that does not interfere with its right folding. Among the different possible tags, polyhistidine-tags are often used for allowing affinity purification of genetically modified proteins [67]. A polyhistidine-tag (His-tag) is a standard amino acid motif in proteins that consists of at least six histidine (His) residues. His-tagged proteins are usually separated from other proteins by immobilized metal affinity chromatography (IMAC) [68]. Supports, i.e. beaded agarose or magnetic particles, are derivatized with chelating groups to immobilize the desired metal ions. Such ions act as ligands for binding and purification of biomolecules of interest. The chelant most commonly used as ligands for IMAC is nitrilotriacetic acid (NTA), which can be loaded with the desired divalent metal, in particular Ni^{2+}. Nickel generally provides good binding efficiency to His-tagged proteins, thus, his-tagged proteins bind the Ni-NTA in the separation column, while other proteins do not or bind only very weakly. Those unspecific proteins can be removed by washing the carrier with an appropriate buffer in the so-called stationary phase. Once separated target his-tagged proteins from the rest of the solution, bounded proteins are displaced from their linker by using a competitor for the NTA: the imidazole. By this eluting phase, is possible to obtain proteins with high purity. The His-tag is commonly cleaved by Tobacco Etch Virus (TEV) protease at the end the purification steps [69], [70]. The same affinity principle used for IMAC can be exploited for efficient

immobilization of proteins on gold surfaces [71]. The immobilization implies the creation of a SAM of thioled nitriloacetic acid (SH-NTA) moieties onto gold surfaces, the partial chelation of Ni^{2+} by NTA moieties and finally the link of His-tagged proteins [72] as schematically shown in Figure 3.22.

This strategy introduces different advantages in comparison to the amine-coupling procedure: it allows maintaining the protein of interest in the right buffer and it guaranties the formation of a monolayer of iso-oriented proteins. It is indeed noteworthy to mention at this stage that the former procedure (amine-coupling) can implicate not only protein N-terminus but also the amine groups of lateral chains of amino acids such as arginine and lysine. This implies that the anchored proteins will be not iso-oriented, as they will be in the case of His-tag functionalization protocol. We therefore decided to take advantage form His-tag method for the conformational analysis of EGFR *wt*. To this aim, we analyzed in hydrated conditions EGFR *wt* protein His-tagged at the N-terminal and immobilized onto the gold surface of CEIRA device nanoantennas via SAM of SH-NTA. In addition, we tested the possibility to follow the dynamic binding of the target protein onto our CEIRA devices, in order to demonstrate the possibility to perform dynamic analyses *in situ*.

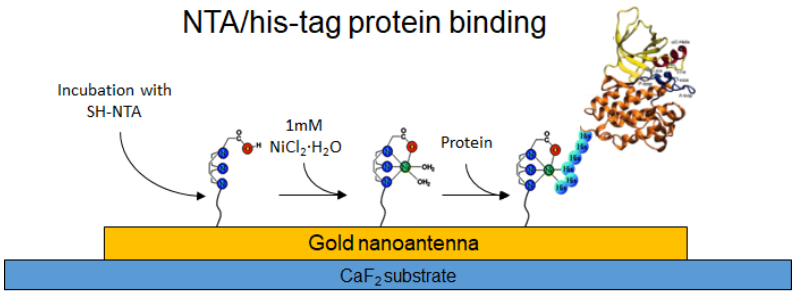

Figure 3.22: Scheme of SH-NTA protein binding steps. After the formation of a SH-NTA SAM onto the gold surface of nanoantennas, the SAM is equilibrated with a solution of $NiCl_2$. Ni^{2+} metal ions coordinate with the NTA molecules and water. Finally, his-tagged proteins are added in solution and they will link to the gold surface via NTA/Ni^{2+}-His-tag affinity.

Materials and methods

Cross-geometry devices for CEIRA-PIR microscopy were fabricated with the same parameters of the former measures in water conditions (See section 3.4). EGFR wild type proteins were expressed as reported in section 3.5.1 "*Protein production*". Proteins were purified by Ni-NTA affinity chromatography (Qiagen) without removing the his-tag at the end of the process.

<u>His-tag functionalization protocol</u>

A scheme of the functionalization protocol is reported in Figure 3.22. After nanofabrication and gold activation by O_2 plasma cleaning, patterned CaF_2 optical windows were incubated 10 μM NTA-EG_6-$(CH_2)_{16}$-SH (SH-NTA hereafter) solution ON at 4°C in order to induce the formation of a SAM onto the gold surface of nanoantennas. Then, CaF_2 substrates were rinsed with 95% ethanol and dried under a nitrogen stream. The SAM of SH-NTA was equilibrated with a 1mM $NiCl_2 \cdot H_2O$ solution for 30 minutes in order to promote first nickel chelation and then subsequent protein immobilization. For preliminary measurements, EGFR *wt* His-tagged proteins were chemically linked to the SAM before the integration into the device. Proteins were immobilized by loading the structures with 80 μL of 100 nM solution of His-tagged EGFR *wt* for 1 hour. Excess of protein was then rinsed with buffer (25 mM HEPES buffer with added 2mM DTT (1,4-Dithiothreitol), NaCl 0.2 M and 5% of Glycerol). Once prepared, samples were integrated within the fluidic devices and immediately measured. For dynamic measurements, functionalized CEIRA substrates were directly loaded within the fluidic devices for *in situ* protein binding measurements, flushing within the device the protein solution.

Fabrication of fluidic device

For preliminary static experiments, the used device for PIR CEIRA measurements is the one reported in Figure 3.15.

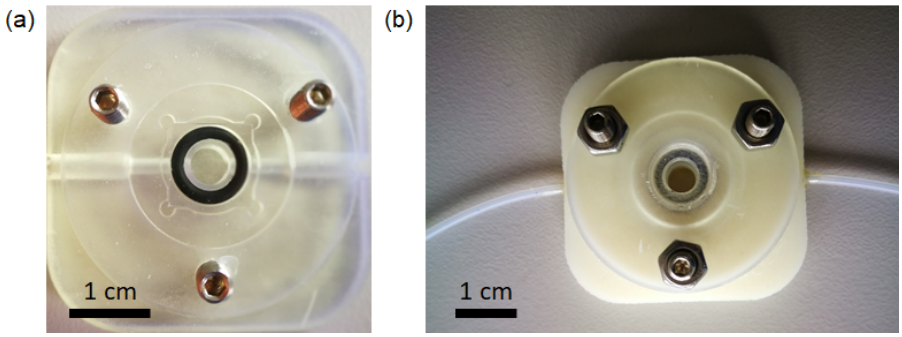

Figure 3.23: Images of the fluidic device for dynamic PIR-CEIRA measurements *in situ*. (a) Picture of the lower part of the device. The inlet and outlet channels are visible through the transparent plastic material. (b) Assembled final device. Buffer exchange was tuned by aspiration trough a micro pump.

In order to perform dynamic analyses, functionalized CEIRA substrates were integrated into a fluidic device. The design of such devices, is based on the same scheme of former static devices, integrated with two channels for the liquid exchange. Such plastic fluidic chambers were fabricated by using 3D printing. In Figure 3.23 the final device is shown.

CEIRA measurements and CEIRA data analyses

Spectra were acquired by PIR-CEIRA microscopy, using the same instrumentation and the same acquisition parameters of section 3.4 "*The plasmonic internal reflection concept*". PIR-CEIRA protein spectra data analysis was made as reported in section 3.4 "*PIR-CEIRA data analysis*".

Results and discussion

EGFR was produced with a tail of 6 histidine residuals at the N-terminal at the Structural Biology Laboratory at Elettra. For the purpose of our measurements, the His-tag was maintained in order to drive the linking of the EGFR onto the gold surface of our plasmonic *cross*-geometry devices via hexahistidine (His_6)-Ni^{2+}: nitriloacetic acid (NTA) methodology. The spectral profile of CEIRA devices was characterized after SH-NTA functionalization in order to verify the potential interfering effects of the SAM moieties in the spectral region of interest for Amide I protein analysis. Resulting spectra, collected in liquid static devices filled with HEPES, were very similar to those obtained for bare antennas in deionized water (see Figure 3.16), with a single band centered at about 1640 cm^{-1}, consistent with the presence of water (See Figure 3.24, blue curve).

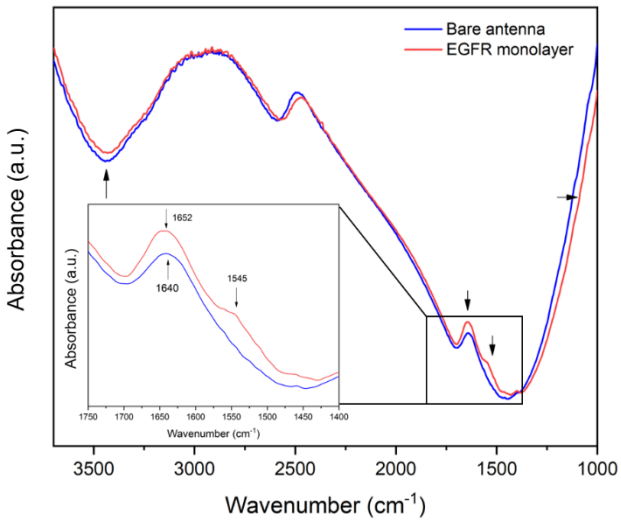

Figure 3.24: PIR-CEIRA measurements on SAM-functionalized antennas (blue spectra) and antennas functionalized with *wt* EGFR via SH-NTA/His-tag (red spectra). Vertical downward arrows highlight protein amide bands, while upward arrows highlight aqueous contributions. The horizontal arrow highlights the resonant shift induced by protein linking to the sensor. In the inset, the position of protein amine bands is highlighted by arrows.

In parallel, ATR-FTIR measurements of a stock solution of SH-NTA were performed, and they confirmed that SH-NTA is substantially free from any intense feature in the L1 ROI.

Only a weak band was visible at 1400 cm^{-1}, which does not interfere with our measurements (data not shown). In Figure 3.24 the resulting CEIRA spectra of His-tag SAM-functionalized antennas (blue curve) and EGFR-functionalized antennas (red curve) are compared. Amide II band of EGFR can be clearly appreciated.

The repeatability of the functionalization has been validated by measuring several arrays on diverse CEIRA chips (Figure 3.25). Different from what reported for the amine-coupling experiments, the good measurements repeatability deduced by Figure 3.25 evidences a comparable coverage of the arrays and that proteins are in the same conformational status.

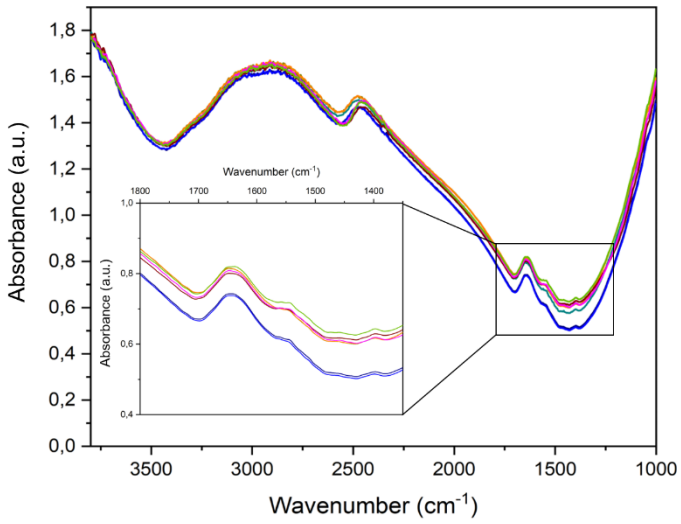

Figure 3.25: PIR-CEIRA measures of EGFR *wt* in HEPES buffer onto different arrays of three different CEIRA devices. In the inset, protein signals are highlighted.

In addition, second derivative analysis confirmed the preservation of the same conformation for all the reported measurements (see Figure 3.26). The second derivative of Amide I of EGFR *wt* is characterized by a prominent contribution peaked at 1654 cm^{-1}, which reflects the abundancy of α-helix components [37]. A well-defined shoulder can be seen at 1634 cm^{-1}, which can be assigned to native β-sheet motifs [73]. This evidence further confirms the hypothesis that protein misfolding and aggregation contribute to the broad spectral band at about 1630 cm^{-1}, identified for EGFR *wt* immobilized by amine-coupling (section 3.5.1). Finally, a small shoulder can be detected at about 1679 cm^{-1}, position usually assigned to turns and loops [74]. Amide II band of protein spectra in second derivative is characterized by a relevant peak at 1545 cm^{-1}, which can be assigned to α-helix moieties and by a second strong contribution ate 1515 cm^{-1}, relevant for Tyr residues [33]. In contrast with the case of model proteins (BSA and ConA) measured in sections 3.3 and 3.4, for EGFR proteins we could not measure a reference spectrum by conventional ATR-FTIR

spectroscopy. Indeed, while BSA and ConA solutions can be prepared in simple water buffer, for preserving the right folding of EGFR it is mandatory to maintain the protein in a physiological-like condition. ATR-FTIR acquisitions are performed by letting the protein solution drying onto the IRE crystal in order to concentrate the sample within the evanescent field. In such a way, also buffer additives concentrate as well, in particular HEPES and glycerol, therefore interfering with the measurements. Also, due to the low concentration of the starting solution (less than 10 μM), protein signatures were overwhelmed from those of the buffer moieties (data not shown). This further highlights the advantage offered by PIR-CEIRA measures, which are not affected by the presence of such buffer moieties.

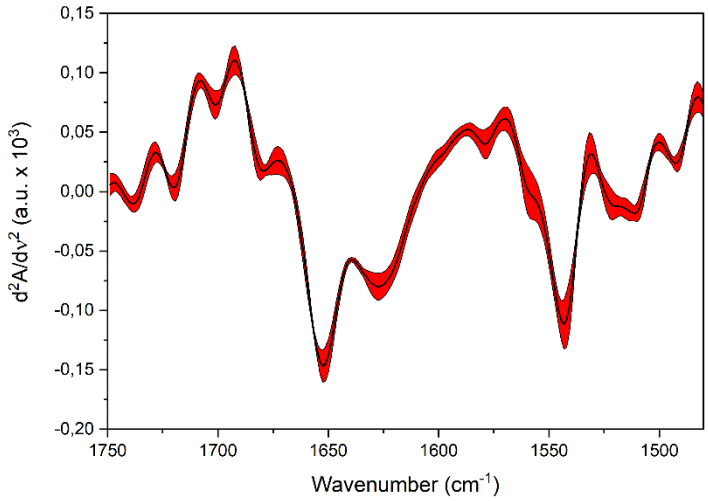

Figure 3.26: Second derivative spectra obtained from raw PIR-CEIRA measures of EGFR *wt* functionalized chips in HEPES buffer. Shadowed area represents standard deviation of measurements, while the black line represents the average spectrum.

As a further evidence of the advantages offered by the protein immobilization method, we tested the possibility to follow the dynamic binding of protein onto the CEIRA chips. Plasmonic devices functionalized only with SH-NTA were integrated in a fluidic device and EGFR *wt* in HEPES buffer was directly injected within liquid chamber. Protein solution was first injected and kept *in situ* for 20 minutes and then fresh buffer solution was flushed inside the device chamber in order to remove the eventual excess of protein non-specifically bounded to nanoantennas. Spectra were acquired continuously and each spectrum was accumulated for 3 minutes (over a total measurement time of 60 minutes). The time evolution of the resonance curves in the amides spectral window is reported in Figure 3.27a. The increase of amide signals during the protein injection is well visible as a consequence of protein accumulation onto the CEIRA substrate. The red-shift of the resonance peak

induced by the variation of the refractive index of the environment surrounding the nanoantennas also proves the protein binding. For better highlighting the dynamic of protein binding, the time evolution of the absorbance signal of Amide I band, evaluated as Amide I band height at 1654 cm^{-1}, was considered and it is shown in Figure 3.27b. It is possible to observe that the Amide I signal rapidly increases in the first 10 minutes after protein injection and it reaches a plateau after 20 minutes, suggesting that after this time a homogenous coverage of all gold structures is achieved. Indeed, after extensive buffer rinsing, the Amide I intensity remains substantially stable.

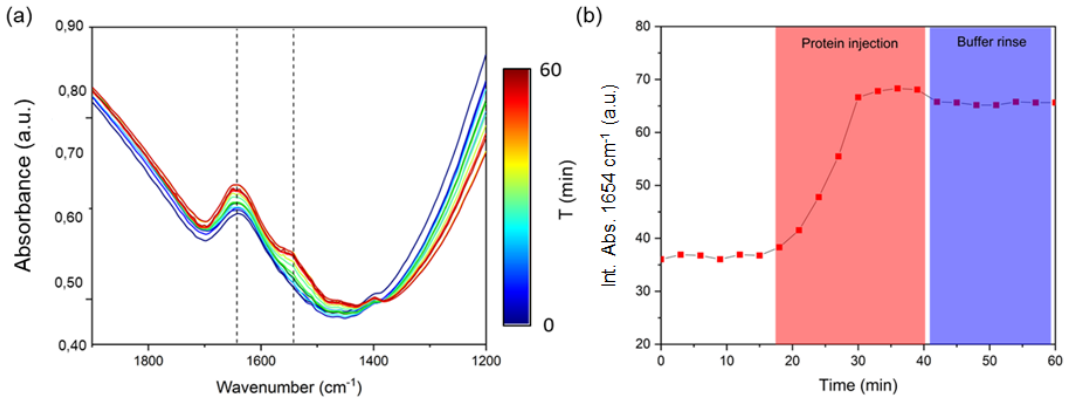

Figure 3.27: Protein dynamic binding onto the CEIRA substrates via SH-NTA/His-tag immobilization. In the figure, the time course of the *in-situ* binding of EGFR *wt* is reported. (a) Increasing of the Amide I-II band signals in the amide spectral window in time. The time scale is of 60 minutes. (b) Quantitative evaluation of the protein binding *in situ* by following the increasing of the amide I band peak intensity (height of 1654 cm^{-1} peak). Protein injection and buffer rinse are highlighted as shadowed areas.

Conclusions

The experimental results here discussed demonstrate that the His-tag immobilization protocol is a powerful method for protein anchoring, which allows to preserve the protein conformational status and to obtain ordered and oriented protein monolayers. Of course, this immobilization method is limited to those proteins that are produced by recombinant technology. It is important to remember that, proteins obtained by recombinant technology are the main samples employed in pharmaceutical and biomedical research: of course, the real target in this field, and in the field of protein conformational studies are proteins isolated from natural sources, but procedures employed for their purification are generally inefficient and time-consuming. The possibility to isolate and amplify the DNA which encodes for functional proteins (both wild type or mutant variants) together with the capability to

integrate it, by means of genetic vectors, in cells in culture provide the possibility to express target molecules in higher quantities and with higher efficiency.

The presented results demonstrate that our CEIRA devices allow performing *in situ* and *operando* protein conformational analysis. It is therefore possible foresee the exploitation of CEIRA-PIR microscopy for dynamically following protein-ligand interaction *in situ*, opening new possibilities in the understanding of the dynamics of these fundamental interactions.

3.5.3 Conformational changes induced by ligand-protein interaction investigated by CEIRA-PIR microscopy

Direct investigation of protein-ligand interactions is of fundamental importance for understanding mechanism and dynamics of functional proteins in health or disease development and treatment [75], [76]. Generally, biophysical methods, such as X-ray crystallography [77] and NMR spectroscopy [78], provide this information, but they are very time and material consuming [79]. Due to these limitations, great interest is devoted to the development of complementary biophysical methods for achieving this goal. As we already discussed in Chapter 1, IR spectroscopy provides specific information on conformational changes in a *label-free* fashion and possibly dynamic way. The use of ATR-FTIR spectroscopy as a tool for detect conformational variations have been widely demonstrated in literature [36], [38], [74]. Moreover, the possibility to detect small variation on protein monolayers, or few layers anchored onto multiple reflection ATR crystals have been recently demonstrated by J. Güldenhaupt *et al* [80]. The authors proposed a novel approach investigating a ligand-induced secondary structural change onto a chaperonin covalently bonded to the surface of a germanium ATR crystal and brought in contact using a fluidic setup with different ligands. Data analysis relied on difference spectroscopy to highlight small variations of the spectra occurring as a consequence of the binding of the ligand. Then the ATR-FTIR data were compared with X-ray crystallography outcomes for their validation. This is a very interesting and thorough work, however, one of the main limitations of their experimental setup is that the use of multiple reflection ATR crystals is mandatory for providing good spectra. Such crystals are quite long and wide, and different hundreds of μL of protein solutions are required for covering all the crystal surface (some cm^2). This protocol is not always compatible with protein of biomedical interest which rate of production is very low and which cannot be produced in high volumes with appropriate concentrations.

With our former experiments, we have demonstrated that PIR-CEIRA approach can reach detection limit up to attomoles, that is orders of magnitude better than ATR sensitivity, by using ultralow quantities of target material and few tens of μL of protein diluted solutions thanks also to the small sensing areas of our devices. In addition, taking advantage from the presented SH-NTA/His-tag functionalization method, we demonstrated the capability of track dynamic processes, envisaging the possibility to follow eventual protein conformational variation upon interaction with a specific inhibitor/ligand.

To confirm this hypothesis, EGFR was induced to interact with a specific tyrosine-kinase inhibitor (TKI), Lapatinib. The structure of Lapatinib is reported in Figure 3.28.

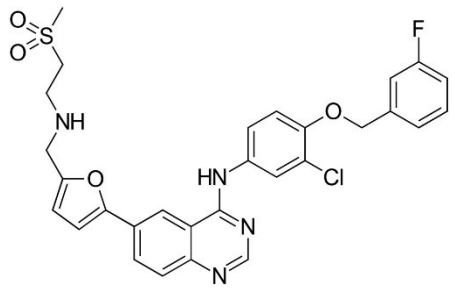

Figure 3.28: Molecular structure of Lapatinib.

It is an oral dual TKI which molecular weight is 581,058 g/mol. It targets EGFR and human epidermal growth factor receptor-2 (HER2) [53], and it is used in different anticancer therapies [81], [82]. Lapatinib favors extended downregulation of phosphorylation activity since it has a slow rate of dissociation from EGFR [83]. The interaction of Lapatinib with EGFR in the last years was studied by X-ray crystallography [84]. In Figure 3.29, the 3D structures of EGFR in its APO form (that is the inactive EGFR lacking the ATP cofactor) and EGFR co-crystallized with Lapatinib, are compared. Significant differences in the orientation of different parts of the protein have been reported upon drug interaction, in particular the position of the N-lobe and the αC motif [84]. These differences mainly involve folding changes which belongs to tertiary structure, while the overall secondary structure of APO-EFGR and Laptinib/EGFR is very similar. As reported in the introduction of section 3.5, the secondary structure of APO/EGFR, estimated by DSSP, is 36% helical structure, and 15% beta sheet strands [63], but these percentages slightly change for Lapatinib/EGFR, and are 34% in helical structures and 14% beta sheet [84].

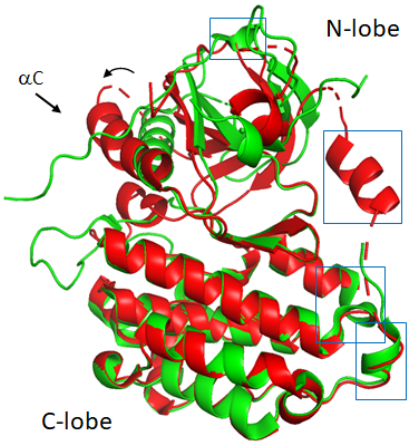

Figure 3.29: Overlap of the XRD structures of APO/EGFR and lapatinib/EGFR (by [84]). Secondary structure differences are highlighted by blue square.

In the following, we aim to demonstrate that PIR-CEIRA microscopy provides the possibility to identify the subtle conformational changes of proteins consequent to protein-inhibitor interaction, by exploiting the plasmonic enhanced amide signatures from Mid-IR CEIRA devices, and resolving their spectral content with difference spectra analyses [85].

Materials and methods

Cross-geometry CEIRA devices were fabricated as reported in section 3.4 "*CEIRA substrates*" and they were mounted onto a fluidic device in order to control the level of the buffer and inject the inhibitor. The *wild type* form of the human kinase domain of EGFR was produced and purified as reported in *materials and methods* of section 3.5.2. Thus, protein was bounded onto the gold surface of plasmonic nanoantennas arrays via SH-NTA/his-tag recognition, as described 3.5.2 "*His-tag functionalization protocol*". Spectra were acquired by PIR-CEIRA microscopy at SISSI beamline, using the same instrumentation and the same acquisition parameters section 3.5.2.

CEIRA data analyses

Raw spectra were corrected for carbondioxide and water vapor using OPUS 7.5 routines (BrukerOptics GmbH). No spline line correction has been done before second derivative analyses. Vector normalized second derivative were computed (Savitzky-Golayfilter, 17 smoothing points) in the region of interest by OPUS7.5 routines. In addition, PIR-CEIRA raw spectra have been cut and baseline-corrected in the 1800-1400 cm^{-1} range, and normalized in the same spectral region. Principal Component Analysed (PCA) was

performed using HyperSpec and stats packages in R statistical environment [86]. After PCA, spectra have been reconstructed using only the first 4 loadings provided by PCA, in order to further reduce the spectra noise. Difference spectra have been done upon spectra noise correction. The average spectrum of the APO/EGFR acquisitions has been used as reference for subtract each EGFR/lapatinib spectrum.

Results and discussion

Plasmonic devices were integrated within the fluidic chambers and EGFR *wt* proteins were loaded onto the CEIRA devices in HEPES buffer. Upon protein binding to the CEIRA substrates surface, they have been immediately measured. In parallel, Lapatinib (GW-572016) ditosylate (Sigma Aldrich) was dissolved in dimethyl sulfoxide (DMSO) and then diluted in deionized water. After the acquisition of EGFR-apo spectra, Lapatinib solution was fluxed within the fluidic device with a 3 molar excess respect to the protein for 30 minutes in order to interact with EGFR protein, and then PIR-CEIRA measures have been performed. At a first glance, the analysis of second derivative spectra was used for identifying conformational details of proteins. Results are reported in Figure 3.30 and demonstrate that the protein status was preserved in both cases (See Figure 3.26 for comparison).

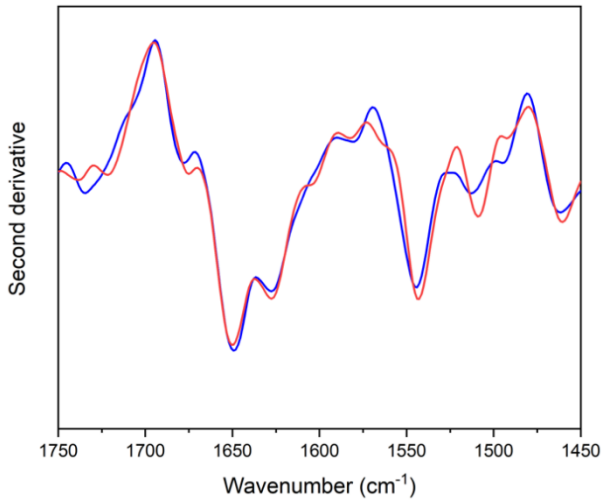

Figure 3.30: Average spectra of the second derivative of EGFR *wt* in its apo (blue spectra) and lapatinib-binded (red spectra) forms. Spectra are calculated on the base of 7 different acquisitions. Reported spectra are not normalized.

In addition, by the analysis of second derivative spectra of Lapatinib/EGFR and EFGR, very tiny spectral variations among the two constructs could be identified. Nevertheless, in order

to exclude the contribution of the spectral noise in second derivative profile modulation, we decided to implement data analysis with more advanced methods.

Principal component analyses (PCA) of the collected spectra was performed (Figure3.31) in order to highlight if reliable spectral variation really could be detected after protein-ligand interaction as a consequence of the minimal secondary structure modifications expected from X-ray data.

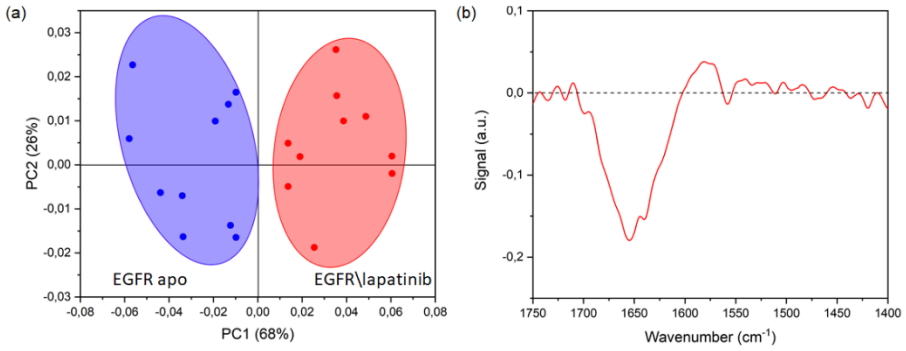

Figure 3.31: (a) Scatter plot of the Principal Component Analysis (PCA) scores. Blue dots represent APO/EGFR measurements; red dots represent lapatinib/EGFR measurements. (b) Profile of the PC1 loading in the spectral region of PCA analysis (Amides region, 1750-1400 cm^{-1}).

PCA shows that Lapatinib/EGFR and EFGR spectra are well separated along PC1 (Figure 3.30a). By considering the profile of the loading 1 (Figure 3.30b), it appears that the separation between the two subgroups is driven by the negative contribution peaked at 1654 cm^{-1}, suggesting a variation of helix structures upon drug treatment.

For a better interpretation of protein spectra changes upon inhibitor treatment, EGFR spectral variations ascribable to Lapatinib interaction have been highlighted by considering IR difference spectra [87]. In particular, the spectrum of APO/EGFR averaged on 8 acquisitions has been subtracted to Lapatinib/EGFR individual spectra acquired on the same ROI-L1 of 8 different chips. Since difference spectra enhance not only the diversity among two spectra but also the spectral noise, a noise correction procedure has been applied before subtraction. Specifically, in order to further reduce the level of spectral noise, spectra have been reconstructed using only the first 4 loadings obtained from PCA, that account for about 95% of the system variance. Absorbance difference spectra are reported in Figure 3.32 and the different colors represent here the diverse chips of the plasmonic device.

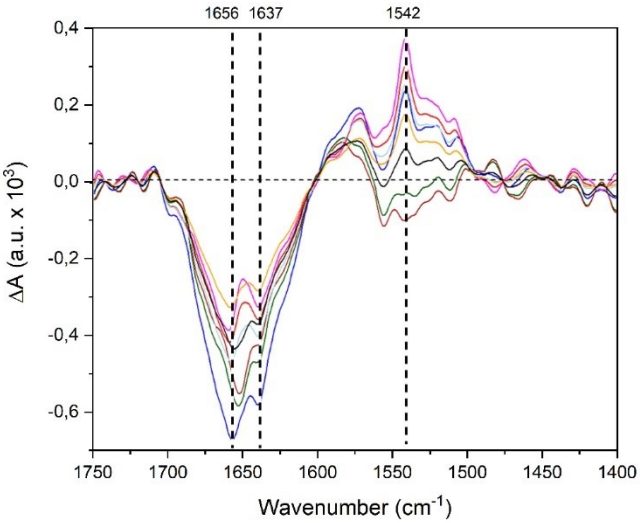

Figure 3.32: Analysis of difference spectra. Each spectrum is the result of the subtraction of the APO/EGFR spectrum, averaged on 8 acquisitions, to eight individual CEIRA Lapatinib/EGFR spectra. Subtraction has been carried out after spectral noise correction. Main differences are highlighted by dashed lines.

As can be seen in the figure, difference spectra present a negative double peak in the region of Amide I, centered at 1656 cm^{-1} (higher peak) and 1637 cm^{-1}. As aforementioned, the 1660-1650 cm^{-1} region is characteristic for Amide I absorption of helical secondary structures, both α helices and 3_{10}-heliches [74], while the peak at 1637 cm^{-1} is characteristic for β sheet structures. Results suggested that in Lapatanib/EGFR there is a decrement of α structures together with a smaller reduction of β structures in comparison to the APO/EGFR form. These results are in accordance with the XRD observation, which highlighted a decrement of 2% of alpha structures, together with a reduction of 1% of beta structures. The absorbance intensity of the spectra variations is in the order of 10^{-3} a.u., which is at less than 5% of the proteins signal. In addition, a positive peak in the Amide II band region was detectable for some chips, centered at 1542 cm^{-1}. As we have seen in the introduction, the Amide II is less sensitive to secondary structures, and variations in its position or components are more related to amino acid side chains contributions. However, shifts of the Amide II band to 1542 cm^{-1} position from native position at 1545 cm^{-1} have been associated in some cases to a damage of the native structure of proteins [88]. By looking at the second derivative analyses reported in Figure 3.29 a slight shift of the main peak of Amide II band is visible. However, such shift (2 cm^{-1}) is below the spectral resolution of our measurements (4 cm^{-1}) and thus was considered it not significate.

Parallel ultraviolet circular dichroism (UV-CD) was performed on the same protein solutions before and after lapatinib interaction. Results are reported in Figure 3.33.

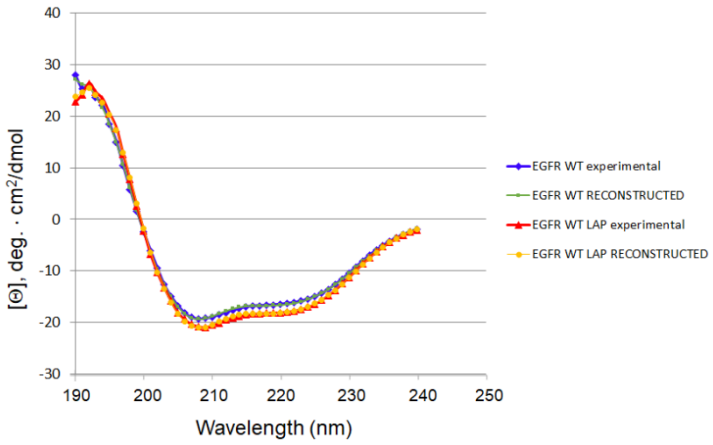

Figure 3.33: Circular dichroism (CD) spectra in the UV range of apo EGFR *wt* (blue line) and EGFR bounded with lapatinib (red line) in HEPES buffer at pH 8.2. Protein solutions were 0,1 mg/mL concentrated. Green and orange lines ware reconstructed spectra generated by Contin/LL software [87] in order to estimate the secondary structure of proteins.

Circular dichroism (CD) is an excellent tool for rapid determination of folding properties of proteins that have been obtained using recombinant techniques or purified from tissues [89]. The most widely used application of protein CD is to determine whether an expressed, purified protein is folded, or if a mutation affects its conformation or stability [90]–[92]. Our acquired CD spectra confirmed that, in both cases, the folding of proteins was preserved (Figure 3.33). As can be seen in Figure 3.33, we could appreciate only minimal differences in the position of the minima of the CD spectra for the two conditions. In order to understand if such differences are significant, experimental curves have been fitted by using Contin/LL software [91] for estimating the secondary structure of proteins in the two different conditions (see over imposed green and orange lines in Figure 3.33). The fitting confirmed that CD could not detect any conformational variation of EGFR after lapatinib interaction (data not shown), despite the evidences provided by X-ray data, highlighting the UV-CD is a technique not enough sensitive for revealing such subtle secondary structure changes.

With our spectral observations, we demonstrated that by using our CEIRA chips is possible to predict the conformational changes induced by inhibitor-protein interactions, in a faster manner employing very low quantities of material with a conformational sensitivity higher than that of conventional circular dichroism. These results demonstrate the extreme conformational sensibility of our device.

Conclusions

With these last experiments, we demonstrated that by using PIR-CEIRA chips is possible to follow the conformational changes induced by inhibitor-protein interactions, in a fast manner and employing very low quantities of material. System allowed a conformational sensitivity higher than that of conventional circular dichroism, also by using conventional FTIR source.

References

[1] F. Neubrech, A. Pucci, T. W. Cornelius, S. Karim, A. García-Etxarri, and J. Aizpurua, "Resonant Plasmonic and Vibrational Coupling in a Tailored Nanoantenna for Infrared Detection," *Phys. Rev. Lett.*, vol. 101, no. 15, p. 157403, Oct. 2008.

[2] F. Neubrech *et al.*, "Infrared Optical Properties of Nanoantenna Dimers with Photochemically Narrowed Gaps in the 5 nm Regime," *ACS Nano*, vol. 6, no. 8, pp. 7326–7332, Aug. 2012.

[3] J. Vogt, C. Huck, F. Neubrech, A. Toma, D. Gerbert, and A. Pucci, "Impact of the plasmonic near- and far-field resonance-energy shift on the enhancement of infrared vibrational signals," *Phys. Chem. Chem. Phys.*, vol. 17, no. 33, pp. 21169–21175, Aug. 2015.

[4] R. Adato, S. Aksu, and H. Altug, "Engineering mid-infrared nanoantennas for surface enhanced infrared absorption spectroscopy," *Mater. Today*, vol. 18, no. 8, pp. 436–446, Oct. 2015.

[5] D. L. Allara, "Handbook of Infrared Spectroscopy of Ultrathin Films By Valeri P. Tolstoy (St. Petersburg State University, Russia), Irina V. Chernyshova (St. Petersburg State Polytechnical University, Russia), and Valeri A. Skryshevsky (Kyiv National Taras Shevchenko University, Ukraine). John Wiley and Sons, Inc.: Hoboken. 2003. xxvi + 710 pp. $250.00. ISBN 0-471-35404-X.," *J. Am. Chem. Soc.*, vol. 126, no. 47, pp. 15633–15634, Dec. 2004.

[6] M. Dufva and C. B. V. Christensen, "Diagnostic and analytical applications of protein microarrays," *Expert Rev. Proteomics*, vol. 2, no. 1, pp. 41–48, Jan. 2005.

[7] G. C. Terstappen, C. Schlüpen, R. Raggiaschi, and G. Gaviraghi, "Target deconvolution strategies in drug discovery," *Nat. Rev. Drug Discov.*, vol. 6, no. 11, pp. 891–903, Nov. 2007.

[8] Z. Tong *et al.*, "Application of biomaterials to advance induced pluripotent stem cell research and therapy," *EMBO J.*, vol. 34, no. 8, pp. 987–1008, Apr. 2015.

[9] K. Manoli *et al.*, "Printable Bioelectronics To Investigate Functional Biological Interfaces," *Angew. Chem. Int. Ed.*, vol. 54, no. 43, pp. 12562–12576, Oct. 2015.

[10] L. S. Penn and H. Wang, "Chemical modification of polymer surfaces: a review," *Polym. Adv. Technol.*, vol. 5, no. 12, pp. 809–817, Dec. 1994.

[11] C. Haensch, M. Chiper, C. Ulbricht, A. Winter, S. Hoeppener, and U. S. Schubert, "Reversible Supramolecular Functionalization of Surfaces: Terpyridine Ligands as Versatile Building Blocks for Noncovalent Architectures," *Langmuir*, vol. 24, no. 22, pp. 12981–12985, Nov. 2008.

[12] M. L. Sham, J. Li, P. C. Ma, and J.-K. Kim, "Cleaning and Functionalization of Polymer Surfaces and Nanoscale Carbon Fillers by UV/Ozone Treatment: A Review," *J. Compos. Mater.*, vol. 43, no. 14, pp. 1537–1564, Jul. 2009.

[13] M. M. M. Bilek *et al.*, "Free radical functionalization of surfaces to prevent adverse responses to biomedical devices," *Proc. Natl. Acad. Sci.*, vol. 108, no. 35, pp. 14405–14410, Aug. 2011.

[14] S. Choi and G. Birarda, "Protein Mixture Segregation at Coffee-Ring: Real-Time Imaging of Protein Ring Precipitation by FTIR Spectromicroscopy," *J. Phys. Chem. B*, vol. 121, no. 30, pp. 7359–7365, Aug. 2017.

[15] M. Thommes and K. A. Cychosz, "Physical adsorption characterization of nanoporous materials: progress and challenges," *Adsorption*, vol. 20, no. 2, pp. 233–250, Feb. 2014.

[16] M. E. Marques, A. A. P. Mansur, and H. S. Mansur, "Chemical functionalization of surfaces for building three-dimensional engineered biosensors," *Appl. Surf. Sci.*, vol. 275, pp. 347–360, Jun. 2013.

[17] E. Pensa *et al.*, "The Chemistry of the Sulfur–Gold Interface: In Search of a Unified Model," *Acc. Chem. Res.*, vol. 45, no. 8, pp. 1183–1192, Aug. 2012.

[18] S. Lupi *et al.*, "Performance of SISSI, the infrared beamline of the ELETTRA storage ring," *JOSA B*, vol. 24, no. 4, pp. 959–964, Apr. 2007.

[19] T. L. McMEEKIN, M. L. GROVES, and N. J. HIPP, "Refractive Indices of Amino Acids, Proteins, and Related Substances," in *Amino Acids and Serum Proteins*, vol. 44, 0 vols., AMERICAN CHEMICAL SOCIETY, 1964, pp. 54–66.

[20] M. Eslamian and F. Zabihi, “Ultrasonic Substrate Vibration-Assisted Drop Casting (SVADC) for the Fabrication of Photovoltaic Solar Cell Arrays and Thin-Film Devices,” *Nanoscale Res. Lett.*, vol. 10, no. 1, p. 462, Dec. 2015.
[21] L. S. Taylor and G. Zografi, “Spectroscopic Characterization of Interactions Between PVP and Indomethacin in Amorphous Molecular Dispersions,” *Pharm. Res.*, vol. 14, no. 12, pp. 1691–1698, Dec. 1997.
[22] M. Stefani, “Protein Folding and Misfolding on Surfaces,” *Int. J. Mol. Sci.*, vol. 9, no. 12, pp. 2515–2542, Dec. 2008.
[23] D. Ami *et al.*, “In situ characterization of protein aggregates in human tissues affected by light chain amyloidosis: a FTIR microspectroscopy study,” *Sci. Rep.*, vol. 6, Jul. 2016.
[24] D. Enders and A. Pucci, “Surface enhanced infrared absorption of octadecanethiol on wet-chemically prepared Au nanoparticle films,” *Appl. Phys. Lett.*, vol. 88, no. 18, p. 184104, May 2006.
[25] M. J. E. Fischer, “Amine Coupling Through EDC/NHS: A Practical Approach,” in *Surface Plasmon Resonance: Methods and Protocols*, N. J. Mol and M. J. E. Fischer, Eds. Totowa, NJ: Humana Press, 2010, pp. 55–73.
[26] A. K. Wright and M. R. Thompson, “Hydrodynamic structure of bovine serum albumin determined by transient electric birefringence.,” *Biophys. J.*, vol. 15, no. 2 Pt 1, pp. 137–141, Feb. 1975.
[27] C. Petibois, G. Déléris, M. Piccinini, M. Cestelli-Guidi, and A. Marcelli, “A bright future for synchrotron imaging,” *Nat. Photonics*, vol. 3, no. 4, p. 179, Apr. 2009.
[28] Y. Yao, A. J. Hoffman, and C. F. Gmachl, “Mid-infrared quantum cascade lasers,” *Nat. Photonics*, vol. 6, no. 7, pp. 432–439, Jul. 2012.
[29] L. M. Miller and P. Dumas, “Chemical imaging of biological tissue with synchrotron infrared light,” *Biochim. Biophys. Acta BBA - Biomembr.*, vol. 1758, no. 7, pp. 846–857, Jul. 2006.
[30] D. Etezadi, J. B. Warner Iv, F. S. Ruggeri, G. Dietler, H. A. Lashuel, and H. Altug, “Nanoplasmonic mid-infrared biosensor for *in vitro* protein secondary structure detection,” *Light Sci. Appl.*, vol. 6, no. 8, p. e17029, Aug. 2017.
[31] K. A. Majorek *et al.*, “Structural and immunologic characterization of bovine, horse, and rabbit serum albumins,” *Mol. Immunol.*, vol. 52, no. 3, pp. 174–182, Oct. 2012.
[32] K. D. Hardman and C. F. Ainsworth, “Structure of concanavalin A at 2.4-Ang resolution,” *Biochemistry (Mosc.)*, vol. 11, no. 26, pp. 4910–4919, Dec. 1972.
[33] A. Dong, P. Huang, and W. S. Caughey, “Protein secondary structures in water from second-derivative amide I infrared spectra,” *Biochemistry (Mosc.)*, vol. 29, no. 13, pp. 3303–3308, Apr. 1990.
[34] G. Zandomeneghi, M. R. H. Krebs, M. G. McCammon, and M. Fändrich, “FTIR reveals structural differences between native β-sheet proteins and amyloid fibrils,” *Protein Sci. Publ. Protein Soc.*, vol. 13, no. 12, pp. 3314–3321, Dec. 2004.
[35] F. Piccirilli, G. Schirò, V. Vetri, S. Lupi, A. Perucchi, and V. Militello, “Decoding vibrational states of Concanavalin A amyloid fibrils,” *Biophys. Chem.*, vol. 199, pp. 17–24, Apr. 2015.
[36] A. Barth, “Infrared spectroscopy of proteins,” *Biochim. Biophys. Acta BBA - Bioenerg.*, vol. 1767, no. 9, pp. 1073–1101, Sep. 2007.
[37] E. Goormaghtigh, J.-M. Ruysschaert, and V. Raussens, “Evaluation of the Information Content in Infrared Spectra for Protein Secondary Structure Determination,” *Biophys. J.*, vol. 90, no. 8, pp. 2946–2957, Apr. 2006.
[38] D. M. Byler and H. Susi, “Examination of the secondary structure of proteins by deconvolved FTIR spectra,” *Biopolymers*, vol. 25, no. 3, pp. 469–487, Mar. 1986.
[39] R. J. Ellis and F. U. Hartl, “Principles of protein folding in the cellular environment,” *Curr. Opin. Struct. Biol.*, vol. 9, no. 1, pp. 102–110, Feb. 1999.
[40] K. Rahmelow and W. Hubner, “Infrared Spectroscopy in Aqueous Solution: Difficulties and Accuracy of Water Subtraction,” *Appl. Spectrosc.*, vol. 51, no. 2, pp. 160–170, Feb. 1997.
[41] A. Matruglio *et al.*, “Graphene liquid cells for multi-technique analysis of biological cells in water environment,” *J. Instrum.*, vol. 13, no. 5, p. C05016, 2018.

[42] S. G. Kazarian, K. L. A. Chan, and F. H. Tay, “ATR-FT-IR Imaging for Pharmaceutical and Polymeric Materials: From Micro to Macro Approaches,” in *Infrared and Raman Spectroscopic Imaging*, John Wiley & Sons, Ltd, 2009, pp. 347–375.
[43] R. Adato and H. Altug, “*In-situ* ultra-sensitive infrared absorption spectroscopy of biomolecule interactions in real time with plasmonic nanoantennas,” *Nat. Commun.*, vol. 4, p. 2154, Jul. 2013.
[44] O. Limaj *et al.*, “Infrared Plasmonic Biosensor for Real-Time and Label-Free Monitoring of Lipid Membranes,” *Nano Lett.*, vol. 16, no. 2, pp. 1502–1508, Feb. 2016.
[45] N. E. Zhou, C. M. Kay, and R. S. Hodges, “The Role of Interhelical Ionic Interactions in Controlling Protein Folding and Stability: De Novo Designed Synthetic Two-stranded α-Helical Coiled-Coils,” *J. Mol. Biol.*, vol. 237, no. 4, pp. 500–512, Apr. 1994.
[46] K. Fu, A. M. Klibanov, and R. Langer, “Protein stability in controlled-release systems,” *Nat. Biotechnol.*, vol. 18, no. 1, pp. 24–25, Jan. 2000.
[47] C. Wang *et al.*, “Improving Protein Stability and Controlling Protein Release by Adding Poly (Cyclohexane-1, 4-Diyl Acetone Dimethylene Ketal) to PLGA Microspheres,” *Curr. Drug Deliv.*, vol. 12, no. 6, pp. 726–735, 2015.
[48] N. Normanno *et al.*, “Epidermal growth factor receptor (EGFR) signaling in cancer,” *Gene*, vol. 366, no. 1, pp. 2–16, Jan. 2006.
[49] Y. Shan, A. Arkhipov, E. T. Kim, A. C. Pan, and D. E. Shaw, “Transitions to catalytically inactive conformations in EGFR kinase,” *Proc. Natl. Acad. Sci. U. S. A.*, vol. 110, no. 18, pp. 7270–7275, Apr. 2013.
[50] B. Zou, V. H. F. Lee, L. Chen, L. Ma, D. D. Wang, and H. Yan, “Deciphering mechanisms of acquired T790M mutation after EGFR inhibitors for NSCLC by computational simulations,” *Sci. Rep.*, vol. 7, no. 1, p. 6595, Jul. 2017.
[51] R. L. Siegel, K. D. Miller, and A. Jemal, “Cancer statistics, 2018,” *CA. Cancer J. Clin.*, vol. 68, no. 1, pp. 7–30, Jan. 2018.
[52] M. G. Oser, M. J. Niederst, L. V. Sequist, and J. A. Engelman, “Transformation from non-small-cell lung cancer to small-cell lung cancer: molecular drivers and cells of origin,” *Lancet Oncol.*, vol. 16, no. 4, pp. e165-172, Apr. 2015.
[53] H. A. Burris, “Dual kinase inhibition in the treatment of breast cancer: initial experience with the EGFR/ErbB-2 inhibitor lapatinib,” *The Oncologist*, vol. 9 Suppl 3, pp. 10–15, 2004.
[54] L. V. Sequist, “First-Generation Epidermal Growth Factor Receptor Tyrosine Kinase Inhibitors in EGFR Mutation; Positive Non-small Cell Lung Cancer Patients,” *J. Thorac. Oncol.*, vol. 3, no. 6, Supplement 2, pp. S143–S145, Jun. 2008.
[55] J. H. Park, Y. Liu, M. A. Lemmon, and R. Radhakrishnan, “Erlotinib binds both inactive and active conformations of the EGFR tyrosine kinase domain,” *Biochem. J.*, vol. 448, no. Pt 3, pp. 417–423, Dec. 2012.
[56] Y. Jia *et al.*, “Overcoming EGFR(T790M) and EGFR(C797S) resistance with mutant-selective allosteric inhibitors,” *Nature*, vol. 534, no. 7605, pp. 129–132, Jun. 2016.
[57] K. Politi, D. Ayeni, and T. Lynch, “The Next Wave of EGFR Tyrosine Kinase Inhibitors Enter the Clinic,” *Cancer Cell*, vol. 27, no. 6, pp. 751–753, Jun. 2015.
[58] C.-H. Yun *et al.*, “The T790M mutation in EGFR kinase causes drug resistance by increasing the affinity for ATP,” *Proc. Natl. Acad. Sci.*, vol. 105, no. 6, pp. 2070–2075, Feb. 2008.
[59] I. Dagogo-Jack, J. A. Engelman, and A. T. Shaw, “Overcoming On-Target Resistance to Tyrosine Kinase Inhibitors in Lung Cancer,” *Annu. Rev. Cancer Biol.*, vol. 1, no. 1, pp. 257–274, 2017.
[60] L.-L. Kong *et al.*, “Structural pharmacological studies on EGFR T790M/C797S,” *Biochem. Biophys. Res. Commun.*, vol. 488, no. 2, pp. 266–272, Jun. 2017.
[61] J. Stamos, M. X. Sliwkowski, and C. Eigenbrot, “Structure of the Epidermal Growth Factor Receptor Kinase Domain Alone and in Complex with a 4-Anilinoquinazoline Inhibitor,” *J. Biol. Chem.*, vol. 277, no. 48, pp. 46265–46272, Nov. 2002.
[62] A. Levitzki and E. Mishani, “Tyrphostins and Other Tyrosine Kinase Inhibitors,” *Annu. Rev. Biochem.*, vol. 75, no. 1, pp. 93–109, 2006.

[63] X. Zhang, J. Gureasko, K. Shen, P. A. Cole, and J. Kuriyan, “An allosteric mechanism for activation of the kinase domain of epidermal growth factor receptor,” *Cell*, vol. 125, no. 6, pp. 1137–1149, Jun. 2006.
[64] W. Kabsch and C. Sander, “Dictionary of protein secondary structure: pattern recognition of hydrogen-bonded and geometrical features,” *Biopolymers*, vol. 22, no. 12, pp. 2577–2637, Dec. 1983.
[65] E. Mitri *et al.*, “Time-resolved FT-IR microspectroscopy of protein aggregation induced by heat-shock in live cells,” *Anal. Chem.*, vol. 87, no. 7, pp. 3670–3677, Apr. 2015.
[66] J. Schmitt, H. Hess, and H. G. Stunnenberg, “Affinity purification of histidine-tagged proteins,” *Mol. Biol. Rep.*, vol. 18, no. 3, pp. 223–230, Oct. 1993.
[67] J. E. Tropea, S. Cherry, and D. S. Waugh, “Expression and purification of soluble His(6)-tagged TEV protease,” *Methods Mol. Biol. Clifton NJ*, vol. 498, pp. 297–307, 2009.
[68] H. Block *et al.*, “Chapter 27 Immobilized-Metal Affinity Chromatography (IMAC): A Review,” in *Methods in Enzymology*, vol. 463, R. R. Burgess and M. P. Deutscher, Eds. Academic Press, 2009, pp. 439–473.
[69] P. D. Gershon and S. Khilko, “Stable chelating linkage for reversible immobilization of oligohistidine tagged proteins in the BIAcore surface plasmon resonance detector,” *J. Immunol. Methods*, vol. 183, no. 1, pp. 65–76, Jun. 1995.
[70] L. Nieba *et al.*, “BIACORE analysis of histidine-tagged proteins using a chelating NTA sensor chip,” *Anal. Biochem.*, vol. 252, no. 2, pp. 217–228, Oct. 1997.
[71] S. G. Patching, “Surface plasmon resonance spectroscopy for characterisation of membrane protein–ligand interactions and its potential for drug discovery,” *Biochim. Biophys. Acta BBA - Biomembr.*, vol. 1838, no. 1, Part A, pp. 43–55, Jan. 2014.
[72] D. Wasserberg *et al.*, “Controlling Protein Surface Orientation by Strategic Placement of Oligo-Histidine Tags,” *ACS Nano*, vol. 11, no. 9, pp. 9068–9083, Sep. 2017.
[73] E. Kauffmann, N. C. Darnton, R. H. Austin, C. Batt, and K. Gerwert, “Lifetimes of intermediates in the β-sheet to α-helix transition of β-lactoglobulin by using a diffusional IR mixer,” *Proc. Natl. Acad. Sci.*, vol. 98, no. 12, pp. 6646–6649, Jun. 2001.
[74] E. Goormaghtigh, V. Cabiaux, and J. M. Ruysschaert, “Determination of soluble and membrane protein structure by Fourier transform infrared spectroscopy. III. Secondary structures,” *Subcell. Biochem.*, vol. 23, pp. 405–450, 1994.
[75] M. A. Williams, “Protein–Ligand Interactions: Fundamentals,” in *Protein-Ligand Interactions: Methods and Applications*, M. A. Williams and T. Daviter, Eds. Totowa, NJ: Humana Press, 2013, pp. 3–34.
[76] R. A. Copeland, “Conformational adaptation in drug-target interactions and residence time,” *Future Med. Chem.*, vol. 3, no. 12, pp. 1491–1501, Sep. 2011.
[77] G. Langer, S. X. Cohen, V. S. Lamzin, and A. Perrakis, “Automated macromolecular model building for X-ray crystallography using ARP/wARP version 7,” *Nat. Protoc.*, vol. 3, no. 7, pp. 1171–1179, Jul. 2008.
[78] F. Castellani, B. van Rossum, A. Diehl, M. Schubert, K. Rehbein, and H. Oschkinat, “Structure of a protein determined by solid-state magic-angle-spinning NMR spectroscopy,” *Nature*, vol. 420, no. 6911, pp. 98–102, Nov. 2002.
[79] J.-P. Renaud *et al.*, “Biophysics in drug discovery: impact, challenges and opportunities,” *Nat. Rev. Drug Discov.*, vol. 15, no. 10, pp. 679–698, 2016.
[80] J. Güldenhaupt *et al.*, “Ligand-Induced Conformational Changes in HSP90 Monitored Time Resolved and Label Free—Towards a Conformational Activity Screening for Drug Discovery,” *Angew. Chem. Int. Ed.*, vol. 57, no. 31, pp. 9955–9960, Jul. 2018.
[81] G. M. Higa and J. Abraham, “Lapatinib in the treatment of breast cancer,” *Expert Rev. Anticancer Ther.*, vol. 7, no. 9, pp. 1183–1192, Sep. 2007.
[82] M. Murtaza *et al.*, “Non-invasive analysis of acquired resistance to cancer therapy by sequencing of plasma DNA,” *Nature*, vol. 497, no. 7447, pp. 108–112, May 2013.
[83] M. Pereira, C. S. Verma, and G. Fuentes, “Differences in the Binding Affinities of ErbB Family: Heterogeneity in the Prediction of Resistance Mutants,” *PLOS ONE*, vol. 8, no. 10, p. e77054, ott 2013.

[84] E. R. Wood *et al.*, “A unique structure for epidermal growth factor receptor bound to GW572016 (Lapatinib): relationships among protein conformation, inhibitor off-rate, and receptor activity in tumor cells,” *Cancer Res.*, vol. 64, no. 18, pp. 6652–6659, Sep. 2004.
[85] C. Zscherp and J. Heberle, “Infrared Difference Spectra of the Intermediates L, M, N, and O of the Bacteriorhodopsin Photoreaction Obtained by Time-Resolved Attenuated Total Reflection Spectroscopy,” *J. Phys. Chem. B*, vol. 101, no. 49, pp. 10542–10547, Dec. 1997.
[86] R. C. Team, “R: A language and environment for statistical computing,” 2013.
[87] K. J. Rothschild, “FTIR difference spectroscopy of bacteriorhodopsin: Toward a molecular model,” *J. Bioenerg. Biomembr.*, vol. 24, no. 2, pp. 147–167, Apr. 1992.
[88] M. KYRIAKIDOU, J. ANASTASSOPOULOU, A. TSAKIRIS, M. KOUI, and T. THEOPHANIDES, “FT-IR Spectroscopy Study in Early Diagnosis of Skin Cancer,” *In Vivo*, vol. 31, no. 6, pp. 1131–1127, Nov. 2017.
[89] N. J. Greenfield, “Using circular dichroism spectra to estimate protein secondary structure,” *Nat. Protoc.*, vol. 1, no. 6, pp. 2876–2890, 2006.
[90] Y.-H. Chen, J. T. Yang, and H. M. Martinez, “Determination of the secondary structures of proteins by circular dichroism and optical rotatory dispersion,” *Biochemistry (Mosc.)*, vol. 11, no. 22, pp. 4120–4131, Oct. 1972.
[91] N. Sreerama and R. W. Woody, “Estimation of Protein Secondary Structure from Circular Dichroism Spectra: Comparison of CONTIN, SELCON, and CDSSTR Methods with an Expanded Reference Set,” *Anal. Biochem.*, vol. 287, no. 2, pp. 252–260, Dec. 2000.
[92] L. Whitmore and B. A. Wallace, “DICHROWEB, an online server for protein secondary structure analyses from circular dichroism spectroscopic data,” *Nucleic Acids Res.*, vol. 32, no. suppl_2, pp. W668–W673, Jul. 2004.

Conclusions

During the PhD, I exploited the Collective Enhanced Infrared Absorption (CEIRA) approach for investigating the secondary structure of proteins of biological and biomedical relevance in aqueous environment and at nanomolar concentrations. Thanks to the collaboration with the Istituto Italiano di Tecnologia (IIT) in Genova under the supervision of Andrea Toma, the fabrication of CEIRA devices has been optimized using EBL, by tuning the resonance response provided by ordered nanoantenna arrays to the most relevant protein vibrational bands in the Mid-IR regime, especially to Amide I band. The first device generation was made by nanoantennas arrays composed by neighboring nanoantennas disposed tip-to-tip and separated by a nanometer gap. This array geometry was defined *line-to-line*. As known from the literature and experimentally verified, the efficiency of such arrays is strongly dependent from the direction of light polarization. Therefore, a second array geometry was fabricated, composed by nanoantennas arranged in centrosymmetric crosses, for this reason called *cross*-geometry. We demonstrated that, this second device geometry did not require polarized light for providing optimal responses, giving an improved signal enhancement in comparison to former *line-to-line* devices.

Due to the limited penetration depth of the enhanced EM field, up to some tens of nm from the nanoantenna surface, we improved methodologies for the effective anchoring of proteins onto gold-nanoantennas. Among the different protocols tested, we selected the *amine coupling* since it guaranteed the formation of a thin and homogenous protein layer. The capability of our CEIRA device to probe secondary structure details of protein monolayers was verified by investigating two model proteins with a distinct secondary structure (BSA, mainly α helix folded and ConA, mainly β sheet folded) in dried conditions. Results demonstrated that IR-SR CEIRA microscopy provides repeatable and reliable insight of protein conformation on dried systems, with attomole sensitivity.

Progressing with the characterization of the same model proteins in aqueous environment, we overcame the issue of water layer confinement by implementing the Plasmonic Internal Reflection (PIR) approach, first proposed by Adato and co-workers, based on an ATR-like geometry of acquisition. The obtained results demonstrated that the PIR-CEIRA approach

allows fast, label-free and reliable measurements of protein in buffer solutions with very high sensitivity.

Thus, we exploited the gained experience to assess the subtle secondary structure modifications undergone by Epidermal Growth Factor Receptor, EGFR interacting with tyrosine-kinase inhibitor (TKI), Lapatinib. However, moving from model proteins to proteins of biomedical interest was not straightforward: first results showed that reliable conformational measurements of EGFR were mostly hindered by the adopted immobilization procedure, too harsh for efficiently anchoring EGFR. Therefore, a new protocol for the immobilization of sensitive proteins was implemented, based onto the NTA/His-tag interaction. The experimental results demonstrated that the His-tag immobilization protocol is a powerful method for protein anchoring, which allows to preserve the protein conformational status and to perform *in situ* and *operando* protein conformational analysis. Of course, this immobilization method is limited to those proteins that are produced by recombinant technology, that for this thesis has been provided by the Protein Production Facility at Elettra. However, such proteins represent the main targets of pharmaceutical research. Thanks to this new immobilization protocol, it was possible to foresee the exploitation of CEIRA-PIR microscopy for dynamically following the interaction *in situ* of EGFR with its specific inhibitor (Lapatinib). With these last set of experiments, we demonstrated that by using PIR-CEIRA chips is possible to follow the conformational changes induced by inhibitor-protein interactions with a sensitivity higher than that provided by CD experiments, in a fast manner, employing ultra-low quantities of material, also by using conventional FTIR source.

It is noteworthy to mention that most of the scientific works cited in this thesis and focused on the exploitation of resonant SEIRA for protein analysis have been done using model proteins and they all envisage the application of the technique for the investigation of critical process such as the protein-ligand interaction. At the best of our knowledge, this capability has been proven for the first time in this PhD thesis, and a scientific paper on the thesis outcomes is in preparation.

In conclusion, our CEIRA devices are a versatile tool for further understanding of the fundamental principles underlying the mechanisms on the base of protein-ligand interaction, and in general for monitoring protein secondary structure modifications. A platform with this kind of sensitivity could open new possibilities in the study of proteins of biomedical interest, in particular for drug-screening purposes, setting as a valuable complement to more

consolidated, but also more time and material consuming, methods such as XRD, SAXS, NMR and cryo-EM.

Certainly, nowadays, the widespread of the technology in biological laboratories is mostly limited by the costly and time-consuming approach for CEIRA device fabrication, based on EBL. The cost of the devices could be largely scaled down by using cheaper nanofabrication techniques for nanoantenna arrays fabrication, and some scientific papers have recently published on the topic. Going down that route, larger areas could be patterned, allowing transferring the technology from FTIR microscopy to FTIR spectroscopy, therefore dramatically reducing the costs of the experimental apparatus. Furthermore, despite the advantages offered by IR-SR, conventional sources are also applicable, as proven in this PhD thesis. New generation IR sources, such as quantum cascade lasers, QCLs, offer nowadays a valuable alternative to conventional IR sources and high potential for miniaturization. Despite emitting in narrow spectral intervals, it is possible to conceive the integration of laser, CEIRA sample unit, and IR detector on a chip, where the plasmonic substrate design is tuned to the laser emission for specific applications. Indeed, despite in the manuscript we concentrated more on the nanoantenna arrays tuned to Amide I–II spectral range, arrays suitable for other spectral region have been fabricated and they could be exploited for different applications, encompassing DNA/RNA sensing, monitoring of formation kinetics of planar biomimetic membranes, and many others.

Acknowledgments

The work presented in this thesis would not have been possible without the huge and constant aid and support of my supervisors, Dr. Lisa Vaccari and Dr. Giovanni Birarda. Particularly, I want to thank Lisa for the opportunity that she gave me in continuing the work after the period of the Master's Degree, and Giovanni for the endless support during these years. I would like to thank also my academic tutor, Professor Alessandro Baraldi.

As mentioned in the thesis, different groups collaborated with us during the PhD period. In particular, special thanks go to Andrea Toma and Francesco De Angelis and all the teams of their laboratories for the support and the access to IIT laboratories in Genova for nanostructures fabrications. I personally thank Angelica Carrara and Andrea Cerea for the help in the fabrication of CEIRA devices. We thank Hans Bechtel of Advanced Light Sources (ALS) Infrared Beamlines at Berkley for the synchrotron radiation s-SNIM measurements on *cross*-devices and Sergiu Amarie of Neaspec GmbH (Martinsried, Bayern) for s-SNIM measurements on *line-to-line*-devices and functionalized gold gratings. We thank Pietro Parisse of Nanoinnovation Laboratory at Elettra for the support for functionalization protocols. We thank Fabio Perissinotto of Nanoinnovation Laboratory at Elettra for the AFM measurements of samples and Loredana Casalis for the access to the laboratories. We thank Alessando Gambitta for the fabrication of static and fluidic devices with 3D printer. We thank Marta Semrau and Paola Storici of structural biology laboratory at Elettra for the production of the kinase domain of Epidermal Growth Factor Receptor (EGFR) in its different forms.

Last, but not least, I would like to thank all the team of SISSI (both branches), for the wonderful time experienced at the beamline: a big hug goes to Alessia Chiara and Nicola, for the talks in the office, to Paola, Andrea, Nidhi, Artur and Federica for the lunches and laugh at the canteen. Thank you to all the people who I met at Elettra in these years, you are too many for being mentioned here, but I will always remember you all.

This thesis is dedicated to my family, who always supports me, and in particular to my wife Elisa, who stays always by my side.

Appendix A

In this section I present some additional activities in bio-spectroscopy carried out during my PhD. Working in a large-scale infrastructure, gave me the possibility to follow different projects and to have access to different beamlines. In particular, in collaboration with the Inelastic Ultraviolet Scattering (IUVS) beamline [1]. We carried out a series of experiments devoted to the study of Nucleic Acids (both DNA and RNA) by means of FTIR and Ultraviolet Resonant Raman (UV-RR). This project was based onto my previous experience at SISSI beamline during the internship for my Master thesis in Medical Biotechnology, which was focused onto the characterization of RNA contribution to the vibrational spectra of mammalian cell. Results of the thesis work have been summarized in a paper. [2] In the implementation of the Master Thesis activities, we faced the challenge to measure, by vibrational spectroscopies, extracted genetic material from cells. These issues were mainly related to the contamination of the extracted solutions by residues of the purification passages performed with commercial kits. During the PhD, I decided to further exploit my biotechnology background to develop tailored protocols for extracting nucleic acids from cell cultures with a higher degree of purity, more compatible with vibrational spectroscopy studies. Such protocols were developed inside the project of the Master thesis in Material Science and Engineering of Katia Latella, a master student of "Università di Genova", of which I was the co-supervisor. Developed protocols have been published in a recent paper [3]. The main results presented in the aforementioned paper are summarized in this appendix.

After a brief introduction onto the nucleic acids (*section 4.1*), the tailored protocols developed for purifying nucleic acids solutions extracted from cells for reliable vibrational spectroscopy measurements are presented in *section 4.2*. A brief conclusion is provided in *section 4.3*.

4.1 Vibrational spectroscopies for nucleic acids research

Nucleic acids are biopolymers essential for all known forms of life, since they encode, transmit and express genetic information [4]. They are classified into two main categories: deoxyribonucleic acids (DNA) and ribonucleic acids (RNA). These macromolecules consist of a large number of linked nucleotides, each composed of a pentose sugar (2'-deoxyribose for DNA and ribose for RNA), a phosphate group, and a nitrogenous base (adenine, cytosine,

guanine and thymine in DNA while uracil substitutes thymine in RNA). The sugar-phosphate alternation through a phosphodiester linkage originates the nucleic acid backbone. In simple terms, DNA and RNA differ for the arrangement since in general DNA molecules are double-stranded while RNA molecules are mainly single-stranded [4] (see Figure 4.1). DNA serves to transmit and encode the genetic information, while the principal role of RNA resides in its expression.

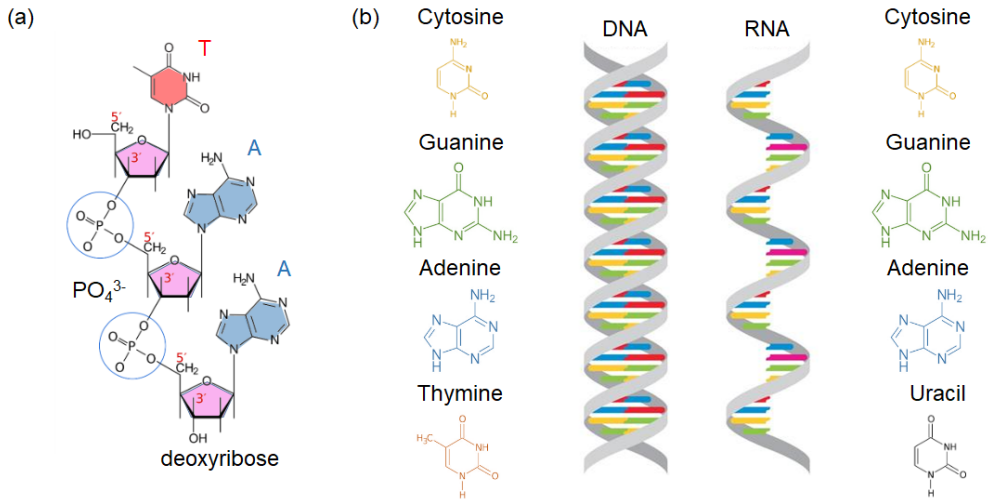

Figure 4.1: (a) Sugar-phosphate backbone of nucleic acids. In the figure the case of deoxyribonucleic acid (DNA) is reported. Sugar moieties are purple colored, nitrogenous bases light blue colored and phosphate linking moieties are highlighted by blue circles. Sugar moieties are linked by the PO_4^{3-} molecules in the 5'-3' direction. (b) Main differences between DNA and RNA strands. Bases that composes the two typologies of nucleic acids are reported. (adapted from [5]).

Vibrational spectroscopies are widely used tools for nucleic acids studies [6]–[9]. In particular, FTIR spectroscopy offers many distinct bands indicative of DNA and RNA sugar-phosphate backbone folding [10], allowing to probe nucleic acids structural changes upon environmental modifications, such as thermal denaturation or dehydration [11]. Specifically, the B and A forms of DNA and RNA have distinct spectral contribution of the asymmetric stretching of phosphate band, centered at about 1240 and 1220 cm^{-1} respectively [2]. On the other hand, Raman spectroscopy is more sensitive the N-heteroaromatic nucleotide rings of nucleic acids [12]. In particular, the use of UV excitation instead of visible light allows to obtain a higher scattering efficiency without intense fluorescence backgrounds [8], [13]. Due to their clear complementarity, FTIR spectroscopy and ultraviolet-resonant Raman (UV-RR) scattering are often jointly utilized to investigate nucleic acids [14]. Spectra of DNA obtained by both UV-RR (blue line) and FTIR (red line) spectroscopy are overimposed in

Figure 4.2 in order to highlight the complementarity of the information provided by the two analytical tools.

The determination of the FTIR and UV-Raman spectral features of isolated DNA and RNA strands have been widely studied and documented in literature [10], [11], [15]. However, most of the published works focus on the analysis of simplified systems, such as short chains of nucleic acids or commercial pure genomic DNA samples, (i.e. Salmon or Calf Thymus DNA) provided as lyophilized sodium salts and thus certified as high-quality materials. However, high-quality commercial DNA is limited only to few sources, while the major interest in molecular biology arises from the vibrational analysis of DNA and RNA isolated from real biological samples, such as living or preserved tissues, primary cells, and virus particles relative to the specific lines of research of each laboratory.

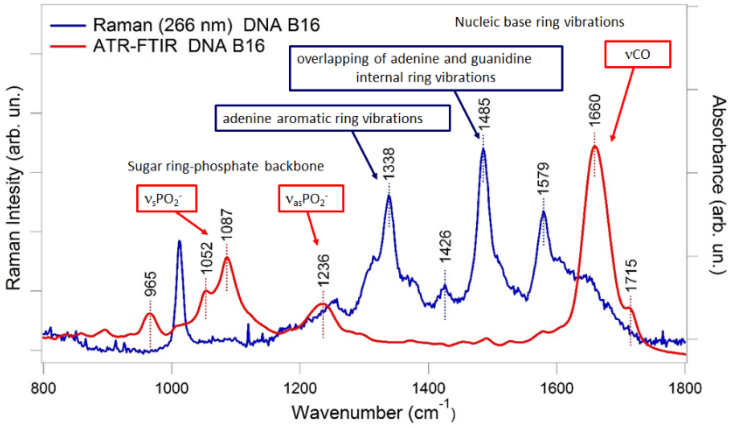

Figure 4.2: The FTIR spectrum of a genomic DNA extracted by B16 cells (red spectrum) obtained by ATR-FTIR spectroscopy is over imposed to a UV-RR spectrum of the same B16 DNA solution. UV-RR spectrum have been acquired with an excitation wavelength set at 266 nm. Image clearly shows the complementarity of the two spectra as a consequence of the different selection rules of the two techniques. In particular, sugar-phosphate backbone information can be obtained by the analyses of the symmetric (ν_s – 1000-1150 cm^{-1}) and asymmetric (ν_{as} – 1180-1280 cm^{-1}) stretching of phosphate moieties of the FTIR spectrum. UV-RR spectrum shows sharp Raman shifts in the spectral region between 1280-1600 cm^{-1}, which are relevant for the contributions of nucleic acids ring vibrations.

Nowadays, several commercial kits are available for obtaining DNA or RNA from different biological sources. However, such commercial extraction kits and UV spectrophotometry assays have been optimized for the removal/detection of molecules that could interfere with the quantification of nucleic acid, their folding, and their activity. These commercial kits make use of several chemicals, some of them patented, that, if not correctly removed, can potentially show up and affect the vibrational spectra of nucleic acid samples, even if classified as high quality from a standard biological perspective. As a consequence, such

chemical residuals can compromise the interpretation of IR and Raman spectra. In particular, in our experiments, DNA and RNA samples of good quality from a biological perspective (certified by means of UV spectrophotometry measures) clearly showed the presence of chemicals interfering with the vibrational measurements, mainly ethanol and guanidinium salts.

During these three years, we developed a fast and cheap strategy for tailoring the extraction protocols in light of the requirements of both vibrational spectroscopies that allowed obtaining nucleic acid samples with ethanol concentration below the detection limit of both techniques and with a significantly reduced spectral interference from guanidinium salts, which was reported in a recent paper which I am the author [2]. A summary of the results is here reported.

4.2 Nucleic acid purification protocols tailored for vibrational spectroscopy

In this work genomic DNA and total RNA solutions were purified starting by different cell cultures of B16 mouse melanoma cell line. Two of the most common commercial kits for nucleic acids purification have been used. In particular DNA was extracted and purified using QIAmp® DNA Blood Mini Kit (Qiagen). RNA was extracted by Isol-RNA lysis reagent (5-Prime), following TRIzol® Plus RNA Purification Kit protocol (Thermo-Fisher Scientific). The general protocols of these two kits are reported in the following and the tailored protocols are reported as well. The optimization of nucleic acids solutions was verified by means of UV-RR and ATR-FTIR spectroscopy. By identifying by vibrational spectroscopy techniques the contaminants moieties in solutions produced by conventional kits, it was possible to selectively implement the passages of protein extraction procedures.

(i) Conventional protocol of QIAmp® DNA Blood Mini Kit (Qiagen)

The QIAamp DNA Mini Kit yields DNA sized up to 50 kilo base (kb), with the major part of the fragments in the range of 20–30 kb. It is a fast procedure and is carried out using QIAamp Mini spin columns in a standard micro centrifuge. The kit provides an optimized buffer (AL buffer) and enzymes (proteinase K) to lyse samples, stabilize nucleic acids, and enhance selective DNA adsorption to the QIAamp membrane. Specifically, 4–5 million of B16 cells were lysed with AL buffer and proteinase K. Ethanol was then added to the lysates, and the suspension was loaded onto a QIAamp spin column. The spin column was incubated at room temperature for 5 min before centrifugation at 8000×g in order to increase the yield. To remove residual contaminants, DNA bound to the QIAamp membrane was washed by two centrifugations (8000 and 12000×g, respectively) with two different washing buffers

(AW1 and AW2). Purified DNA was finally eluted with 60 µl of RNase and DNase free water, used as elution buffer, from the QIAamp Mini spin column. The DNA solution was equilibrated to room temperature (15–25 °C).

(ii) Standard RNA purification by using Isol-RNA lysis reagent (5-Prime)

For RNA purification we followed the TRIzol® Plus RNA Purification Kit protocol (Thermo-Fisher Scientific). Isol-RNA lysis reagent is a ready-to-use monophasic solution of phenol and guanidine thiocyanate, designed to isolate high quality total RNA from cell and tissue samples of human, animal, plant, yeast, or bacterial origin [16]. It maintains the integrity of the RNA due to highly effective inhibition of RNase activity while disrupting cells and dissolving cellular components during sample homogenization.

For our purposes, an entire cell plate was lysed by adding 1 ml of sol-RNA lysis reagent and then homogenized. Hence, 200 µl of chloroform were added, and the homogenate was let to separate into a clear upper aqueous layer (containing RNA, ~700 µl), an interphase, and a red lower organic layer (containing DNA and proteins). RNA was then precipitated from the aqueous layer with 800 µl of isopropanol. The precipitated RNA was washed with ethanol 70% v/v in order to remove impurities and then resuspended in 60 µl of RNase and DNase free water for use in downstream applications. All passages were performed maintaining the temperature at 4 °C to avoid RNA degradation.

(iii) Optimized DNA extraction protocol

The optimized protocol is based on the standard QIAamp DNA Mini Kit with the following modifications: (a) the cultured cells were split into two aliquots (about 3 million of cells each) before lysis, and each aliquot was loaded on an individual column; (b) after the use of the second washing buffer (AW2), columns were spun-dry twice at full speed in order to remove all the liquid part from the silica membrane; and (c) before the final elution, a pre-elution step with 2–3 µl of elution buffer was done, the column was centrifuged (8000×g), and the eluted material was discarded.

(iv) Optimized RNA extraction protocol

The optimized protocol relies on the standard one with the following modifications: (a) after phenol/chloroform separation, the clear upper aqueous layer containing RNA (~700 µl) was divided into two different aliquots: (i) an “upper aliquot” composed by the first 300 µl, (ii) a “lower aliquot” by the subsequent 300 µl. The remaining 100 µl closer to the interface have been discarded. (b) In order to remove ethanol residuals, RNA precipitates were dried in a

biological hood for 20' and resuspended in 60 µl of RNase and DNase free water before data acquisition.

Results and discussion

Yield and quality of nucleic acids were evaluated after extraction and purification procedures, both standard tests and tailored for vibrational analysis, by conventional measurements with Nanodrop® [17] by means of the $A_{260/280}$ and $A_{260/230}$ ratios. This method relies on the measurement of the optical density of nucleic acid solutions in the UV-range.

	Concentration [µg/µL]	$A_{260/280}$	$A_{260/230}$		Concentration [µg/µL]	$A_{260/280}$	$A_{260/230}$
DNA	0,25	1,86	1,99	**RNA**	2,71	2,05	1,89
DNA	0,23	1,87	1,97	**RNA**	2,92	2,08	1,99
DNA	0,31	1,82	1,89	**RNA**	3,14	2,03	1,92
DNA	0,27	1,84	1,94	**RNA**	2,47	2,07	1,97
DNA	0,24	1,79	2,01	**RNA**	2,71	2,01	1,97
DNA	0,31	1,82	2,03	**RNA**	2,65	1,97	1,99
DNA	0,21	1,78	2,00	**RNA**	3,01	1,98	2,03
DNA	0,26	1,83	2,02	**RNA**	4,07	2,04	2,01
DNA*	0,15	1,86	1,98	**RNA**	3,71	1,97	1,97
DNA*	0,13	1,87	1,97	**RNA***	2,02	2,00	2,02
DNA*	0,11	1,77	2,05	**RNA***	2,14	2,03	1,87
DNA*	0,17	1,84	1,99	**RNA***	1,47	2,02	2,10
DNA*	0,14	1,79	2,01	**RNA***	1,71	2,01	2,02
DNA*	0,13	1,82	2,03	**RNA***	1,65	1,99	1,99
DNA*	0,10	1,85	2,02	**RNA***	2,10	1,98	2,03
DNA*	0,16	1,83	2,00	**RNA***	2,07	2,02	2,01

Table 4.1: DNA. Concentrations and UV absorbance ratios $A_{260/280}$ and $A_{260/230}$ of different solutions of DNA extracted from B16 cell cultures with standard (DNA) and tailored (DNA*) protocols. Each solution was extracted from a full plate of confluent B16 mouse melanoma cells. **RNA.** Concentrations and UV absorbance ratios $A_{260/280}$ and $A_{260/230}$ of different solutions of RNA extracted from B16 cell cultures with standard (RNA) and tailored (RNA*, selecting the "upper aliquot") protocols. (Reported from [3] S.I.)

Both DNA and RNA absorb maximally at 260 nm, whereas most proteins at 280 nm [18]. Therefore, the ratio A260/280 is used to assess the purity of both DNA and RNA in terms of removal of any protein material. In biology, a ratio of ~1.8 is generally accepted as "pure" for DNA, whereas a ratio of ~2.0 is generally accepted as "pure" for RNA [19]. If the ratio is appreciably lower, in both cases, it may indicate the presence of protein, phenol, or other

contaminants that absorb strongly at, or near, 280 nm. The $A_{260/230}$ ratio is the second parameter used to evaluate the quality of nucleic acids. Some molecular contaminants absorb at 230 nm, such as residues of guanidinium hydrochloride, phenol, or ethylenediaminetetraacetic acid. For a generally accepted pure nucleic acid, the $A_{260/230}$ ratio often sets in the range of 1.8–2.2. If the ratio is appreciably lower, this may indicate the presence of co-purified contaminants [17]. $A_{280/260}$ and $A_{260/230}$ values of our nucleic acid solutions extracted from B16 cells with both standard and optimized extraction procedure were in the standard range (see Table 4.1).

Once evaluated the quality of the samples with conventional assays, their vibrational characterization was performed with both ATR-FTIR and UV-RR spectroscopies before and after the optimization of the extraction protocols. More details for spectroscopic measures parameters and instrumentation can be found in [3] "Materials and methods". FTIR measures were performed first. ATR-FTIR spectra were acquired using a Germanium IRE in continuous mode, during the drying process of a droplet of a solution 0.1 mg/mL of DNA or RNA. Figure 4.3a reported the main FTIR results obtained for B16 genomic DNA, while Figure 4.3b shows the results obtained for extracted total RNA.

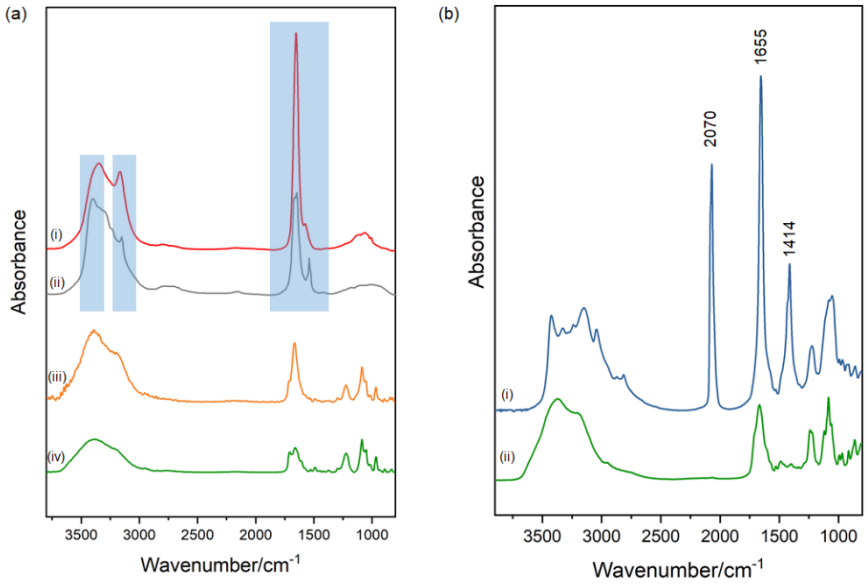

Figure 4.3: ATR-FTIR spectra of different DNA and RNA solutions extracted from B16 cell cultures. Measures have been performed letting the solution drying. (a) ATR-FTIR acquisitions on B16 DNA samples: (i) B16-DNA extracted with the standard protocol, (ii) guanidinium hydrochloride salt, from [20]; (iii) B16-DNA extracted with the tailored extraction protocol; (iv) commercial certified high-purity Salmon DNA. The vertical blue bars highlight the guanidinium hydrochloride vibrational peaks. (b) ATR-FTIR spectra of RNA extracted from B16 cells at different stages of purifications. (i) B16 RNA extracted with the standard protocol;

(ii) B16-RNA extracted with the tailored extraction protocol. Peak positions of the main contaminant spectral features are highlighted in the picture. (Adapted from [3].)

In curve (i) in Figure 4.3a is shown a representative spectrum of DNA extracted with the standard protocol after water evaporation. By comparing this spectrum with that of a standard pure genomic DNA (spectrum (iv), lyophilized salmon DNA from Sigma Aldrich), they look very different. In particular, strong mutual features peaked at 3166, 3150, 1670, and 1564 cm^{-1} (highlighted by blue boxes), which corresponds mainly to N─H, C═N, and C─N bonds, dominate the spectra of DNA samples obtained by the conventional extractive method. They can be ascribed to the co-precipitation with DNA of guanidinium hydrochloride during the isolation steps [20]. For qualitative comparison, we reported in Figure 4.3a the spectra of guanidinium hydrochloride (Curve (ii), from reference [19]). Guanidinium hydrochloride is a compound commonly employed for protein denaturation, and it is one of the constituents of the lysis buffer (AL buffer) used in the extractive protocol. As a matter of fact, proteins readily dissolve in concentrated solutions of guanidinium hydrochloride (4 M or higher) and become biologically inactive because they lose their secondary structure, while secondary structure of nucleic acids is not affected to the same extent [20]. The strong intensity of the guanidinium bands centered at 1670 and 1564 cm^{-1} compromises the interpretation of DNA features in the 1800–1500 cm^{-1} range [10]. The guanidinium spectrum is also characterized by a broad and relatively weak band in the 1300–1000 cm^{-1} spectral range. This band covers the contributions of asymmetric and symmetric stretching bands of phosphates moieties of DNA and RNA, which are the most important for obtaining information on nucleoside-specific interaction and nucleic acid conformation [3], [10]. These bands are characteristic of vibrations localized to the sugar and sugar-phosphate moieties that give rise to vibrational modes sensitive to backbone conformation and sugar puckering.

Reconsidering the extraction protocol, we noticed a close dependence between the intensity of guanidinium signals and the concentration of the starting cellular material: the higher the concentration, the most the guanidinium signals were intense. Indeed, Qiagen spin columns have a limited load capacity, and the standard protocol suggests limiting the cell growth to 4–5 million of cells. Lysates from larger cell populations saturate the columns and, consequently, could increase the quantity of salt in the final solutions. Therefore, we decided to reduce the number of loaded cells to 3 million, in order to further diminish the chance of spin column saturation. Moreover, two additional spin-dry steps after the conventional washing ones and a pre-elution step have been introduced, in order to minimize the residual

contaminants into the column before the final elution. By comparing Curves (iii) (DNA by tailored protocol) and (ii) (DNA by standard protocol), we can appreciate that the strong absorption bands of guanidinium at 3166 and 1670 cm^{-1} are greatly reduced, whereas the contribution at 1564 cm^{-1} is under the limit of detection. Also, the 1500–1200 cm^{-1} range seems to be free from any guanidinium contribution. The comparison between the spectrum B16-DNA extracted with our tailored-protocol (iii) and that of standard certified high-quality Salmon DNA (iv) shows that the overall shape of DNA samples obtained with our tailored protocols is consistent with the standard one, especially in the range of the phosphate stretching vibrations (1250-900 cm^{-1}) which is the most meaningful region for the characterization of nucleic acids structures. Nevertheless, spectra comparison reveals that guanidinium salts are still present in the extracted DNA. Instead, protein contamination of our DNA solutions can be excluded, because the most characteristic protein bands, usually peaked at ~1650 and ~1540 cm^{-1} (Amide I and II, respectively), cannot be detected in our spectra.

Overall, we can conclude that the optimized extraction protocol and the adopted FTIR sampling method allow getting reliable information on the most diagnostic spectral bands of cellular B16-DNA, thanks to the reduction of the spectral contribution of interfering agents, in this case guanidinium hydrochloride.

On the other side, B16 total RNA solutions were characterized according with the same sequence of measures used for DNA. For isolation and purification of RNA from cells, we adopted a protocol based on the conventional phenol/chloroform extractive method, particularly efficient for RNA extraction [21]. The aforementioned method leads to the separation of the cell lysate into three phases: an upper clear aqueous layer of 700 µl containing RNA, a second very thin solid interphase containing DNA, and a lower red organic layer containing lipids, proteins, and other chemical moieties. During the recovery of the aqueous layer containing RNA, a small fraction of the organic lower solution, containing the larger fraction of contaminants, can be sampled as well. As a matter of fact, acquired B16-RNA ATR-FTIR spectra showed absorption bands peaked at 1414, 1655, 2070 cm^{-1} and above 2970 cm^{-1} (See Figure 4.3b, curve (i)), that could be assigned to guanidinium. It is indeed known that Isol-RNA lysis reagent (5-Prime) contains guanidinium thiocyanate, which is one of the most effective protein denaturants, able to efficiently denature endogenous ribonucleases [21]. Under this hypothesis, the strong absorption peak at ~2070 cm^{-1} has been assigned to the S─CN stretching of guanidinium thiocyanate [22].

Thus we decided to divide the aqueous layer into two portions: (a) a first portion composed by the first 300 µL of aqueous solution, called "upper aliquot," which is far from the white interphase; and (b) a "lower aliquot" composed by the following 300 µL. As postulated, the "lower phase" was still contaminated with guanidinium thiocyanate moieties (data not shown). Instead, the spectrum of the "upper aliquot" was free from guanidinium main contributions (See Figure 4.3, curve (ii)). Unfortunately, we could not match the acquired spectra with a standard during our experiments. However, B16-RNA spectra from the sample extracted with the tailored protocol are consistent with those reported in literature, especially in the region of phosphates vibrations, where all the fingerprint peaks of ribose-phosphate backbone are clearly discerned.

For both RNA and DNA tailored protocols, the repeatability of the optimized extraction procedure was verified by measuring several prepared solutions (see Figure 4.4).

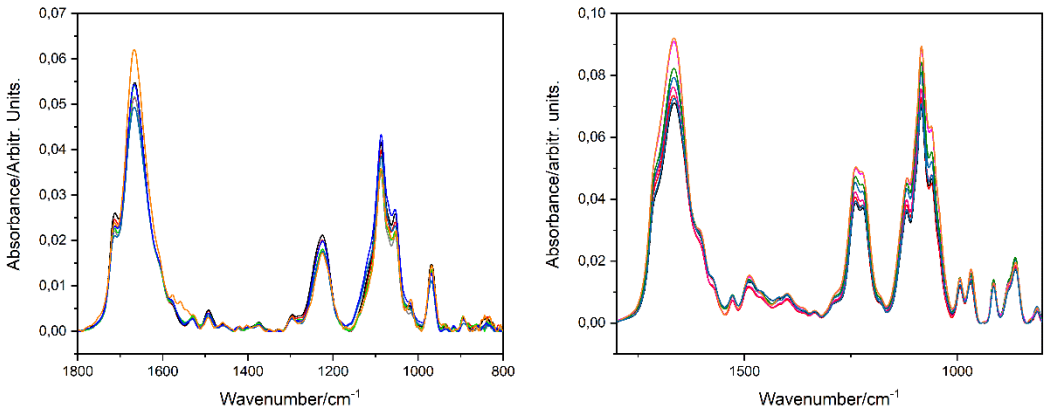

Figure 4.4: Left panel: ATR-FTIR spectra acquired on seven different B16 genomic DNA solutions obtained by using the optimized extractive protocol based on the Qiamp DNA blood Mini Kit (Qiagen). (b) Right panel: ATR-FTIR spectra acquired on eight different B16 genomic RNA solutions obtained by using our optimized extractive protocol based on the TRIzol RNA purification kit (Thermo Scientific). (Reported from [3] S.I.)

After ATR-FTIR characterization, we performed a series of UV-RR measurements directly on both B16-DNA and RNA samples obtained with tailored protocols.

Results for DNA solutions are summarized in Figure 4.5. Curve A in Figure 4.5 represent the typical Raman spectrum of B16-DNA. Despite the good quality showed by FTIR spectroscopy for such solutions, UV-RR clearly shows the presence of some Raman extra peaks centered at about 869, 1049, 1086, and 1455 cm^{-1} (highlighted in red in the Figure 4.5). Such peaks have been assigned to the presence of ethanol [23]. Ethanol, indeed, is used in DNA isolation procedures in order to prevent DNA dissolution in water (as in salting-out and organic extractions) and to optimize the binding of DNA to solid state supports. In

particular, in the used kit, ethanol promotes DNA aggregation and precipitation onto the silica of the spin columns during the first isolation step, and it is present in both the washing buffers, to elute contaminants without losing DNA. The presence of ethanol should be detected also by FTIR spectroscopy. However, since we analyzed protein samples onto ATR crystals after the solution drying, since ethanol evaporated with the solvent. Indeed, side ATR-FTIR acquisitions, made in continuous mode during solvent evaporation, clearly showed that in the first acquisitions the presence of ethanol was detectable (data not shown). It has to be mentioned that the ethanol rate of evaporation is slower for UV-RR measures, due to the acquisition geometry (i.e., the evaporation rate from the IRE crystal is faster than from UV cuvette).

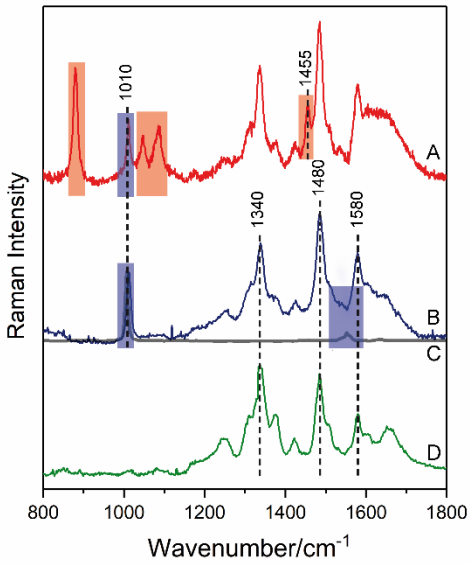

Figure 4.5: UV-RR spectra of B16-DNA aqueous solutions obtained by tailored extractive method before (a) and after (b) the ethanol removal procedure. Spectra of guanidinium hydrochloride salt (c, from [20]) and Salmon DNA aqueous solutions (d) are also reported in the same panel for comparison. The red and blue boxes highlight the main vibrational peaks assigned to ethanol and guanidinium hydrochloride, respectively. (Reported from [3])

Analyzing to the spectrum A of in Figure 4.5, a strong band centered at about 1010 cm^{-1} is detectable and can be associated to the residual guanidinium hydrochloride [20], [24]. Although the guanidinium main peak falls in a spectral region free from other relevant vibrational features, the ethanol peak centered at 1455 cm^{-1} overlaps with the C─C and C─N stretching vibrations of adenine and guanine (1455–1480 cm^{-1}) [12]. In 2017, Cammisuli et al. proposed a simple way for minimizing ethanol contamination in UV-RR measurements [23], taking advantage from the temperature stability of DNA. In Cammisuli et al., DNA

solutions were first warmed up to 80 °C in order to accelerate ethanol evaporation and then cooled down to room temperature for acquiring spectra of DNA solution free from ethanol. This procedure has been applied to our DNA solutions. An example of the UV-RR B16-DNA spectra obtained by applying also the described evaporation treatment is shown in Figure 4.5 curve B. No peaks at 869, 1049, 1086 and 1455 cm^{-1} can be detected in this spectrum, confirming the ethanol evaporation. Nevertheless, the contamination from guanidinium still persists. For comparison, the Raman spectrum of dehydrated guanidinium hydrochloride has been reported in Figure 4.5 curve C, normalized to the maximum of the peak at 1010 cm^{-1} of curve B. The spectrum of guanidinium hydrochloride shows the presence of a second peak centered at ~1550 cm^{-1} with intensity near ~0.15 times the peak at 1010 cm^{-1}. Due to its small intensity, the contribution of guanidinium can be considered negligible in this spectral region. B16-DNA spectrum has several Raman features in the spectral range 1100–1800 cm^{-1}, the most intense of which are centered at 1340, 1480, and 1580 cm^{-1}. According to the literature [25]–[27] the first peak can be assigned to adenine aromatic ring vibrations, the second one to an overlapping between adenine and guanidine internal ring vibrations, whereas the latter to adenine and guanine-CH_2 scissoring. To further confirm the correspondence between the aforementioned spectral features and the DNA UV-RR vibrational spectra, in Figure 4.5 curve D is the spectrum of the commercial Salmon DNA aqueous solution acquired with same parameters. The direct comparison between spectrum B (B16 DNA) and spectrum D (reference) shows a good level of analogy in spectral peak positions and relative intensities. All these considerations confirm that, thanks to our tailored protocol, we could obtain a complete and reliable UV-RR spectrum of DNA from B16 cell line.

Then, UV-RR measurements were done onto B16-RNA solutions obtained from the “upper aliquot”. Figure 4.6a shows a typical Raman spectrum collected from the B16-RNA aqueous solution. Differently from what observed for DNA, there is a strong fluorescence signal that overlaps with the Raman features of RNA. In order to remove the fluorescence background, we measured a blank solution, prepared by following the same steps suggested by TRIzol protocol using water, instead of the cellular lysate. The UV-RR spectrum obtained for this solution (reported in Figure 4.6a as a red line) has a slope similar to the one of RNA spectrum. Therefore, by subtracting the blank spectrum from the experimental profile of RNA (Figure 4.6b), we could retrieve the spectrum of it. The Raman signal obtained by following this procedure is in accordance with the UV-RR RNA spectra reported in literature [28].

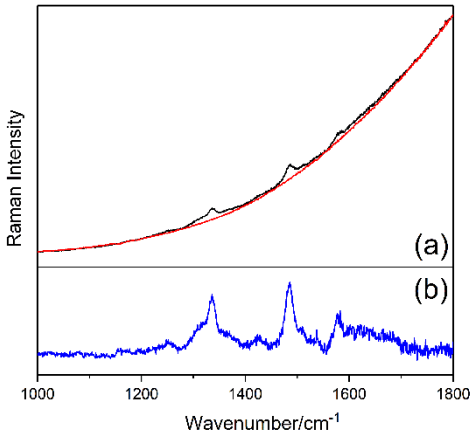

Figure 4.6: (a) Ultraviolet-resonant Raman spectrum of aqueous solution of B16-RNA extracted with the optimized protocol (back curve). The spectrum called "blank solution," showing a severe fluorescence background, is plotted in red. (b) Spectral difference (multiplied by 5) between the spectrum of B16-RNA and the blank solution spectrum. (Reported from [3]).

It is noteworthy to consider that the presence of such a strong fluorescence implies that some contaminants may be still present. However, their concentration is sufficiently low to not show any vibrational feature overlapped with the ones of RNA, thus to not affecting the quality of the data in any significant way.

4.3 Conclusions

In conclusion, our tailored protocols can be applied to extract genetic material from general biological samples, preserving a good yield and improving at the same time spectroscopic quality for routine applications, allowing to collect vibrational data more reliable and thus promoting the use of vibrational techniques for molecular biology studies. In particular, the ability to obtain nucleic acid samples from different biological sources and to reveal subtle changes associated to their structure opens new possibilities for the study of nucleic acids modifications, such as those induced by epigenetic disorders.

References

[1] F. D'Amico *et al.*, "UV resonant Raman scattering facility at Elettra," *Nucl. Instrum. Methods Phys. Res. Sect. Accel. Spectrometers Detect. Assoc. Equip.*, vol. 703, pp. 33–37, 2013.

[2] P. Zucchiatti *et al.*, "Contribution of Ribonucleic Acid (RNA) to the Fourier Transform Infrared (FTIR) Spectrum of Eukaryotic Cells," *Anal. Chem.*, vol. 88, no. 24, pp. 12090–12098, Dec. 2016.

[3] P. Zucchiatti *et al.*, "The quality is in the eye of the beholder: The perspective of FTIR and UV resonant Raman spectroscopies on extracted nucleic acids," *J. Raman Spectrosc.*, vol. 49, no. 6, pp. 1056–1065, Jun. 2018.

[4] B. Alberts, A. Johnson, J. Lewis, M. Raff, K. Roberts, and P. Walter, *Molecular Biology of the Cell*, 4th ed. Garland Science, 2002.

[5] fisiologicamente, "Biologia | Gli acidi nucleici: DNA e RNA," *Fisiologicamente*, 14-Oct-2016. .

[6] T. Shimanouchi, M. Tsuboi, and Y. Kyogoku, "Infrared Spectra of Nucleic Acids and Related Compounds," in *Advances in Chemical Physics*, John Wiley & Sons, Ltd, 2007, pp. 435–498.

[7] D. M. Byler and H. Susi, "Examination of the secondary structure of proteins by deconvolved FTIR spectra," *Biopolymers*, vol. 25, no. 3, pp. 469–487, Mar. 1986.

[8] L. Ashton *et al.*, "UV resonance Raman spectroscopy: a process analytical tool for host cell DNA and RNA dynamics in mammalian cell lines," *J. Chem. Technol. Biotechnol.*, vol. 90, no. 2, pp. 237–243, Feb. 2015.

[9] D. A. Guzmán-Embús, M. Orrego Cardozo, and C. Vargas-Hernández, "Genomic DNA characterization of pork spleen by Raman spectroscopy," *J. Appl. Phys.*, vol. 114, no. 19, p. 194704, Nov. 2013.

[10] B. R. Wood, "The importance of hydration and DNA conformation in interpreting infrared spectra of cells and tissues," *Chem. Soc. Rev.*, vol. 45, no. 7, pp. 1980–1998, Apr. 2016.

[11] M. Banyay, M. Sarkar, and A. Gräslund, "A library of IR bands of nucleic acids in solution," *Biophys. Chem.*, vol. 104, no. 2, pp. 477–488, Jun. 2003.

[12] F. D'Amico *et al.*, "Oxidative damage in DNA bases revealed by UV resonant Raman spectroscopy," *The Analyst*, vol. 140, no. 5, pp. 1477–1485, Mar. 2015.

[13] S. A. Asher, "UV Resonance Raman Spectroscopy for Analytical, Physical, and Biophysical Chemistry," *Anal. Chem.*, vol. 65, no. 4, p. 201A–210A, Feb. 1993.

[14] G. Kister, G. Cassanas, and M. Vert, "Effects of morphology, conformation and configuration on the IR and Raman spectra of various poly(lactic acid)s," *Polymer*, vol. 39, no. 2, pp. 267–273, Jan. 1998.

[15] J. M. Benevides, S. A. Overman, and G. J. Thomas, "Raman, polarized Raman and ultraviolet resonance Raman spectroscopy of nucleic acids and their complexes," *J. Raman Spectrosc.*, vol. 36, no. 4, pp. 279–299, Apr. 2005.

[16] D. C. Rio, M. Ares, G. J. Hannon, and T. W. Nilsen, "Purification of RNA using TRIzol (TRI reagent)," *Cold Spring Harb. Protoc.*, vol. 2010, no. 6, p. pdb.prot5439, Jun. 2010.

[17] P. Desjardins and D. Conklin, "NanoDrop Microvolume Quantitation of Nucleic Acids," *J. Vis. Exp. JoVE*, no. 45, Nov. 2010.

[18] V. A. Huss, H. Festl, and K. H. Schleifer, "Studies on the spectrophotometric determination of DNA hybridization from renaturation rates," *Syst. Appl. Microbiol.*, vol. 4, no. 2, pp. 184–192, 1983.

[19] S. R. Gallagher, "Quantification of DNA and RNA with absorption and fluorescence spectroscopy," *Curr. Protoc. Cell Biol.*, vol. Appendix 3, p. Appendix 3D, May 2001.

[20] O. D. Bonner and C. F. Jordan, "The infrared and Raman spectra of guanidinium salts," *Spectrochim. Acta Part Mol. Spectrosc.*, vol. 32, no. 5, pp. 1243–1246, Jan. 1976.

[21] P. Chomczynski and N. Sacchi, "Single-step method of RNA isolation by acid guanidinium thiocyanate-phenol-chloroform extraction," *Anal. Biochem.*, vol. 162, no. 1, pp. 156–159, Apr. 1987.

[22] N. Dizge, E. Demirbas, and M. Kobya, “Removal of thiocyanate from aqueous solutions by ion exchange,” *J. Hazard. Mater.*, vol. 166, no. 2, pp. 1367–1376, Jul. 2009.
[23] F. Cammisuli, L. Pascolo, M. Morgutti, A. Gessini, C. Masciovecchio, and F. D’Amico, “Avoiding Ethanol Presence in DNA Samples Enhances the Performance of Ultraviolet Resonance Raman Spectroscopy Analysis,” *Appl. Spectrosc.*, vol. 71, no. 1, pp. 152–155, Jan. 2017.
[24] K. Mishiro, F. Hu, D. W. Paley, W. Min, and T. H. Lambert, “Macrosteres: The Deltic Guanidinium Ion,” *Eur. J. Org. Chem.*, vol. 2016, no. 9, pp. 1655–1659, Mar. 2016.
[25] S. P. A. Fodor and T. G. Spiro, “Ultraviolet resonance Raman spectroscopy of DNA with 200-266-nm laser excitation,” *J. Am. Chem. Soc.*, vol. 108, no. 12, pp. 3198–3205, Jun. 1986.
[26] S. P. A. Fodor, R. P. Rava, T. R. Hays, and T. G. Spiro, “Ultraviolet resonance Raman spectroscopy of the nucleotides with 266-, 240-, 218-, and 200-nm pulsed laser excitation,” *J. Am. Chem. Soc.*, vol. 107, no. 6, pp. 1520–1529, Mar. 1985.
[27] J. R. Perno, C. A. Grygon, and T. G. Spiro, “Ultraviolet Raman excitation profiles for the nucleotides and for the nucleic acid duplexes poly(rA)-poly-(rU) and poly(dG-dC),” *J. Phys. Chem.*, vol. 93, no. 15, pp. 5672–5678, Jul. 1989.
[28] J. M. Benevides, M. Tsuboi, J. K. Bamford, and G. J. Thomas, “Polarized Raman spectroscopy of double-stranded RNA from bacteriophage phi6: local Raman tensors of base and backbone vibrations,” *Biophys. J.*, vol. 72, no. 6, pp. 2748–2762, Jun. 1997.