UNIVERSITÉ DE BOURGOGNE

FACULTÉ DES SCIENCES

Laboratoire Interdisciplinaire Carnot de Bourgogne

Département Nanosciences

# Spectral and temporal distribution of biomolecules by Dynamic SERS

Thèse présentée à l'Université de Bourgogne pour obtenir

le grade de Docteur en Physique

par

## Thibault BRULÉ

Soutenue le **24 Octobre 2014** devant la commission d'examen
composée de :

| | | |
|---|---|---|
| FÉLIDJ N. | Pr., Université Paris VII Diderot | Rapporteur |
| ADAM P.M. | Pr., Univ. Technologie de Troyes | Rapporteur |
| LAMY DE LA CHAPELLE M. | Pr., Université Paris XIII Nord | Examinateur |
| HOERBER H. | Pr., Université de Bristol | Examinateur |
| DEREUX A. | Pr., Université de Bourgogne | Examinateur |
| FINOT E. | Pr., Université de Bourgogne | Directeur de thèse |
| BOUHELIER A. | DR. CNRS, Univ. de Bourgogne | Co-encadrant de thèse |

*Rester immobile ne sert à rien. Il faut choisir entre progresser ou régresser. Allons donc de l'avant et le sourire aux lèvres.*

**Lord R. Baden-Powell**

# Remerciements

Ce travail de thèse n'aurait pas été celui qu'il est devenu sans la présence et le soutien de bon nombre de personnes que j'aimerais remercier ici même.

Tout d'abord, merci à Gilles Bertrand et plus encore Alain Dereux de m'avoir accueilli au sein du Laboratoire Interdisciplinaire Carnot de Bourgogne à Dijon, et plus précisément au sein de l'équipe Optique Submicronique et Nano-Capteurs dont je remercie le chef d'équipe Eric Lesniewska. Je remercie également Hans Jauslin et tous les membres du conseil de l'école doctorale Carnot-Pasteur de la confiance qu'ils m'ont accordé en m'attribuant un financement de thèse du Ministère de l'Enseignement Supérieur et de la Recherche.

Je tiens ensuite à remercier celui qui fut tel un guide durant toute cette thèse, qui m'a repoussé dans mes retranchements mais m'a aussi élevé plus haut aux moments venus. Je parle ici de mon directeur de thèse, Eric Finot. Nos discussions, souvent quelques peu utopiques, resteront pour moi de grands moments poussant toujours plus loin la réflexion et l'interprétation du travail réalisé et à venir. De la même façon, je suis reconnaissant envers Alexandre Bouhelier, co-encadrant de ma thèse, qui m'a laissé toute la liberté d'action dont j'avais besoin mais qui était là aux moments où j'avais besoin d'être épaulé notamment en termes de réglages optiques ou de fabrication. Son point de vue a toujours enrichit mon travail et je l'en remercie.

Je remercie aussi les membres du jury de ma thèse qui ont accepté de libérer du temps pour juger et enrichir mon travail. Je remercie donc tout particulièrement les deux rapporteurs Nordin Félidj et Pierre-Michel Adam, ainsi que les examinateurs Marc Lamy de la Chapelle, Heinrich Hoerber et Alain Dereux.

J'ai une pensée pour les différents personnels permanents que j'ai cotoyés durant

3

ces trois ans: Aymeric pour ses discussions, ses conseils, son écoute, Laurent pour son écoute et ses conseils en chimie et microfluidique notamment, Gérard pour tous ses conseils et ses analyses, Jean Claude pour sa perspicacité, Juan pour sa bonne humeur et sa fidélité inconditionnelle à l'Espagne, Elly pour sa disponibilité, Carmen pour ses échanges et son écoute, Eric B., Kamal et Christian M. pour leur bonne humeur et leur présence.

Je tiens aussi à remercier toutes les personnes que j'ai cotoyées en donnant des enseignements. Merci à Christophe de m'avoir fait confiance, notamment en m'octroyant la responsabilité d'un module de Licence 3 et Benoit qui s'est porté volontaire pour me seconder dans cette tâche.

Ma vie au laboratoire ne s'est pas résumé à des rencontres avec des personnels permanents mais aussi des "gens de passages", plus ou moins longtemps d'ailleurs. Certains ont compté, et compteront, plus que d'autres. Je pense ici à ceux qui m'ont été proche durant un certain temps et avec qui j'ai partagé de si bons moments, au bureau mais aussi en dehors : Christian, toi qui est celui qui a subi le plus mes humeurs, notamment en fin de rédaction, avec qui j'ai passé tant de pauses café/goûter, tant discuté et échangé, et tellement d'autres choses, n'oublies pas, rien n'est fini ; Delphine, alias Tatie Crème, Floriane et Arnaud, vous qui m'avez tellement conseillé, ouvert les yeux, changé les idées, fait revenir sur terre mais aussi valorisé, etc, vous le savez, même après vous resterez ; Marie, Mickaël, Stéphane, Paco, Mélanie, vous qui m'avez supporté et m'avez sorti de ma routine dans le bureau ; Hélène, Jérémie, Padmnabh, vous qui avez participé à mon projet et avez permis de réussir cela ; Jean-Emmanuel de prendre la suite de ma thèse et permettre de continuer les travaux que j'ai entamé dans l'équipe ; Thomas, Karim, Simon, Pauline, Ece, Serkan, Jean, Antonin, Olivier, David, Aurélien, Mingxia, Johan, Rémi, Houssam, Michael, Nicolas, Florent, et j'en oublie, qui avez été là, près de moi, et sans qui rien n'aurait été pareil. Merci.

Je tiens à remercier les gens que j'ai pu rencontrer pendant ma thèse dans le cadre des différents projets auxquels j'ai participé, qu'il s'agisse de SPEDOC, contrat européen en collaboration avec l'INSERM de Dijon, COSINGO et l'ICFO en Espagne et l'EPFL de Lausanne, notamment Olivier et Andrea avec qui j'ai pu partager sur l'analyse statistique. Je pense aussi à Emilie, Nicolas et Loïc du CSGA-INRA de Dijon avec qui j'ai travaillé sur quelques molécules odorantes. Enfin, je remercie Zoya et Lise pour leur dynamisme et leur joie de vivre que l'on a partagé quelques semaines.

J'ai aussi une pensée pour ceux que j'ai rencontré en dehors du laboratoire durant ces trois années dijonnaises, et notamment toutes les personnes extraordinaires que j'ai appris à connaitre et à découvrir à travers mon engagement chez les Scouts et Guides de France du territoire Bourgogne Nord, aussi bien dans l'équipe territoriale (Marie-Hélène, Laurent, Stéphanie, Laurent, Thomas, Emmanuel, Pauline, Cécile, Emilie,...)

que les chefs et cheftaines, les RGs et leurs équipes de groupe.

Et pour finir, je remercie mes proches qui m'ont soutenu avant, pendant et qui continueront après cette thèse. Je parle bien sûr de mes parents, mes frères, mes grands-parents, oncles et tantes, ma belle-famille, mais surtout de mon épouse Chloé qui m'a supporté dans les bons et les mauvais moments et mon fils Timéo qui m'apporte tellement de bonheur au quotidien.

A tous, Merci.

# Contents

# Introduction

For forty years, the development of sensor suitable for biological samples remains a major challenge. Despite remarkable progress in terms of either sensitivity or selectivity, many fundamental problems have not been yet studied to combine the two specificities.

The detection of traces of biomolecules in a mix with a universal efficiency is indeed the dream of people working in sensors domain. Different techniques have been developed in different scientific domains, from electronic microscopy to cantilever deflection, from NMR to Elisa tests. Here, our choice is the non-invasive optical method, and more specifically Raman spectroscopy.

Much works based on Raman spectroscopy have been completed in many scientific domains as atmospheric studies, biosensing, etc. The decisive advantages of Raman spectroscopy are in its high selectivity and the very low invasivity of the technique. Unfortunately, the performance of concentration detection is limited by the Raman cross-section. The response time needed for a good Signal-Noise ratio is also long, more than ten minutes. The discovery in 1970s of Surface Enhanced Raman Scattering or Spectroscopy (SERS) has revolutionized the detection of trace molecule. When Van Duyne *et al.* bring a molecule into contact of a metallic surface to investigate the influence of this surface, they observed a huge enhancement of Raman spectra. In fact, the metallic structures enhance the Raman cross-section based on its plasmonic vibration.

Since then, many improvements were purposed in literature, especially concerning the metallic structures used to enhance the Raman signal. The localized plasmonic effect is higher when metallic nanostructures, made of in gold or silver, are closed one to the others. SERS should be able to combine both the high selectivity from Raman spectroscopy with the high sensitivity due to the high local enhancement of this signal

by plasmonic resonance.

The detection of SERS active molecules was mostly focused on dye molecules either for quantifying the local enhancement of hot spot or as a marker of large molecules. This work aims at determining the chemical fingerprint of isolated biomolecules based on the temporal and spectral fluctuations of Raman scattering analysis.

Up to now, no commercial technique is available. This is principally due to different points. Firstly, the substrate preparation induces a significant cost, principally due to electron beam lithography. Secondly, the miniaturization and parallelization of the study are still low developed. One solution for this point is the use of microfluidic chip. However, the use of microfluidic implies that the time of acquisition is reduced. In fact, under the flow, the sounded molecules pass through the plasmonic resonant spot. The residence time of the molecule in the spot is very fast, around 10 ms for a single molecule, and so induces a time of acquisition of the same order of magnitude. Thirdly, the sensitivity is also linked to the nature of the targeted molecule. Indeed, there are overall two different types of molecule that should be sounded in SERS: the resonant and the non-resonant molecules. The first ones are Raman active molecule presenting a higher Raman cross section than non-resonant molecules. Most of the biomolecules are non-resonant.

All of this implies also the necessity of a highly efficient and low cost plasmonic substrate coupled with a fast acquisition setup and a specific spectral analysis to identify the targeted molecule in a mix or a complex medium as blood serum.

So in the first chapter will present the state of the art concerning the development of SERS based nanobiosensors. In Chapter 2, the developed experimental setup and its data analysis will be detailed. Thirdly, single molecule detection will be proved and the quantification of the concentration will be investigated. In Chapter 4, the selectivity of the SERS substrate towards sorting will be demonstrated. Finally, in the last chapter, the detection and identification of single proteins in complex media will be completed.

## State of the art of development of SERS based nanobiosensors

## 1.1 Introduction

A lot of developments have been done in SERS based nanosensors. In this part, the concept of sensor will be firstly presented. Then, I describe the concept of molecular plasmonic, and more particularly its application in the case of SERS, and how SERS was applied for sensing. Finally, I will explain our choices in terms of enhancing substrate, microfluidic, molecules of interest and analysis methods use after.

## 1.2 Nanobiosensors

### 1.2.1 Definition of a biological sensor

Study of biological samples implies specific points of attention. In fact, these samples are very sensitive to the experimental conditions modifications and are easily affected by mechanical impacts or environmental evolutions. Specifically, two types of approaches stand out.

**Chemical approach: Biochip**

If one has to detect several proteins, an array of protein sensors will be required. To construct such an array consisting of hundreds or thousands of sensors will lead to an enormously bulky and costly sensor, which will be virtually impractical to fabricate and use. The multiplexing capability offered by nanosensors is a vast improvement for

real-time protein composition monitoring. An array of thousands of nanosensors, each coated with a different functional group and hence tuned to a prefixed analyte, can be used in a single device. Such an ultra-miniaturized, low-power device supplemented with signal-processing and pattern-recognition algorithms will exhibit effective discrimination capabilities among target analytes, producing a unique signature or fingerprint on exposure to a specimen containing a mixture of chemical and biological characters.

The above situation is comparable with the historical development of computers. The primitive computers were so big that they required full rooms for their storage, using high power and producing heat. With advancement of technology, the size decreased reaching desktop personal computers, then laptops, and then palmtop computers. Power consumption was drastically reduced in this progression reaching very low levels in battery-operated computers. At the same time, the capabilities of the computers also improved. This was enabled by fast strides made by very-large-scale and ultra-large-scale ICs, very-Large-scale Integration (VLSI), and ultra-Large-Scale Integration (ULSI).

## Physical approach: Spectroscopy

Multifunctionality is a versatile feature in nature. The human tongue is a taste sensor and also used in speech. Human nose is used for smelling and breathing. Similarly many other organs serve multiple functions. The nose principle can be presented as a spectroscopic method. Indeed, different molecules can be detected without specific chemical reaction but simply by a spectral fingerprint determination of the molecule. This spectrum can be a mass spectrum, an absorption spectrum, a light scattering spectrum.

This approach allows to more or less completely delete the chemical functionalization needed to differentiate different biomolecules as proteins.

## Time evolution of biosensors

A little contemplation will tell that macro and microsensors have many limitations. They cannot be used in many situations. Small size and light-weight sensors help in making portable instruments, which are essential for military and aerospace applications besides mobile and handled consumer products.

To consider an example, a predicted application of nanosensors is in the detection of cancerous cells in the human body by injecting quantum dots. These quantum dots are crystals of semiconductor materials having sizes in the nanometer scale and emitting fluorescent radiation. These crystals are made of cadmium selenide (CdSe), cadmium sulfide (CdS), or indium gallium phosphide (InGaP) and coated with suitable polymers that safeguard human cells from toxic action of cadmium and also allow attaching

molecules that enable tracking of cell processes and cancers. Obviously, such an application cannot be imagined without developing nanosensors. Macro- or microsensors cannot be inserted in the human body without harming it or disturbing its normal functioning. So it is the small size of nanosensors that is advantageously exploited here. Futuristic advances can open unprecedented perspectives for the application of nanosensors as molecular-level diagnostic and treatment instruments in medicine and as networks of nanorobots for real-time monitoring of physiological parameters of a human body.

System biology, which is currently taking off as research discipline to explore the basic principles of living systems by quantitative modelling of inter- and intracellular processes, will starve for nanosensors and tools to provide data for model verification. Implantable devices like autonomous nanorobots or multifunctional endoscopes (medical devices consisting of a long, thin, flexible or rigid tube that has a light and a video camera, whereby images of the inside of the patient's body can be seen on a screen) for minimal invasive diagnostics, health monitoring, drug delivery, and many other intra-corporal (within a corpus, the body) tasks need ultra-miniature sensors to fulfil their missions while miniaturizing invasiveness. As limitations in downscaling of conventional systems are foreseeable, new materials with new properties on the nanoscale will emerge to fulfil sensor tasks in ultra-miniaturized sensor systems.

**Complexities in protein investigation**

Study a protein is more complex than study an amino acid. Indeed, a protein can be described in three degrees of complexity. First, a protein is an assembly of amino acid. This is called the peptidic chain of the protein. It corresponds to the primary structure of the protein. Then, the peptidic chain forms what is called the secondary structure. This structure is the arrangement of the amino acid in helix ($\alpha$-helix), sheet ($\beta$-sheet), coil, turn ($\beta$-turn) ... Each arrangement is formed by many amino acid. Finally, the structuration of the secondary structures to form the final structure of the protein is called the tertiary structure.

This complex organization of the amino acid composing the protein induces a spatial distribution of the amino acid on the surface of the protein or, at the opposite, a radical chain can be hiddened inside the protein. Because of this complex form and the much localized effect observed in SERS, we expect a Raman signal stemming from only few neighboring amino acids.

Another difficulty in protein study concerns the different conformations that the protein can explore. These conformations are defined as different tertiary structures of a same protein. Even if multidomain proteins are most prone to this type of modifications, monodomain proteins can also be modified by the presence of an additional molecule (e.g. ATP in the case of HSP [1]) or a change of the pH of the solution [2, 3].

This last parameter is critical for proteins due to the electric potential of each amino acid which can rapidly change with some pH fluctuations.

All of these factors might explained why there are so many fluctuations between the SERS spectra presented in literature for BSA, as shown on Fig. 1.8. Indeed, the adsorption site and conditions are critical in protein SERS study.

## Societal aspect and expectations

Stress or Heat Shock Protein 70 (HSP70) accumulates in cells in response to many different sources of stress including anticancer treatment, and helps the cell to survive these otherwise lethal conditions. One group has recently demonstrated that tumorigenicity, metastatic potential and resistance to chemotherapeutical drugs all correlated with the amount of HSP70 expressed in cancer cells. Conversely, it has been shown by the same group that HSP70 depletion by siRNA or anti-sense strategies decreases the tumorigenicity of cancer cells and can even eradicate a tumor. More recently the group has selected short peptides called aptamers (8-13 amino acids) by their ability to interact with and inhibit the protective properties of HSP70. These aptamers sensitize cancer cells to die.

Clinical studies have demonstrated that HSP70 is expressed in cancer patient's intratumoral and/or in the serum [4]. This expression has been correlated with poor prognosis. Research studies have demonstrated that only cancer cells (but not normal cells) express HSP70 on the cell surface. Therefore, HSP70 is an interesting target to be tracked in the organism of a patient to diagnose cancer, monitor its progression as well as treatment efficacy. In practice, the higher the sensitivity of the detection is, the earlier the cancer can be identified and treated. Such higher sensitivity is not available today not only in clinical or point of care environment, but also at the level of oncology research institutes. Indeed, today diagnosis still relies mainly on microscopic (but not molecular) cues, when the tumor is already composed of several millions of cancer cells. For this reason, recent improvements in nanotechnology have spurred increasing interest in cancer detection. Early detection would benefit to the development of new cancer therapies based on peptide aptamers that could be delivered sooner and at lower doses. On a longer term, treatment follow-up provided by an efficient sensing method may enable a more personalized therapy, thereby minimizing the painful side effects of cancer therapy.

In vitro diagnostics is a key element of modern health care and is of increasing importance due to the need for more effective and efficient health care. At the moment, only 50% of diseases are correctly diagnosed in medical laboratories. One-third of all tests are presently performed in point-of-care (POC) settings outside centralized laboratories. The main focus of POC is to complement the efficient diagnosis of conventional medical laboratories that have centralized large and expensive equipment to

shorten the time of clinical decision on therapy, as well as to lower test costs, especially for monitoring purposes.

The POC instrument development currently aims at reduced size and cost to target the market of bedside and outpatient care, but also to become used in emergency centers close to the accident site. This also asks for an easy to use handling strategy to be applied by untrained people. A major objective of POC is the measurement of biomarkers, which play an important role for understanding the development of human diseases and for the identification of the risk for disease. Biological markers are used in Medicine to establish physiological states of an organism, or as indicators for particular disease states. The most traditional parameter used as marker may be the heartbeat, but also blood pressure and body temperature are used since a long time. Recent developments in Cell Biology led to the discovery of a broad variety of molecular biomarkers, which still can be classified from a medical point of view into prognostic and diagnostic markers, with the latter indicating probable effects of treatment and the others the individual disease development independent on the type of treatment. Molecular biomarkers are often peptides found only in very small quantities in blood samples or other body fluids, in the earlier stages of disease making it difficult to detect them before the patient develops severe symptoms.

Specific biomarkers are known in all medical fields: e.g. Hematology; Diabetes; Rheumatology; Therapeutic drug monitoring; Infections agents (HIV, mononucleosis and influenza); Sepsis (C-Reactive Protein, Procalcitonin).

The main tools in molecular diagnostics to detect biomarkers are so called assays measuring the presence of specific molecules, in low concentrations, identified as a marker of disease or risk. The markers used for the assessment can be either from the proteome or the genome, but they also can be hormones, lipids or sugars. The enzyme-linked immunosorbent assay (ELISA) is the most wide spread "wet-lab" type of solid-phase enzyme immunoassay. In 2012 immunoassays already have been a worldwide $17 billion industry at a growth rate of 2-3%. They are based on the use of specific antibodies against an "analyte", which in many cases is a protein antigen.

Recently, several POC strategies have been developed for operating either locally in emergency rooms or via telemedicine to improve cost-effectiveness and help to rectify the rural-urban mortality inequities. Most of the POC strategies for cardiovascular diseases rely on Molecular Biology based devices such as ELISA tests calibrated for identifying one or more out of ten known cardiomarkers. For each biomarker, the minimal level to be detected is below 200 pg/ml. These low levels cannot be measured with ELISA tests in less than 2 hours. This period is much too long undermining the relevance of molecular biology based detection in becoming widely accepted or applied in bedside testing. Hesitations in implementing POC tests may arise also from the
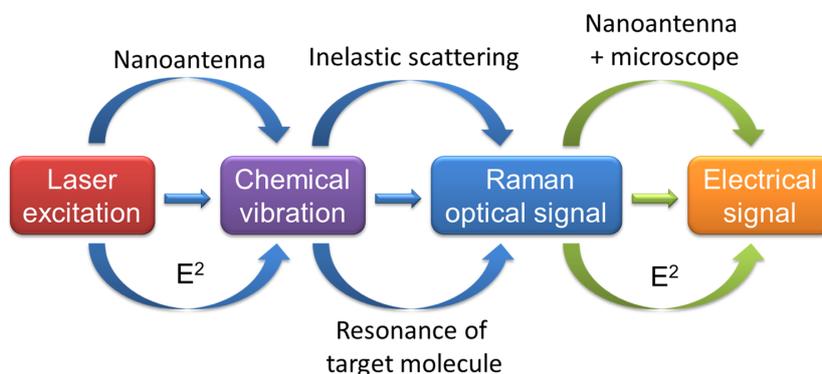
Figure 1.1 – Schematic description of the SERS phenomenon.

qualitative and semi quantitative nature of many POC tests or concerns about low-end accuracy compared to more sensitive laboratory tests.

### 1.2.2 The characteristics of a sensor

First let us enquire into the meanings of the terms sensors and transducers. The word "Sensor" is derived from the Greek word *sentire* meaning to perceive. It is a device that converts a physical stimulus such as mechanical motion, heat, light, sound, and magnetic or electric or radiant effect into an electrical signal, which is measured or recorded by an observer or an instrument. It is used for various purposes including measurement or information transfer.

Similarly, the word transducer has originated from the Greek word *transducere*, which means "lead across". It is a device usually electrical, electronic, electromechanical, electromagnetic, photonic, or photovoltaic that converts power from one system to another either in the same or different form. A transducer is a device that is actuated by energy from one system and supplies energy usually in another form to a second system.

What are the differences between a sensor and a transducer? A sensor differs from a transducer in that a transducer converts one form of energy into other form whereas a sensor converts the received signal into electrical form only. A sensor collects information from the real world. A transducer only converts energy from one form to another.

In our case, SERS is a mix between a sensor and a transducer. Indeed, the SERS phenomenon can be separated in three processes, as shown on Fig. 1.1. The blue processes are transducing and the green one is sensing. Firstly, it consists in the transfer of the optical energy coming from the laser excitation to the chemical vibration of the target molecule. This effect is enhanced by the nanoantenna with an efficiency in pro-

portion to $E^2$. The second process is the energy transfer from the chemical vibration to the emission of a Raman optical signal by simply inelastic scattering. This energy is proportional to the optical resonance of target molecule. These two first processes are clearly transducing effects. The last process is the conversion of the Raman optical signal in electrical signal by an Avalanche PhotoDiode or a spectrometer counting the emitted photons received through the nanoantenna and the microscope objective. This process is a sensing process consisting simply by a measure of the photon rate.

The idea will be also to increase the transducing efficiency and improve the sensing abilities of the system.

Many terms related to sensors will be repeatedly used in the text. It is therefore beneficial to briefly recapitulate the definitions of these terms, and look on SERS specificities.

**Sensitivity**: It is the change in output value per unit change in input variable (the measurand). In SERS, the mesurand are the wavenumber and the intensity. However, the choice of wavenumbers is non-negotiable because it depends of the setup and detectors. The only parameters involving sensitivity evolution is the sounded intensity. This intensity depends on the enhancement properties of the substrate and the collection efficiency of the sensor [5].

**Selectivity**: Ability to detect the measurand in the presence of similar species. It is a measure of the discriminating ability of the sensor with respect to interfering analytes, which are sources of noise in the output. In SERS, it is the ability to distinguish two species simply based on their spectral fingerprints. This involves a clear spectral analysis to sort spectra.

**Resolution**: It is the smallest measurable change in analyte (measurand) value that can be detected by the device. In SERS, it consists in the Limit Of Detection (LOD) of the system. This limit is defined as the smallest concentration of a specific molecule that can be detected in a reasonable response time. This LOD is one of the main limitations in classical SERS experiments [5].

**Response time**: It is the time taken by the sensor to reach $1/e$ or 63% of the final value of the sensed variable (output value). In SERS for biosensing, a reasonable response time is around ten minutes in hospital field. Unfortunately, many SERS experiments are based on spectral acquisitions around a minute, which involves a low statistical response and also reduce the reliability of the result. A reduction of the spectral acquisition time or of the Raman photon

detection time is so necessary.

**Calibration characteristic**: It is the curve obtained by plotting the sensor output along the abscissa and the measurand (analyte values) along the ordinate. In SERS, the reliability of such parameter is questioned. Indeed, this parameter involves the definition of a characteristic SERS spectrum for each molecule. But this point is not clear in literature comparing the different spectral fingerprint for a same molecule, as in the case of BSA for example [6, 7, 8, 9, 10, 11, 12].

**Linearity**: It is the degree to which the calibration curve of the sensor matches with a specified straight line approximating the same. In SERS, this can be translated as the link between concentration and Raman intensity.

**Repeatability**: It is the reproducibility of sensor output readings at given measurand values. One more time, in SERS, this involves the spectral repeatability for a same target molecule, which is not really clear.

**Stability**: It is the degree to which the calibration curve of the sensor remains unchanged over a period of time so that the sensor need not be recalibrated. In SERS, this can be directly associated to the temporal repeatability but also the spatial repeatability of the measurement.

**Drift**: it is the shift or: translation in calibration curve of the sensor with respect to time. In SERS, this drift can be associated to the photoluminescence effect. Indeed, the use of a metallic surface induces a photoluminescence of it. This one is very sensitive to the enhancement and can involve spectral modifications.

**Allowed ambient parameters**: These are permissible maximal values of ambient parameters such as temperature, pressure, relative humidity, light or illumination, etc. under which the device can operate satisfactorily. In SERS, the main parameters are temperature, laser power and flow rate of microfluidic setup and SERS active sample preparation.

## 1.2.3   Definition of a nanosensor

Nanomaterials have sizes comparable with those of biomolecules in proteins, virus (a small infectious agent), cells, nucleic acids (the building blocks of living organisms), etc., and thus can form easy interfaces between biomolecules and readout instruments. Equality or comparability of size-nanomaterials and biomolecules is exploited

in nanosensors. Many nanosensors are based on the interfacing of nanomaterials with biomolecules. These are known as nanobiosensors.

Sensor miniaturization and device integration, based on reproducible fabrication processes and large-scale production, are the top prerequisites low-cost products, and these requirements are fulfilled by nanosensors so that research on nanosensors will lead to cheaper devices.

Nanosensors allow for building an entirely new class of devices that provide the elemental base for "intelligent sensors" capable of data processing, storage, and analysis. These sensors will provide high accuracy, ultrahigh sensitivity, extreme specificity, real-time in vitro information with greater speed, having multi-analyte options, requiring smaller quantity of sample and minimal sample preparation, durable, safe, and portable. Some members of the nanosensors generation of devices have already qualified in laboratory tests, and are starting to appear in the marketplace. But there remains a long way to tread and several issues to be solved. Nonetheless, the potential is enormous.

**What is a nanosensor?**

Definition of nanosensors must be clear. In fact, any sensor that uses a nanoscale phenomenon for its operation is a nanosensor. Any sensor fabricated by nanotechnological methods is a nanosensor, that is, nanosensor is a nanotechnology-based sensor. To define the term, let us agree that any sensor characterized by one of the following properties will be labelled as a nanosensor:

(i) the size of the sensor is in nanoscale: this is answer by the use of nanoparticles or lithographed nanostructures.
(ii) the sensitivity of the sensor is in the nanoscale: the sounded volumes and concentrations are closed to nanogram or nanomolar, nay single molecule.
(iii) the spatial interaction distance between the sensor and the object is in nanometers.

From definition (i), nanosensors involve signal transformation from the environment using nanostructures, that is, structures having at least one dimension in the lateral direction less than 100 nm. Hence, in this definition, sensors must comply with nanotechnology based on geometrical dimensions of the sensor.

At present, the most commonly observed nanosensors exist in nature in the biological world as natural receptors of outside stimulation. Animals like dogs have a strong sense of smell that functions using receptors that feel nanosized molecules. Various fishes use nanosensors to perceive minuscule vibrations in the surrounding water; several insects detect sex pheromones using nanosensors. Similar to animals, many plants also use nanosensors to detect sunlight. In the artificial world, most film cameras have

used photosensors at the nanosize for years. Traditional photographic film uses a layer of silver ions that become excited by solar energy and cluster into groups, as small as four atoms apiece in some cases, scattering light and appearing dark on the frame. Various other types of films can be made using a similar process to detect other specific wavelengths such as IR, ultraviolet, and x-rays.

**Hopes in nanosensing, and particularly nanobiosensing**

We have shown that many sensors are based upon the adsorption of target analytes and adsorption depends on surface area and surface chemistry. This chemical functionalization enables a high selectivity but involves very specific chemical treatment of the surface to be associated to the target analyte and also limit the number of able sounded analytes. Thus, systems like ELISA or Biacore operate on this chemical functionalization. They are very sensitive, very selective but not universal. The idea is so to use the physical approach based on spectroscopy, combined with the optical approach used in Biacore and other Surface Plasmon Resonance systems allowing high sensitivity with an optical, and so non invasive technique.

Different optical spectroscopic methods applied on biomolecules exist. Thus, we should have infrared spectroscopy based on the adsorption of the infrared light by the biomolecules sounded. Unfortunately, water is very absorbent in infrared and water is the principal component of biological entities as cells. Moreover, infrared source for spectroscopy are very expensive. Another spectroscopy is the Raman spectroscopy. The interest of this technique comes particularly from its wavelength window in visible range in which sources are cheap and water, the main solvent in biology, is very low absorbent. However, Raman scattering is a rare effect and has to be enhanced to be reliable.

In ELISA or Biacore, the sensitivity is very high but is based on large number of molecules. Thus, the single molecule resolution cannot be reached with such methods. The dream is so to have a technique at sensitive enough to detect a single molecule. The non-invasivity of this sensor seems one more time critical to ensure such sensitivity.

## 1.2.4   Nanobiosensing: what do we really expect?

To summarize, we have shown that in nanobiosensing the main expectations are based on a high selectivity to sort and identify proteins, amino acids, cells, viruses ..., and a high sensitivity to reduce the concentrations and also the size of target analytes. In this aim, the most interesting techniques seem to be the optical spectroscopies. However, the main limitation of optical sensors comes from the low sensitivity of these techniques. The is the reason why we chose to develop molecular plasmonic sensing techniques adapted to spectroscopies, and more precisely the Surface Enhanced Raman

Spectroscopy. The limits of this technique in sensing have been previously summarized but we will see how scientific developments could answer to these limitations.

## 1.3    Molecular spectroscopy

Different types of nanosensing are proposed based on molecular plasmonic. These are called Surface Enhanced techniques. We can so find the Surface Enhanced Fluorescence, in which the fluorescence signal of a fluorescent molecule is enhanced. This allows reducing the concentration of fluorophore necessary to obtain a good picture with a high SNR. However, the target molecule should be labeled with a fluorophore, or should be intrinsically fluorescent.

An alternative is to use the Raman signal emitted by the molecule probed after its excitation with a monochromatic light source, such a laser. This is called Surface Enhanced Raman. Different studies can be made. First, by filtering a specific wavenumber previously determined corresponding to the target molecule, a mapping of its spatial localization in time or its distribution is possible [13, 14, 15, 16, 17, 18]. This is called SERImaging. Second, by scattering the Raman signal also enhanced with a spectrometer, the characteristic spectrum of the probed molecule is obtained. This is called SERSpectroscopy. It is the effect that we will use in our experiments.

### 1.3.1    Raman spectroscopy

In SERS, the studied signal is the Raman signal of the probed molecule. The Raman signal is the result of a non-linear interaction between the incident light, most of the time a monochromatic light, with the probed molecule (see Fig. 1.2).

The incident light is absorbed by the molecule to achieve a virtual state of energy of this molecule. The dropout of the molecule induces the emission of a photon. The wavelength of this photon is most of the time the same than the incident light. The phenomenon is so called Rayleigh scattering. We talk about scattering because the speed of the phenomenon is close to an instant effect. But rarely, the emitted photon is the result of an inelastic scattering. This is called the Raman effect.

Two cases should happen. The first, which is the most favorable, is called Stokes case. It corresponds to the excitation of the fundamental state of energy of the molecule with the incident light and the emission of a photon to achieve an upper state of energy corresponding to a characteristic vibration state of the probed molecule. The difference between the incident wavelength and the emitted one is called Raman Shift and given in $cm^{-1}$. The shift obtained in this case is positive.
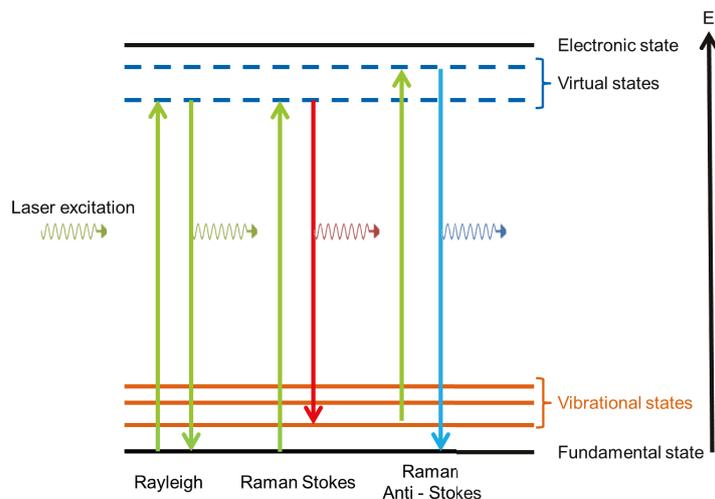
Figure 1.2 – Schematic description of the Raman effect.

The second case, less probable, is called Anti-Stokes. In this one, the molecule is still on a vibration state before the excitation by the incident light. Then, the emitted photon allows achieving the fundamental state of the molecule or a lower state of vibration. In this case, the Raman shift is negative.

A molecule can also be associated to a characteristic Raman spectrum. In fact, each vibration is associated to different vibration states that can be sounding, or not (depending of the selections rules), by Raman spectroscopy. Unfortunately, the Raman effect is a very rare effect. Only very few photons are Raman's photons in a flow of Rayleigh's. And this is accentuated for non-resonant molecules, such are most of the biological molecules. This is the reason why the use of an antenna, and more specifically the Localized Surface Plasmon Resonance of a metallic antenna, is necessary to enhance this Raman signal, thereby increasing the probability of sounding this signal.

## 1.3.2   Localized Surface Plasmon Resonance

New characterization techniques based on nanosciences, and more specifically a well-known effect at this scale: the Localized Surface Plasmon Resonance (LSPR) based on plasmonic has emerged. In this part, I will show some details about such nanosensors developed on the basis of LSPR.

The sensing of nanoscale elements implies the use of a localized effect. The main interest of such sensing method is the work on very low quantity of analyte. One of the objectives is also the determination of the different components of a molecular mix without, or limited, external impact. That is why the development of non-destructive
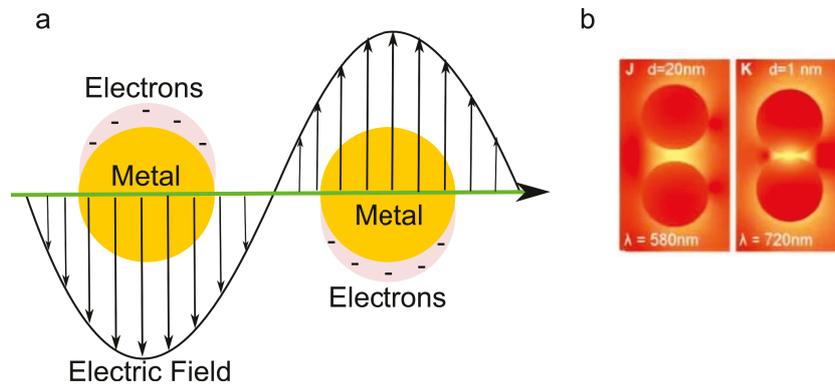
26

Figure 1.3 – (a) Schematic view of a Localized Surface Plasmon Polariton. The coupling between the electric field and the metallic nanoparticles induces collective oscillations of electrons of particles. (b) Near -Field maps for two neighboring gold spheres of radius of 60nm. Adapted from [19].

nanosensors based on optical effect is still effective. The so-used localized optical effect is called Localized Surface Plasmon Resonance.

A Localized Surface Plasmon results from the electronic vibration of metallic nanoparticles as shown on Fig. 1.3. This vibration induces a dipole to which one can associate a resonating frequency. By coupling at least two nanoparticles at a close distance around few nanometers, the plasmon generated is localized within the gap between the nanoparticles. The resonance frequency depends on the nanoparticles positions, their shape and also $\varepsilon$, the dielectric permittivity: the roughness, the distance between both particles, the metallic compounds,... The changes in the shape and $\varepsilon$ of the nanoparticles induce a shift in resonance. This shift is used to detect the shape and $\varepsilon$ modifications, as in SPR.

The coupling of two nanoparticles or nanostructures closed one to the other is called a nanoantenna. In fact, the ElectroMagnetic (EM) Field exciting the LSPR of these nanoparticles is enhanced between both nanostructures in the gap where the plasmon is localized. This enhancement improves the sensibility used in nanosensing.

## 1.3.3   SERS: Surface Enhanced Raman Spectroscopy

The coupling between Raman spectroscopy and high Electromagnetic field local enhancement by LSPR is called SERS, for Surface Enhanced Raman Spectroscopy. Discovered by Jeanmaire and Van Duyne in 1977 [20], the topic SERS is experiencing a renewed interest since ten years due to its potentialities for molecular sensing (Fig. 1.4).

One of the characteristic of SERS is in the definition of an Enhancement Factor
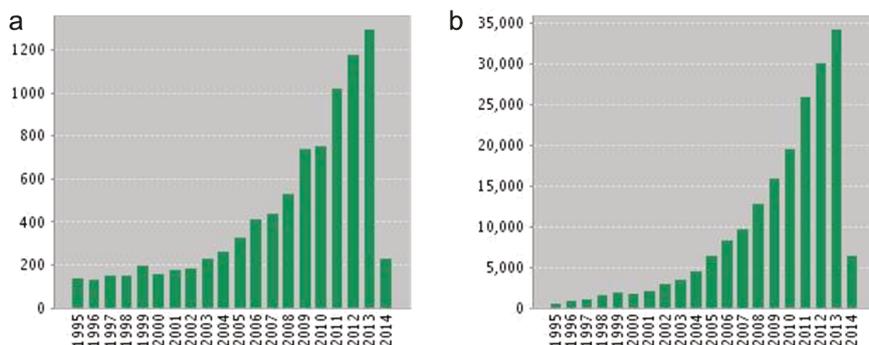
Figure 1.4 – Bibliographic statistics on (a) published items per year and (b)
citations per year. Statistics provided by Web of Knowledge for "SERS"
as entry.

(EF) of the Electromagnetic field. In fact, this local enhancement of the EM field,
induced by the LSPR of a metallic nanoantenna, leads the enhancement of the Raman
scattering of a molecule very close to the hot EM spot. The EF is decisive for detecting
diluted molecules.

**Enhancement Factor Determination**

The impact of the presence of the nanoantenna exalting the Raman signal can be
quantified using the Enhancement Factor EF defined by many ways. Definition has
been first given by P.Hildebrandt [21] but the main work on this topic has been done
by E.LeRu and P.Etchegoin [22]. They have demonstrated that different definitions
could be given for the enhancement factor. The most intuitive and easy to determine
is the one named Average Enhancement Factor (AEF) defined by:

$$AEF = \frac{I_{SERS}/C_{SERS}}{I_{Raman}/C_{Raman}} \tag{1.1}$$

Where $I_{SERS}$ and $I_{Raman}$ are respectively the spectral intensity in SERS and classical
Raman configurations, and $C_{SERS}$ and $C_{Raman}$ are respectively the concentrations of
the solution sounded in SERS and classical Raman configurations.

This AEF is easy to determine in classic SERS experiments, but it is difficult to
determine the $C_{SERS}$. In fact, this concentration should be, by definition, the con-
centration sounded by SERS. In the Raman case, the concentration is the one of the
solution, but in SERS case, the concentration is localized around the hot spot. In the
case of colloids in solution, it could be relatively simple to determine this concentration
contrary to the surface concentration. This definition is not so good, and the problem
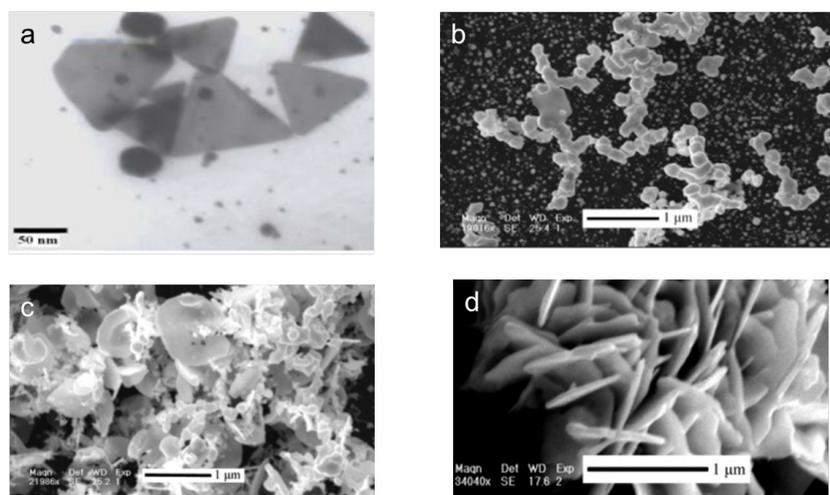is even more complicated in the case of single molecule.

Figure 1.5 – (a) TEM image of gold nanotriangles. Adapted from [23]. (b-d)
SEM images of gold nanoparticles. Adapted from [24]. (b) Nanowires,
(c) Nanosheets and (d) Nanoflowers.

Other definitions are purposed in the single molecule case [22] but are very hard
to implement experimentally because take account of the Raman cross section of the
target molecule.

### 1.3.4   Plasmonic substrates

One of the main limitations of the development of LSPR sensors for the industry is
the reproducibility of low cost and very sensitive substrates. Two main approaches have
been developed. I will present in this part the bottom-up and top-down approaches
applied to SERS substrates.

**Bottom-up approach: Metallic colloidal nanoparticles**

In most of publications, the plasmonic capabilities of metallic coupled nanoparticles
are the basis of SERS effect. These particles are made out of precious metals like gold
or silver. The interest of colloids lies in ease of preparation. In fact, nanoparticles with
various shapes [23, 24] can be produced in large quantities, as shown on Fig. 1.5.

The different shapes are interesting to investigate the spectral enhancement.

Unfortunately, such nanoparticles in a colloidal solution are subject to aggregation.
Particles are thus chemically coated with different additives. These functionalizations
change the spectral signature [25]. Raman spectrum of the probe molecule may be
then altered. An alternative is to functionalize the particles with a specific ligand.
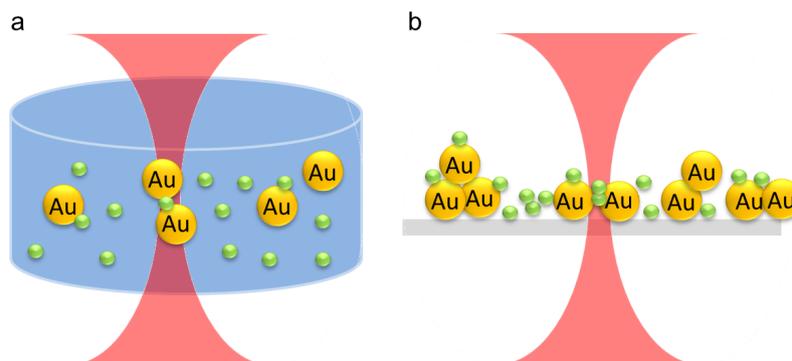
Figure 1.6 – Schematic description of the different configurations for SERS excitation.(a) Liquid case and (b) Solid case.

The media of metallic colloids can be solid or liquid (see Fig. 1.6). In the liquid state, the particles are dispersed in a solvent allowing their dissolution and the target dissolution. The SERS signal is obtained by the excitation of the particles closed one to another. In this case, a control of the homogeneity of the solution is critical [26]. In the solid case, the particles are deposited onto a surface (glass or silicium in most cases). The solvent of the solution is evaporated to stay with the nanoparticles aggregated, or not, depending of the presence/absence of the functionalization. A SERS signal can so be measured by illuminating the aggregates [27].

**Top-down approach: Lithography**

The other approach is a method based mostly on e-beam lithography or etching. The substrate is obtained by structuring a bulk film to obtain the nanostructures supporting the LSPR. These structures are also bare due to the evaporation or etching of the metal.

Different shapes can be designed to modify the orientation of the EM field. The nanostructures are designed with a spatial resolution closed to ten nanometers (aligned nanorods (as on Fig. 1.7) but also honeycombs). This type of nanostructure is relatively easy to model by FEM or FDTD calculations contrary to chemically synthetized nanoparticles with more complex shape. Such calculations highlight the plasmon resonance (see Fig. 1.7 (b-c)). The resonance is mostly a function of the gap between the structures. The major interest of this type of structure is the reproducibility.

Unfortunately, the gap between the nanostructures should be very small, and so very hard to achieve with electron beam lithography. An alternative is the use of Reactive Ion Etching [29, 30, 31, 32, 33] or the use of chemical trick as the oxidation of a chromium layer [28]. The other drawback of the top-down approach is the high cost generated due to the waste of time, loss of matter and the need of expensive
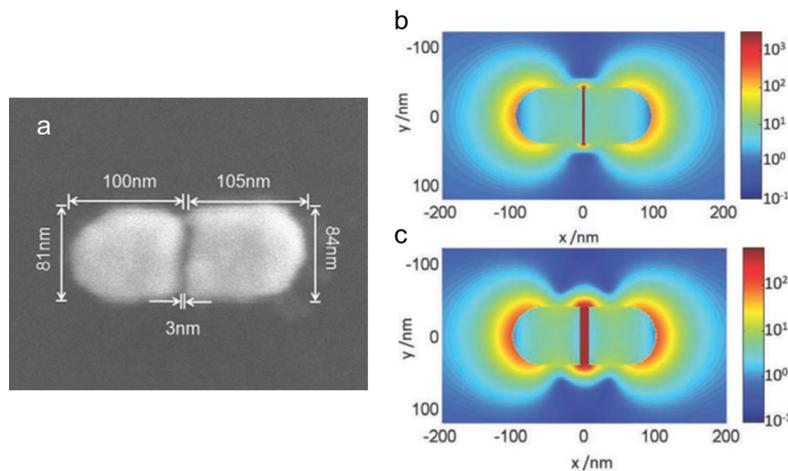
Figure 1.7 – (a) Lithographed Gold nanorods observed by SEM with a 42k x magnification. (b-c) FDTD calculation of the steady state field intensity distribution for dimer structure with 3 nm gaps (b) or 12 nm gaps (c). Adapted from [28].

equipment.

## 1.4  SERS based sensing

Historically, SERS has been demonstrated as the enhancement of the magnitude of the Raman scattering signal when a scattering molecule is placed on or near a roughened noble-metal substrate [20]. Since this discovery, the technique has progressed from model system studies of pyridine on a roughened silver electrode, state-of-the-art surface science to finally real-world sensors applications.

The development of SERS based sensor involves question of the reproducibility, but also of the stability and ability to be adapted for biosensing or the question of sensitivity.

### 1.4.1  Reproducibility over large area

Concerning the reproducibility, an approach allows answering this question: working on large area. This approach results from the use of top-down substrate to generate SERS signal. Indeed, the spatial resolution reliable in electron beam lithography with a high reproducibility (around 10 nm) is not good enough for an efficient enhancement with a single nanostructure.

The choice is so to broaden the studied area to multiply the number of hot spots

and increase the Signal to Noise Ratio (SNR). This broadening can be applied temporally, by a simple increase of the acquisition time, but also spatially. Thus, different solutions are proposed. Firstly, the aggregation of nanoparticle on a surface is a good solution to multiply the number of hot spot [21, 34, 35, 36]. However, the reproducibility on such experimental conditions is reduced because sizes of the aggregates are not well controlled. The second solution is the fabrication of metallic arrays by lithography and reactive ion etching [31, 37]. This improves the reproducibility but lower cross-section due to a large expansion of the hot spots distribution.

In such configurations, the definition of an Average Enhancement Factor is the main parameter to characterize the SERS substrate. The AEFs are so distributed between $10^3$ and $10^6$. This is good enough for the detection of resonant molecules as dyes but, concerning the biomolecules, it is still limited because of the low SERS cross-section of such molecules (around $10^{-14}$).

## 1.4.2   Biological sensing

The concept of sensor is more difficult for the study of non-resonant molecules. In fact, the spectral reproducibility of low concentrated solution of resonant molecule seems coming from this resonance effect. A large part of the non-resonant molecules corresponds to biomolecules such amino-acids, peptides, proteins ... These molecules are interesting as target for early stage detection of many diseases for which these molecules are the molecules of interest.

Many experiments in SERS were performed on such biomolecules by tagging the biomolecule with a resonant marker as developed as [38]. The analysis of the spectral modification of the biofunctionalization previously fixed on the metallic nanostructures was developed in [39]. Finally, few SERS experiments on biomolecules were conducted as we can call "Label-Free", i.e. without specific functionalization or tag, as developed in [6] or  [7]. This point of view was clearly the most interesting for us since it does not need any special previous chemical development of specific linker for the target molecule.

However, in such configuration, SERS seems not so reproducible compared to resonant molecules. Indeed, many experiments have been done on BSA (Bovine Serum Albumin), one of the most prevalent protein in bovine (more than 70 % of serum composition) with different substrates. And their results in different spectra for a same molecule at high concentration (mM), as shown on Fig. 1.8.

This is coherent with the main problem of SERS on non-resonant molecules: the signal recorded for such molecules is very sensitive to the surface effect like adsorption, resulting in a spectrum composed by the closest vibrations to the surface. This is
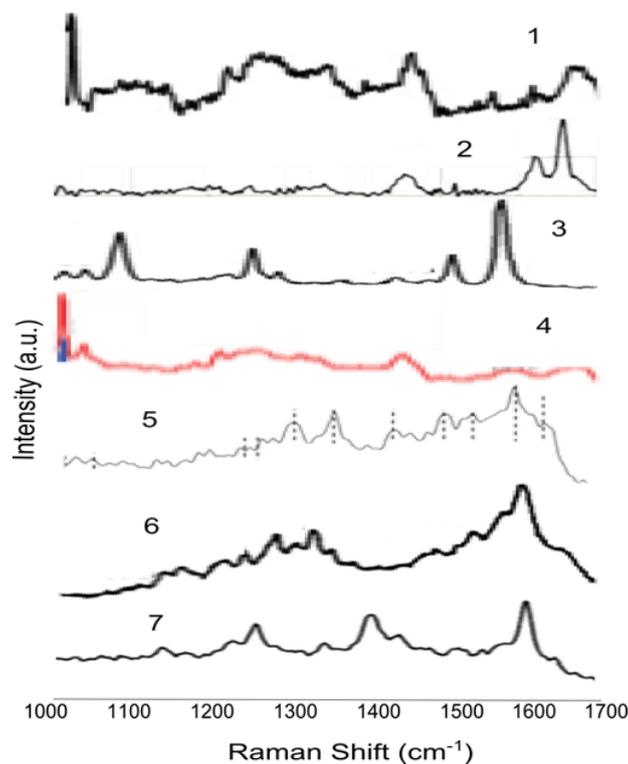
Figure 1.8 – Bovine Serum Albumin (BSA) SERS spectra adapted from literature and obtained on different SERS substrates. (1) Adapted from [6], (2) adapted from [7], (3) adapted from [8], (4) adapted from [9], (5) adapted from [10], (6) adapted from [11], and (7) adapted from [12].

not impacting for resonant molecules because the effect of the resonance reduces this impact to the spectral noise in this specific case. Thus, the differences between the spectra of a non-resonant molecule come from the different sites of adsorption of the molecule on the metallic surface. It becomes difficult to associate a non-resonant molecule with a SERS spectrum because the spectrum will depend of the electrochemical state of the metallic surface.

### 1.4.3 Single molecule regime

In single molecule regime, it is important to look on the temporal fluctuations of Raman spectra in surface enhancement circumstances. Working under flow or in air, the spectra record at different times but in the same experimental conditions appear and disappear during the slot of acquisition. These fluctuations append also for the study of only one molecule. The fluctuations could be simply in terms of intensities ratios between the different spectra or in terms of Raman bands, i.e. Raman peaks positions.
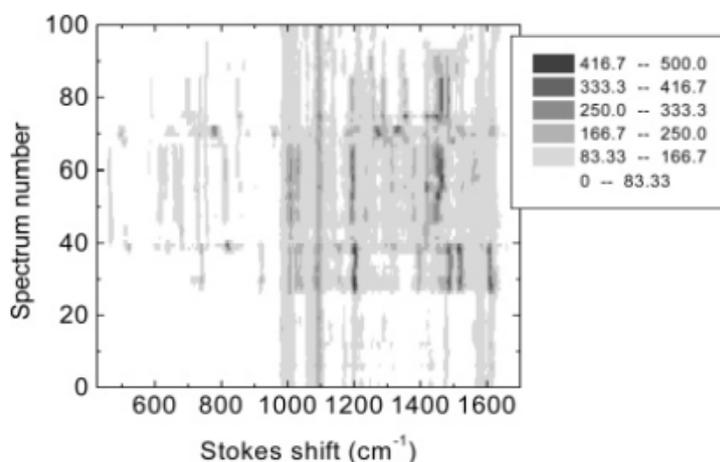
Figure 1.9 – Time evolution of the Raman spectrum of a single 4-mercaptopyridine molecule on silver. Adapted from [40].

In our experiments, if we look in terms of time acquisition, we observe the same effect, as shown in Fig. 1.9. Indeed, it is clearly proved by this experiment that in single molecule regime appear a lot of spectral fluctuations in time. The question of the reproducibility is legitimate, due to these times fluctuations.

Since 90's these time fluctuations have been explained by single molecule regimes [41, 42]. Kneipp *et al.* confirmed there analysis based on Poisson statistics describing the intensity distribution of the fluctuations. To obtain such resolution, a very high EF is necessary. Thus, Kneipp *et al.* predicted that an EF close to $10^{14}$ is necessary to be in single molecule regime for dye molecules.

Le Ru *et al.* detailed in 2007 [22] a complete study of the enhancement factor and its implications in single molecule regime. A recent paper of the same group [43] explained that a minimum enhancement factor in single molecule regime between $10^6 - 10^8$ is required.

Different approaches are so proposed in literature to obtain the SERS single molecule regime based on colloids:

- First, the easiest method consists of mix a solution of target analyte with a colloidal solution in a drop. SERS events appears when the different particles are governed by the Brownian motion in the drop interacting at very close distance between two nanoparticles and one target molecule [44]. The problem of such method is in the low probability of triple interaction exactly where the laser is focused. To increase this probability, the use of a less focused laser or a higher target analytes concentration is necessary but clearly reduces the resolution.
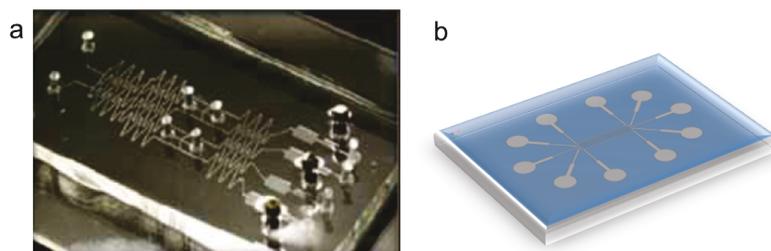
34

Figure 1.10 – (a) Microfluidics picture with complex channel design. Adapted from [47]. (b) Scheme of our microfluidic platform.

Another limitation of this technique is in the evaporation of the drop that is low controlled.

- Secondly, in order to increase the probability of interaction and drastically reduce the evaporation, the use of a microfluidic chip in Y configuration was proposed [12, 45, 46]. The interest in this method stays in the high probability of interaction between colloids and target molecules. Like in the previous method, the main limitation of this method is the indispensable chemical coating on the nanoparticular surfaces. This coating is mandatory to avoid the nanoparticles aggregation. Fluctuations come also simultaneously from the motion of both target molecules and colloidal particles. This induces a low repeatability of the EF due to constant modifications of the nature of the hot spot.

- Finally, a good repeatability is ensured using colloids deposited on glass or silicon surface capped by a microfluidic chip. Target analytes are then injected in microfluidic channels. The only fluctuating origins come from the motion of target analytes in the SERS hot spot, and this hot spot can be well localized on the surface for further measurements.

### 1.4.4 SERS and microfluidic

Benefiting from the recent advances in microtechnology, the use of microfluidic devices to conduct biomedical research and create clinically useful technologies has a number of significant advantages. Instead of operating in air (e.g.: DNA chips), measurements are performed in liquids which is obviously more biologically relevant. Besides, the volume of fluids within the channels is very small (usually several nanoliters) which is 6 orders of magnitudes less than ELISA tests. As a consequence, the amounts of reagents and analytes used is reduced, a significant improvement for expensive reagents such as aptamers and/or specifically designed antibodies. In addition, the fabrication techniques used to construct microfluidic devices are relatively inexpensive and are very amenable both to highly elaborate, multiplexed devices and also to mass production. In a manner similar to that for microelectronics, microfluidics

technologies enable the fabrication of highly integrated devices for performing several different functions on the same substrate chip.

In the context of a sensing platform, three dependent parameters can be optimized: flow rates, volumes, and the diffusion distance above the sensing area. Flow rates can be finely tuned using three approaches, two active and one passive. The passive approach to tuning flow rates simply involves tuning flow resistance and flow pressure to regulate flow velocities. Using standard fabrication approaches, flow rates can be tuned over several decades starting at zero flow rates and capping at high flow-rates with velocities of tens of cm per second. Active control involves the use of integrated micromechanical valves, which allow flow to be shunted at will in a microfluidic channel. One approach involves using intermittent, rather than continuous flow by closing and opening a single valve at defined frequencies. Valves can be actuated at frequencies of up to 50Hz and the open and closed state of a valve can be controlled with millisecond precision. A more sophisticated active control makes use of a peristaltic microfluidic pump, which can precisely control pL flow volumes. Finally, the device has to be designed in such a way that an analyte molecule will come into contact with the sensing area at least once (optimally more than once) to allow efficient capture and detection. Obviously a slow flow rate would permit this. But as discussed previously, low flow rates directly conflict with the requirement for high flow volumes. Another possible design feature to increase contact probability during transit would be to reduce the channel height above the sensing area. Fluidic channels can easily be fabricated to be 1-2 micrometers in height, drastically reducing the likelihood that an analyte will flow by the detector without coming in direct physical contact with it.

In most cases, the microfluidic chips are based on PolyDiMethylSiloxane (PDMS), but can also be of glass. The advantage of PDMS is its ease to be used compared with glass but some PDMS residues should pollute the studied solution. The bonding of PDMS with a glass or silicon substrate is simply realized using oxygen plasma activating the substrate surface to react with the activated surface of the PDMS chip.

Some substrates using microfluidic are developed to sort the molecules in a mix by weight or size with small serpentines (Fig. 1.10 (a)). Other microfluidics chip have been created to control the mix of two separates solutions and study this mix knowing the ratios of each solution for different positions in the channel [48].

## 1.5   Our choices

Now that the state of the art of SERS-based biosensors is exposed, we can justify strategic choices in our experimental approach:

1. the orientation towards the single molecule detection using high efficiency raspberry-
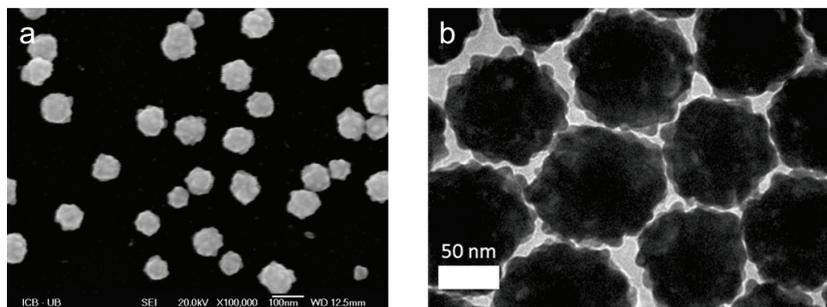
Figure 1.11 – Chemically synthesized nanoraspberry observed in SEM (a) or TEM (b).

like GNPs.

2. the gradual approach in the choice of molecules from resonant molecules such as dye towards amino acid and then proteins.

3. the most innovative breakthrough arises from the data analysis. The statistical analysis was a significant milestone for interpreting the huge amount of data.

### 1.5.1 Substrate and microfluidic

Among all the host of possible SERS active template, we have identified three main priorities.

The substrate should be first economic. The approach based on electron beam lithography was thus excluded from benefit of GNPs. GNPs offer also the advantage of being more efficient in terms of enhancement (up to $10^{12}$). Single molecules were detected historically using GNPs.

However the mobility of GNPs in colloid represents an important source of noise and fluctuations.

In this context, we opted to the most SERS efficient GNPs know at the beginning of my work: the Chebyshev particles. Roughness of GNPs creates a surface texture with bumps. Those defects which seem theoretically difficult to accommodate generate extremely hot spot at the surface. The surface texture of GNPs can generate many shapes of the particle such as nanoflowers, nanoraspberries, nanotriangles ... Our choices burns out to nanoraspberry-like GNPS, the radius of bumps (5 nm) being close to the size of proteins, as shown on Fig. 1.11.

In order to limit as low as possible the temporal fluctuation, we develop a chemical rate for immobilizing the GNPs on a flat substrate, so that the gap between GNPs is

fixed. GNPs are then capped by a microfluidic channel. This configuration helps to ensure the "chemical free" approach. Bare GNPs can probe directly the biomolecules driven by the microfluidic avoiding pollutant from the atmosphere. The assembly, the preparation and the manipulation will be described in Chapter 2.

## 1.5.2   Gradual analysis of molecules of interest

We believe that a gradual approach in the complexity of the molecular structure should be favored for a comprehensive understanding of spectra of single molecule. Resonant molecules (Methylene Blue) have been used to test the sensitivity and the reproducibility of the experimental approach, the SERS substrate, the optical bench of measurement and the microfluidic. Such approach is detailed in Chapter 3 where we discuss the limits of detection in terms of concentration.

The selectivity and the ability of the sensor for sorting the molecules have been demonstrated with more challenging molecules. Amino acids are not all resonant in visible range. The spectral identification of the cysteine, as an example, is presented in Chapter 4.

In Chapter 5, we dealt with the protein identification. We will first examine the "standard" one: the BSA known to be particularly robust over pH and temperature but which gives controversial SERS spectra. We ultimately present the protein of interest, the cancer biomarker HSP-70. Tracking the changes in conformation of this protein constitutes already a major challenge. Finding this biomarker in the whole blood is a real difficulty and an extra greater challenge posed by the European Project FP7 "SPEDOC".

## 1.5.3   New SERS analysis tools

As presented before, temporal fluctuations are intrinsically linked to the single molecule regime of SERS. To rise up the reproducibility of this technique and better understand the role of a molecule in SERS process, it was important to develop a routine analysis method that might take all the spectra into account without losing information. Since the averaging of the signal was clearly not appropriate, we have scanned the different statistical approaches from the simplistic to the most comprehensive. These statistical tools are, for some, based on intensity fluctuations, and another one on multiplexing variations of the intensities associated to the different wavenumbers. Described in Chapter 2, we applied these methods along the manuscript to discuss the benefits and limits.

Simply based on the intensity fluctuations we first calculate the normalized variance of the intensity for each wavenumber in dynamic experimental conditions. The Mandel

Factor is then a first indicator of the most fluctuating Raman bands.

Another method based on the Probability Density Function has been proposed to estimate the occurrence of spectra.

The last track of reflection is a multivariate analysis called Principal Component Analysis. This method is currently used to compare different SERS spectra resulting from different molecules [6, 49]. The definition as a multivariate method implies that this method is based on the intensity fluctuations of each wavenumber but taking account of the fluctuations for all wavenumbers simultaneously.

## 1.6    Conclusion

As shown in Fig. 1.4, studies based on SERS are growing. This is due to the high interests of this technique coupling a high efficiency, due to the enhancement of the signal, with a clear specificity and a high sensitivity. A lot of efforts concern the plasmonic substrate developments to answer to each wanted specificities. Unfortunately, the SERS physical processes are not clearly well-known. Some people talk about a chemical enhancement effect [43, 50, 51, 52] based on the chemical interaction between the molecule of interest and the metallic substrate, as others talk about a physical effect simply based on polarizability of the molecule of interest in the electromagnetic hot spot, because of the classical definition of the Raman effect as a polarization-dependent effect.

In this thesis, the objectives will be to better understand the SERS effect, to apply this effect to understand the molecule-metallic surface interactions and to develop experimental and analysis tools to ameliorate the reproducibility, the sensitivity, the response time and the resolution of SERS sensors.

Dynamic characterization of a single molecule by SERS

## 2.1 Introduction

This chapter presents the development of a specific experimental setup including its specific statistical analysis to be applied for the dynamic characterization of single molecules by SERS. This includes first the preparation of the SERS active substrate; secondly, how we have incorporated the fluid handling using microfluidics device and how we assemble the various optical modules for adapting a confocal microscope to a Raman spectrometer for dynamic studies; thirdly, which statistical methods have been applied to analyze this amount of data.

## 2.2 SERS-active substrate

As discussed in the state of the art, we have chosen to work with gold nanoparticle (GNP) based substrate to increase the enhancement performance. Here I detail the chemical synthesis of our raspberry like GNPs to control the size, the shape and the surface state. Once the GNPs prepared, we immobilize these particles onto a glass surface. Here, I give some recommendations that can help you to get a homogenous chemical self-assembling as a monolayer with an inter-particular distance below 5 nm.

### 2.2.1 Raspberry-like GNPs synthesis

The synthesis is adapted from [53] and is processed with absolutely clean glasses and nanofiltered water. The first step consists of the reduction of gold salt ($HAuCl_4.3H_2O$, Molecular Weight = 393.83 g/mol) in aqueous solvent by a zwitterion at pH 7.4.
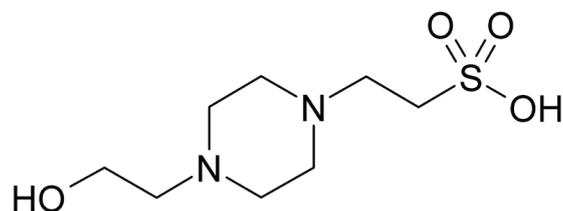
Figure 2.1 – Chemical scheme of HEPES molecule. Free radicals on the nitrogens of piperazine cycle allow the $Au^{3+}$ reduction.
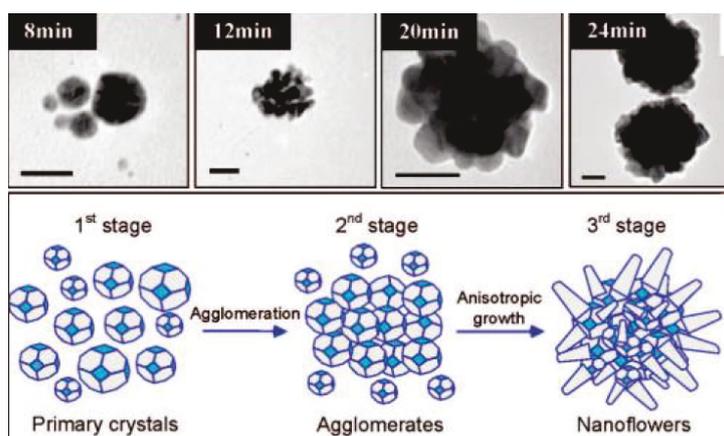


Figure 2.2 – Schematic illustration of the hypothetic growth mechanism for Au nanoflower formation in HEPES solution [53]. Smaller gold crystals are formed first and then tend to assemble into larger clusters, which can later grow with shaper surface texture.

The chosen zwitterion is HEPES (4-(2-hydoxyethyl)-1-piperazin ethane sulfonic acid, Molecular Weight = 238.3 g/mol), chemically schematized on Fig. 2.1.

This step can be declined in 4 points:

1. A mother solution of gold salt at 0.25M (0.493 mg/ 5 mL) is prepared. This solution can be stored in the freezer for many months.

2. An aqueous solution of HEPES at 0.1M (2.38 g/ 100 mL) is prepared and the pH is adjusted at 7.4 by just adding 1 mL of NaOH 1M and completing to 100 mL.

3. In a 500 mL beaker, 500 $\mu$L of gold salt solution are diluted in 400 mL of water under stirring.

4. 25 mL of HEPES solution are added and stirring during at least 2 hours at room temperature. The reaction is schematized on Fig. 2.2.

In terms of colors, the solution turns from pale yellow to uncolored with add of HEPES. It stays uncolored during approximately 30 minutes before becoming progres-
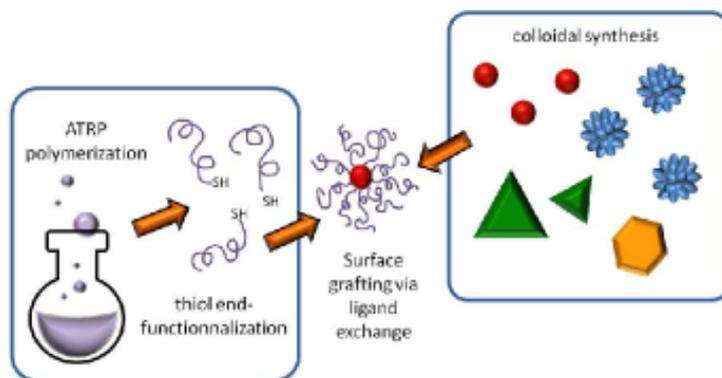
Figure 2.3 – Strategy of synthesis of the PS-grafted gold nanoparticles, further described in [54]. The highly efficient ligand exchange process can be applied to colloidal metallic particles of different size, shape and composition.

sively pale lavender and finally ocean blue in transmission and brownish in scattering.

The colloidal solution is centrifuged at 13,000 rpm during 1 or 2 minutes, depending of the centrifuge. The supernatant is removed with a Pasteur pipet. It is easily recognizable because of its pink-purplish color due to the little spherical particles. The tube is then filled in with previous solution keeping the concentrated part. And this tube is centrifuged one more time, and so on.

When all of the colloidal solution has been centrifuged, the concentrated part is rinsed with water and dispersed with ultrasonic waves to clean off the excess of HEPES. This step is repeated if necessary. Unfortunately, before the functionalization, the solution is less stable and can be stored no more than one week in fridge before rinsing.

## 2.2.2   Surface functionalization of GNPs with PS-SH

This work is based on [54]. The principle of functionalization of colloidal particles by a polymeric chain is well-schematized in Fig. 2.3. In our case, GNPs are functionalized with Polystyrene (PS) chains having thiol-end (SH). It is worthnoting that PS-SH is very hard to obtain but the Post-Doc working on SPEDOC project has synthetized some of it during her thesis. The functionalization is spontaneous at room temperature when the polymeric chains and the nanoparticles are mixed in solution. The main problem is in the very low solubility of PS-SH in water. That is why it has been chosen to work with an intermediate solvent as acetone, which is miscible with water and in which PS-SH is efficiently soluble.

A solution of excess of PS-SH (around 0.3 g) is thus prepared in approximately

60 mL of acetone under stirring in a glass bottle with a plastic plug. 5 mL of concentrated colloidal solution is then slowly added. The solution becomes progressively trouble because PS-SH is not soluble in water. An add of acetone is able to arise the solubility if necessary. Then, the mix is staying stirring during one night under hood. After some hours appears a plastic deposit at the bottom and on the side of the bottle. This deposit is extracted by centrifugation taking off the uncolored supernatant.

Once those nanoparticles are capped with PS, they become insoluble in polar solvents as water, alcohol or acetone. Their surfaces become hydrophobic inducing a good solubility in non-polar, or low polar, solvents as chloroform. The particles are dispersed one more time in a small volume of chloroform ($CHCl_3$) and centrifuged during 3 minutes at 13,000 rpm to take off the excess of PS-SH localized in the supernatant part. This step is repeated to ensure the total suppression of PS-SH in excess. To control this, an addition of isopropanol in the supernatant results in a trouble supernatant if it still contains PS. The functionalized GNPs are then diluted in chloroform and stored in glass vial. They are then dried at room temperature by simple chloroform evaporation. These capped nanoparticles obtained are very stable and can be stored in dry atmosphere for at least two years.

### 2.2.3   Polymer masks for GNPs deposition

The assembling of the GNPs substrate and a microfluidic chip implies some constraints. The main constraint comes from the nature of the assembling, which is based on an activation of the glass slide with oxygen plasma. The presence of GNPs dispersed on the totality of the surface risks reducing the efficiency of the assembling. The choice is then to deposit the nanoparticles only where will be localized the microfluidic channels, i.e. where the glass slide will not be in contact with the microfluidic wall.

Channels were designed by UV-lithography. A polymer resist Shipley S-1813 is spin coated on the glass slide and exposed to a 400 nm light with a global power of 70 $W/cm^2$ using a MJB-21 UV masker. The resist chosen is positive because the used mask is the same than for microfluidic chip molding. After exposition the resist is developed by dissolving the exposed resist in a basic solvent (MF319). This steps are schematized in Fig. 2.5 (1-2).

It finally results in a glass slide capped with Shipley, except in the exposed area drawing the microfluidic design.

### 2.2.4   GNPs deposition and "cleaning"

Once the glass slide has been set up the GNPs can be deposited. Thus, a drop of GNPs diluted in chloroform is deposited on the slide in the channel, and also outside
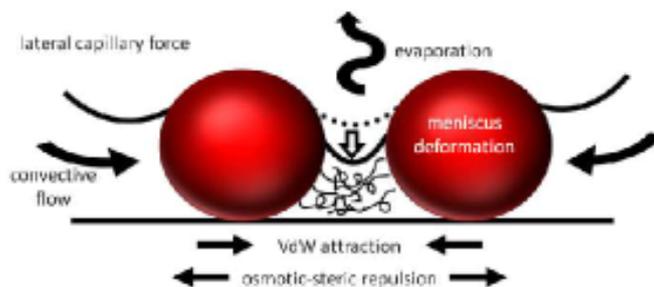
Figure 2.4 – Illustration of the evaporation-induced self-assembly process between two neighboring uncharged nanospheres, in a low-polarity solvent, occurring when the thickness of the solvent layer becomes comparable to the particle diameter. The lateral capillary force acts through convective flow pushing the particles together in response to the deformation of the meniscus between the particles, raising the energy cost caused by the surface tension of the solvent.

(see Fig. 2.5 (3-4)). The concentration of the nanoparticle is difficult to determine just on the basis of the volume of chloroform used. It can only be estimated when the solution changes its color. Evaporation and deposition are more homogenous if the slide is previously deposit on a water film to ensure a better thermal transfer by limiting the condensation. The importance of a good control of the evaporation step is schematized in Fig. 2.4 and results on a nanoparticle film, as shown in Fig. 1.11 (b).

After the evaporation of chloroform a blue layer stays due to the nanoparticle film. The UV-resist is lift-off by a simple immersion of the slide in acetone under low stirring (Fig. 2.5 (5)). Nanoparticles are then present only in the channel designed by UV-lithography. Unfortunately, the so-obtained GNPs are still functionalized with PS-SH chains. The idea is so to expose the glass slide to oxygen plasma during more than 2 hours in a Plasma Cleaner at the highest power (Fig. 2.5 (6)). The oxygen atoms are implies in the etching of the PS-SH capping by breaking the chemical link between thiol and gold atoms.

Finally, after this ionic exposure, the result is a simple glass slide on which are present gold nanoparticles without capping according to the microfluidic design used (Fig. 2.5 (7)). All steps of the deposition and preparation of the substrate are summarized in Fig. 2.5.

## 2.3   Microfluidics and Confocal set up

The SERS substrates detailed before were coupled with two techniques: the microfluidic platform and a confocal microscope, which was built up in the lab.
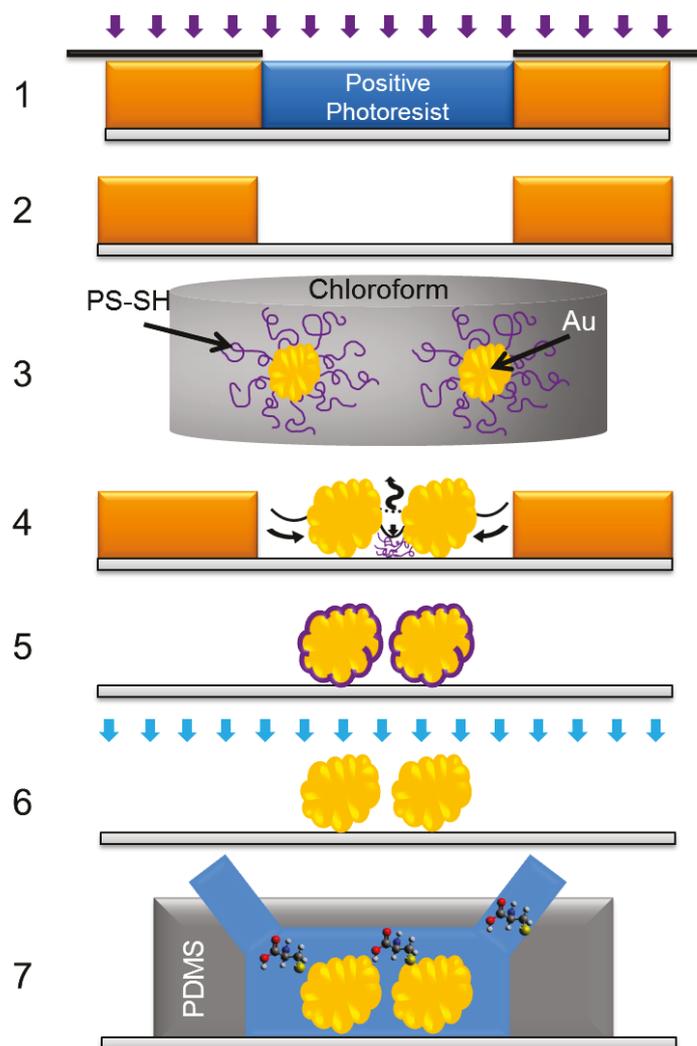
Figure 2.5 – Schematic illustration of the procedure used to integrate the colloidal gold nanoflowers inside the channels of the microfluidic chip. Channel-shaped "boxes" are engraved in a resist layer using UV lithography, and the PS-coated nanoflowers are introduced inside those boxes through solvent evaporation. Lift-off and subsequent oxygen plasma gas exposure lead to channel-shaped monolayers of raspberry-like GNPs with a bare gold surface deposited onto the glass slide, which is then assembled to the PDMS cell.

## 2.3.1 Microfluidic chip molding and alignment

The microfluidic platform used to flow the molecules of interest on the GNPs is a home-made PDMS molded one. A paste of PDMS is prepared by mixing one volume of silicon elastomer curing agent with 10 volumes of silicon elastomer. This paste is
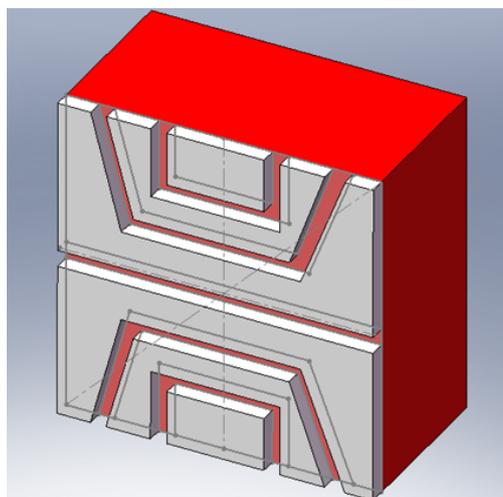
Figure 2.6 – Schematic representation of the microfluidic chip constituted by five channels used in the different experiments.

deposited in a ring on a polymer mask previously obtained by UV exposition of SU-8 polymer layer on a silicon wafer, using the UV-mask used before. The ensemble is then heated at $120°C$ during 45 minutes, before cooling at room temperature and demolding as shown schematically in Fig. 2.6.

Channels of the home-made PDMS microfluidic cell were aligned with the microstructured substrate. The PDMS and the glass were activated by using an $O_2$ Plasma Cleaner. Finally, the substrate is ready to be used in the experiment. A flow controlled pumping system based on a peristaltic pump was used to flow the solution through the microfluidic channel onto the SERS active areas.

## 2.3.2 Characterization confocal setup for dynamic studies

Dynamic SERS measurements are performed using a custom-built confocal Raman setup (Fig. 2.7).

**Excitation of the biomolecules**

A 784 nm laser line is chosen to avoid as much as possible photoluminescence of the gold nanoparticles in visible wavelengths, and the absorption of the light by the water, that mostly occurs in infrared range. The light polarization is controlled using a polarizer and turned with a half-wave plate. To illuminate the sample with a homogenous spot, the beam is expanded before the entry into the microscope by an optical beam expander based on an afocal association of two lenses. The beam is then directed to a Nikon inverted microscope, and more specially a filter cube composed by a laser clean up centered at 785 nm for which the transmission is able only for
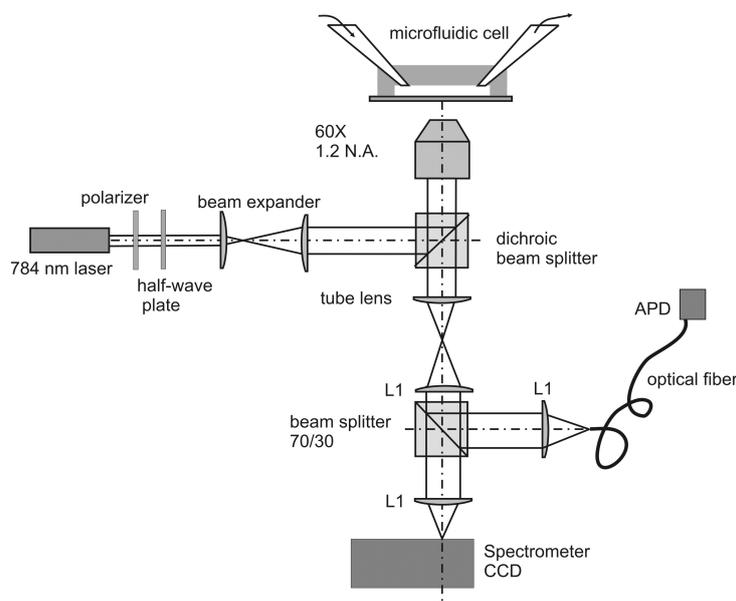
Figure 2.7 – Schematic representation of the confocal optical setup

the 784 nm $\pm$ 3 nm, as shown on Fig. 2.8. The laser is then reflected by a dichroic mirror with a very good efficiency. The specifies of the dichroic mirror are presented on Figure. 2.8.

The collimated beam is finally focused on the glass sample with a Nikon water-immersion objective x60, with a high Numerical Aperture of 1.2. The power of the excitation is controlled just before the objective to take into account the different mirrors and beam splitters placed on the beam path. The excitation power is then adjustable between 100 $\mu$W and 1 mW.

**Detection of the emitted signal**

The signal recorded on the different detectors is the backscattered light from the sample. This signal consists mainly of three contributions: the Rayleigh signal and the reflection of the laser on the surface, the photoluminescence of gold and the Raman backscattering.

The Rayleigh signal, as shown in Fig. 1.2, is at the same wavelength than the laser excitation, i.e. 784 nm. Thus, after travelling back through the objective, the beam is reflected by the dichroic mirror and redirected towards the back side of the microscope, passing through a 50/50 beam splitter before reacting finally with a fibered photomultiplier. This signal enables an easy mapping of the surface to localize the nanoparticles on the surface.
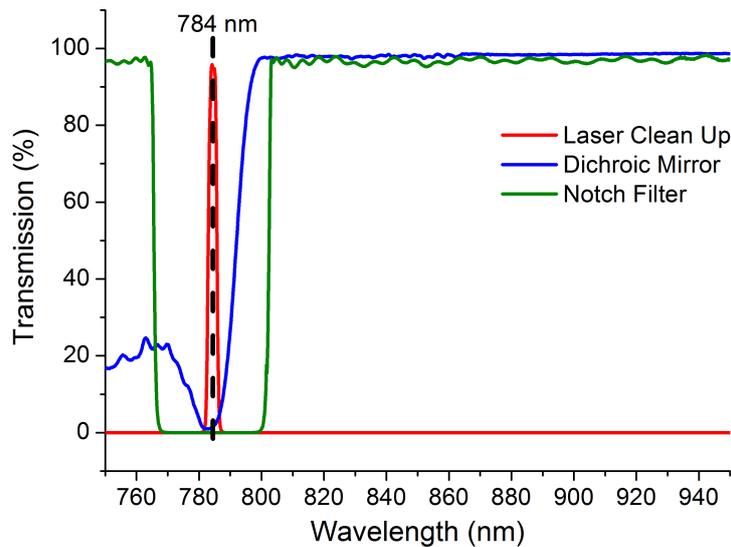
Figure 2.8 – Transmission characteristics of the different filters compound the filter cube.

Another backscattered signal is the photoluminescence of gold. In fact, due to the laser excitation, the electron of the valence band of gold are excited and some energy coming from the laser is also transferred to gold atoms. The backscattered light reacting with gold is shifted to higher wavelength than the excitation. This light is transmitted through the objective and through the dichroic mirror and then filtered by the notch filter. This results in a large bump in the spectrum record after filtering. However, the photoluminescence effect is reduced using a 784 nm excitation laser. Indeed, this effect is more pronounced at lower wavelengths around 500-600 nm corresponding to the resonance of GNPs, i.e. to wavelengths that gold absorbs more.

The third signal is the signal of interest: the Raman scattering. As described on Fig. 1.2, the Raman signal is coming from the excitation of a molecule by a monochromatic light, in my case the laser. Two different types of Raman scattering should happened: the Stokes and the Anti-Stokes. It is well-known that the Anti-Stokes effect is rarer and less intense than the Stokes one. That is why it has been chosen to work with the Stokes Raman scattering. This signal is so transmitted through all filters on the way of the backscattered light.

After the microscope, the backscattered signal transmitted through the filters is divided. 70 % of the signal is focalized in a multimode fiber plugged on an Avalanche PhotoDiode (APD) which has a very good temporal resolution (lower than 1 ms). The other 30 % are directed and focalized on a spectrometer. This spectrometer scatters the light in wavelength with a 0.12 nm spectral resolution.

The next step is to find an efficient hot spot. The surface is scanned by moving
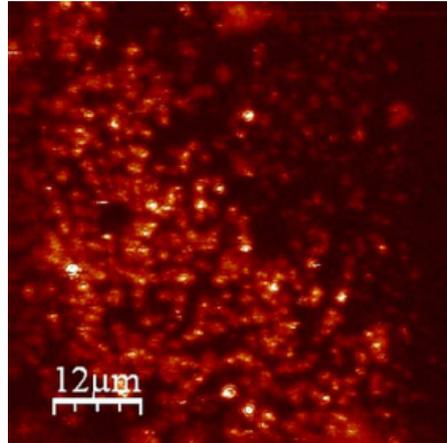
Figure 2.9 – Map of Raman signal measured on the APD. 256 x 256 pxls.
Scan speed = 1 line per second

the substrate using a piezo stage. Thus, the map reconstructed from the Raman signal measure on the APD allows localizing the hot spots precisely as shown in Fig 2.9. We choose one hot spot, we let the laser on it and we record all of the interest signals (APD intensity and spectra). Classically, our acquisition are so composed between 512 and 2048 spectra of 1 s or from 516,096 to 2,621,440 Raman intensities on the APD.

## 2.4 Statistical analysis of the dynamic studies

After choosing the hot spot and recording the APD and the spectrometer signals, it is necessary to find a systematic method of analysis. In fact, the high number of spectra recorded implies the necessity of using a statistical method to resume the complete record with a minimum of variables.

There are also principally three types of statistical methods: The first is the Mandel M-Factor determination. This method is based on the intensity variations of each wavenumber. The second is the Probability Density Function. This one is based on the intensity distributions of each wavenumber during the acquisition. The last method is the Principal Component Analysis. It is based on the intensity fluctuations of the whole spectrum. This last method is called multivariate analysis.

### 2.4.1 Mandel M-factor

The first statistical method is so called the Mandel M-factor determination. This factor is a good indicator of the intensity fluctuations and can be calculated at the dif-

50

ferent wavenumbers. Historically, this factor has been defined by Mandel to solve the problems of temporal fluctuations of the fluorescent signal of resonant molecules [55, 56].

Mathematically, this factor is based on the calculation at each wavenumber $\nu$ of the variance $\sigma_\nu^2$ and the average of the intensity $< I_\nu >$ in time [56, 57] and a simplified definition has been proposed by Margolin et al. [58]:

$$M_\nu = -1 + \frac{\sigma_\nu^2}{< I_\nu >} \tag{2.1}$$

A null value of M corresponds to a non-fluctuating system, i.e. a Poissonian distribution of intensity. A positive M parameter indicates a super-Poissonian statistics.

The resulting spectrum informs about the most fluctuating Raman bands during the acquisition.

## 2.4.2 Probability Density Function

The second statistical method is the Probability Density Function (PDF). This method is based on the intensity distribution during the acquisition. In fact, the PDFs $p_\nu(A_\nu)$ of the photon rates $A_\nu$ are estimated experimentally at each wavenumber $\nu$. Histograms are built using a bin number n defined by the root of the number N of acquired spectra and bin sizes $\Delta A_\nu$ obtained by the intervals between the minimum $A_{min,\nu}$ and the maximum $A_{max,\nu}$ of the photon rates.

$$n = \sqrt{N} \tag{2.2}$$

$$\Delta A_\nu = \frac{A_{max,\nu} - A_{min,\nu}}{n} \tag{2.3}$$

$p_\nu(A_\nu)$ are defined by the number of events $n_\nu(A_\nu)$ having is photon rate $A_\nu$ comprised in the interval between $A_\nu$ and $A_\nu + \Delta A_\nu$ and divided by n and $\Delta A_\nu$.

$$p_\nu(A_\nu) = \frac{n_\nu(A_\nu)}{n\Delta A_\nu} \tag{2.4}$$

Hence, the integral of $p_\nu(A_\nu)$ over the entire range of photon rates $A_\nu$ is equal to one.

For this statistics is associated a specific representation based on the spectral intensity vs wavenumber with a colored map associated to the probability distribution in logarithmic scale. An application of this method is shown in Fig. 2.10.
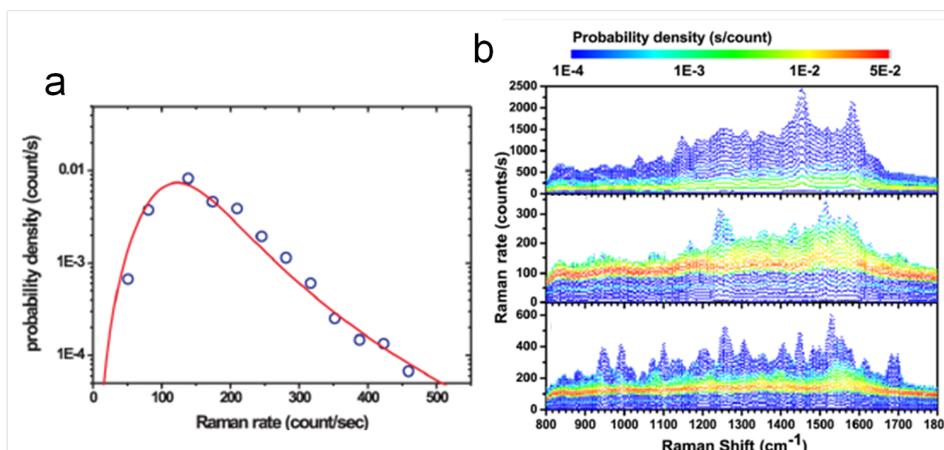
Figure 2.10 – Example of an application of Probability Density Function adapted from [25]. (a) At 1540 cm$^{-1}$ line or (b) all lines for different temporal areas. The colorbar represents the Probability Density.

The application of this method is complementary to Mandel factor. The most probable events are colored in red whereas rare events are in blue. In a more conventional SERS acquisition, the rare and bright events are usually either masked in the ensemble average using long time acquisition or only partially analyzed by extracting the highest intensity spectrum. The visual representation summarizes the systematic statistical treatment enabling a probability density assessment of both intensity fluctuations and heterogeneities of the spectral shape.

## 2.4.3   Principal Component Analysis

The third statistical method is the multivariate one. Beyond identifying the fluctuating lines with the M-factor or the PDF, a complementary approach is to sort the spectra into different categories according to their degree of similarity. The identification of the number of families of independent spectra provides a deeper insight into the source of the fluctuations. This decomposition is conducted through a statistical processing of the data using Principal Component Analysis (PCA) [49].

The Principal Component Analysis (PCA) method is a well-established technique used in analytical spectroscopy. PCA consists in applying specific linear transformations to convert a set of observations of possibly correlated variables into a set of values of linearly uncorrelated variables called principal components. PCA works well only for data that are a linear combination of independent sources, as it happens here with the additive contributions of the SERS signals from multiple types of molecules. PCA can be applied by eigenvalue decomposition of a data covariance (or correlation) matrix or singular value decomposition of a data matrix, usually after mean centering (and normalizing or using Z-scores) the data matrix for each attribute. Most of the

steps detailed after are schematically represented in Fig. 2.11.

The goal is to transform a given data set $X$ of dimension $M$ to an alternative data set $Y$ of smaller dimension $L$. Equivalently, we are seeking to find the matrix $Y$, where $Y$ is the Karhunen–Loève transform (KLT) of matrix $X$:

$$\mathbf{Y} = \mathbb{KLT}\{\mathbf{X}\} \tag{2.5}$$

Our data set comprising a set of $M = 512$ spectra and we want to reduce the data so that each observation can be described with only L spectra that we hope less numerous than M, $L < M$. One spectrum is arranged as a set of $N = 1024$ wavenumbers $\mathbf{x}_1 \ldots \mathbf{x}_N$ with each $\mathbf{x}_n$ representing a single grouped observation of the M variables.

We dispose then to a single matrix $X$ of dimensions $M \times N$.

We then find the empirical mean $u[m]$ along each spectrum $m = 1, ..., M$.

$$u[m] = \frac{1}{N} \sum_{n=1}^{N} X[m, n] \tag{2.6}$$

We proceed by centering the data as follows:

$$\mathbf{B} = \mathbf{X} - \mathbf{uh} \tag{2.7}$$

Where $h$ is a $1 \times N$ row vector : $h[n] = 1$ for $n = 1, \ldots, N$

The $M \times M$ empirical covariance matrix $C$ is calculated from the outer product of matrix B with itself:

$$\mathbf{C} = \mathbb{E}\left[\mathbf{B} \otimes \mathbf{B}\right] = \mathbb{E}\left[\mathbf{B} \cdot \mathbf{B}^*\right] = \frac{1}{N-1}\mathbf{B} \cdot \mathbf{B}^* \tag{2.8}$$

Where $\mathbb{E}$ is the expected value operator, $\otimes$ is the outer product operator, and $*$ is the conjugate transpose operator. Since $B$ consists entirely of real numbers, the "conjugate transpose" is the same as the regular transpose.

The matrix $V$ of eigenvectors which diagonalizes the covariance matrix $C$:

$$\mathbf{V}^{-1}\mathbf{CV} = \mathbf{D} \tag{2.9}$$

Where $D$ is the diagonal matrix of eigenvalues of $C$. This step used a computer-based algorithm for computing eigenvectors and eigenvalues.

Matrix D will take the form of an $M \times M$ diagonal matrix, where

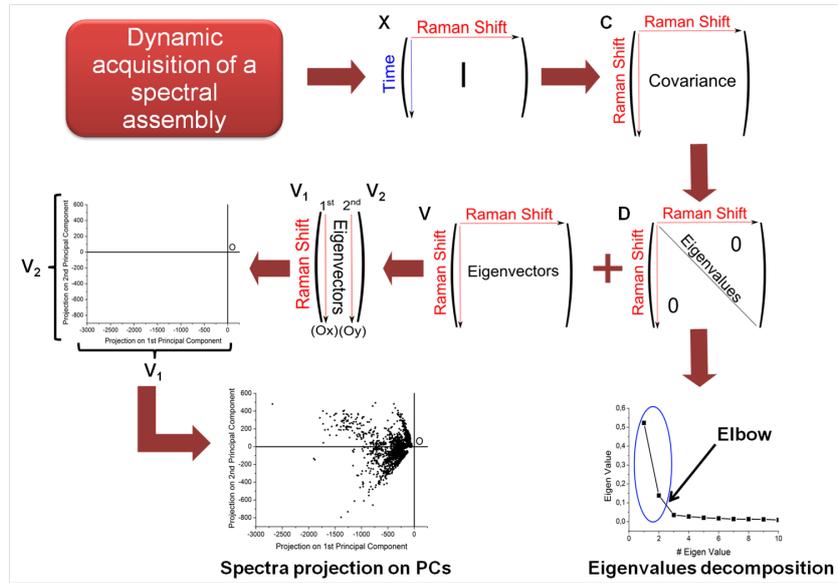$$D[p, q] = \lambda_m \qquad \text{for } p = q = m \tag{2.10}$$

Figure 2.11 – Schematic representation of the most important steps of Principal Components Analysis as applied in our case. Each matrix is notified with its name given in the different equations.

is the $m^{th}$ eigenvalue of the covariance matrix C, and

$$D[p, q] = 0 \qquad \text{for } p \neq q. \qquad (2.11)$$

Matrix $V$, also of dimension $M \times M$, contains $M$ column vectors, each of length $M$, which represent the $M$ eigenvectors of the covariance matrix $C$.

The eigenvalues and eigenvectors are ordered and paired. The $m^{th}$ eigenvalue corresponds to the $m^{th}$ eigenvector.

The columns of the eigenvector matrix $V$ and eigenvalue matrix $D$ are then sorted in order of importance in order of decreasing eigenvalue magnitude (i.e. in absolute value). The eigenvalues represent the distribution of the source data's energy among each of the eigenvectors, where the eigenvectors form a basis for the data.

By normalizing each eigenvalue by the sum of all eigenvalues, they are expressed as percentage of variance. The number of Principal Components needed to efficiently describe the experiment is determined by looking at the evolution of the eigenvalues after sorting them from the maximal to the minimal eigenvalues.

This number is determined by the elbow method (Fig. 2.11). In fact, on the eigenvalue evolution figure, an elbow appears. The eigenvectors necessaries are so the one corresponding to the eigenvalues higher than the elbow position, and the other eigen-

vectors are less impacting and can be also neglected to describe the experiment.

The necessary Principal Components compose the eigenspace that will be used to project each spectrum of the experiment. Each spectrum is decomposed on the basis formed by these first Principal Components to obtain $\mathbf{X_{PCs}}$, thereby approximating $\mathbf{X}$. Also if 3 PCs are necessaries, it results in

$$\mathbf{X_{PCs}} = \alpha_1.V_1 + \alpha_2.V_2 + \alpha_3.V_3 \simeq \mathbf{X} \tag{2.12}$$

The projection of each $\mathbf{X_{PCs}}$ in the eigenspace allows to highlight correlations between the different spectra. Thus, spectra projected along a same axis are correlated, the only variable parameter is the total intensity of each spectrum. In summary, this method can be schematically represented as on Fig. 2.11.

## 2.5   Conclusion

This chapter was devoted to present the experimental methods we have developed specifically to monitor accurately the fluctuations of Raman signal expected from single molecule. It includes four innovative key issues:

1. A SERS active substrate based on Chebyshev GNPs with stable high enhancement of the electromagnetic field.

2. The implementation of the microfluidics enabling to flow various concentration of low volume of solution.

3. A confocal microscope enabling to monitor the dynamic of the Raman fluctuations on one hot spot featuring a high temporal resolution of 1 ms.

4. A statistical analysis of the whole series of data, namely between 512 and 2048 spectra of 1 s or from 516,096 to 2,621,440 Raman events.

Performance of the nanosensor

## 3.1  Introduction

This chapter aims at studying how SERS on a single hot spot can be used to quantify the concentration of an analyte. We chose to use a dye to maximize the chance of success for detecting ultra-low concentration. The Methylene Blue is indeed a resonant molecule in visible range of which the spectral assignment is already well established.

This chapter will answer five critical issues:

- What is the reproducibility of the detection and under which conditions statistic properties can be deduced from a single but relatively long time series, in other words is the phenomenon ergodic?

- What are the most significant parameters to follow? The intensity of the signal, the number of events or the characteristic time attributed to the lifetime of the molecule in the hot spot?

- Can we define a sensitivity of this nanosensors within a wide linear regime between lower and higher concentration limits? Is the estimation of the EF necessary in this context?

- What are the forces driven the molecule in the hot spot? Is it a random motion like a diffusion process or an adsorption process with a characteristic residence time?

- Does the definition of an enhancement factor at the single hot level make sense?

We will first discuss the experimental conditions to have a reproducible sensor. Next we examine the sensitivity of the sensor by analyzing the time series of the APD signal. Fluctuations will be analyzed by two approaches : the intensity distribution and the noise analysis in the frequency space.

## 3.2 Experimental methods

### 3.2.1 Experiment plan

A specific experimental plan was designed to investigate the effect of both the EFs of the hot spot and the changes in concentration of Methylene Blue. The concentration was ranged over 12 orders of magnitude from 10 aM up to 10 $\mu$M and the flow rate imposed by the peristaltic pump is 2 $\mu$L/min. 15 spots easily retrievable on the SERS substrate have been probed. Those spots were located by means of lithographed marks deposited on the substrate.

The running order of concentrations was from lowest to highest molecule content. We added one spot to analyze when increasing the concentration by one order of magnitude. Typically, for 1 fM the study is made on 5 hot spots, 6 spots for 10 fM, 7 for 100 fM, ..., to achieve with 15 spots for the 10 $\mu$M Methylene Blue solution.

For each series on one spot at a given concentration, we recorded 512 spectra of one second resulting in 655,360 data points on the Avalanche PhotoDiode for 512 s acquisition.

### 3.2.2 Spurious signal

Special care was taken to avoid the surface contamination or the pollutant of the solution. We have found out two main sources: one is coming from the surfactant of the colloids ; the second aspect is the release of the polymer constitutions of the PDMS microfluidics.

It is very important to keep to a high level of cleanliness of the nanoparticles. GNPs can be cleaned sequentially by oxidizing organics using plasma cleaner and by rinsing off thoroughly with deionized water.

We give in Fig. 3.1 four consecutive times series of the rinsing of nanoparticle with DI water in microfluifluics. The first step is characterized by a significant decrease of the base line in the first 50 s. Then the baseline decreases slightly in 1500 s to reach a value close to zero. This trend is even clearer by following the first mode of the intensity distribution. A broad and a non–central distribution is a strong indication of
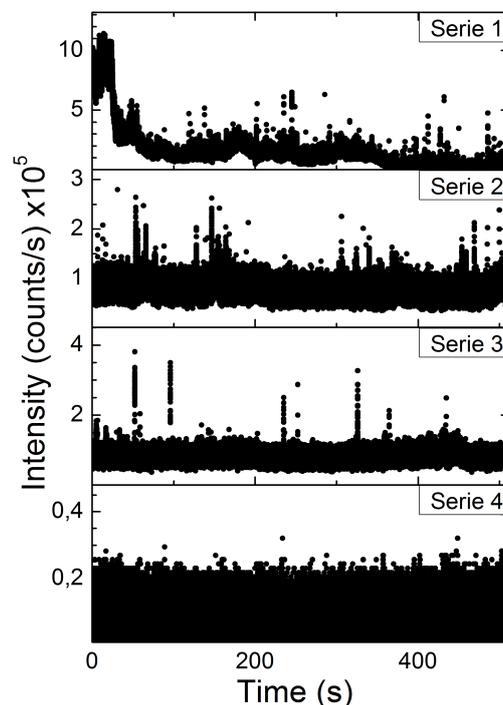
Figure 3.1 – Temporal fingerprints of MB recorded on the APD during 512 s
with 1280 points/s for different concentration of Methylene Blue on the
spot 2.

contaminant acquired in pure water.

The kinetics of desorption induced by the rinsing process is shown in microfluidics:
it is quite long to obtain a clear flat baseline. Note that the baseline is punctuated
by many short peaks, on a small extent at the beginning, at the maximum level at
600 s and since then those short events become to disappear again. This clean state is
confirmed spectrally with zero Raman response. The first mode of the intensity must
be also centered close to zero.

Spectrally, two types of spectra were discerned. Spectra corresponding to the base-
line were attributed by characteristic bands between 1440 and 1600 cm-1 to the PS-SH
(Fig. 3.2 (a)) (yellow area). Those spectra highly pronounced at the beginning of the
experiment disappear with the rinsing. Spectra corresponding to the short events were
attributed by two characteristic peaks in the grey area Fig. 3.2 (b) of some monomer
dimer or trimer of the PDMS molecule. Those Raman bands are only visible with the
release by a bright new microfluidics device.

Now that we have described the potential adverse effects to which attention should

59
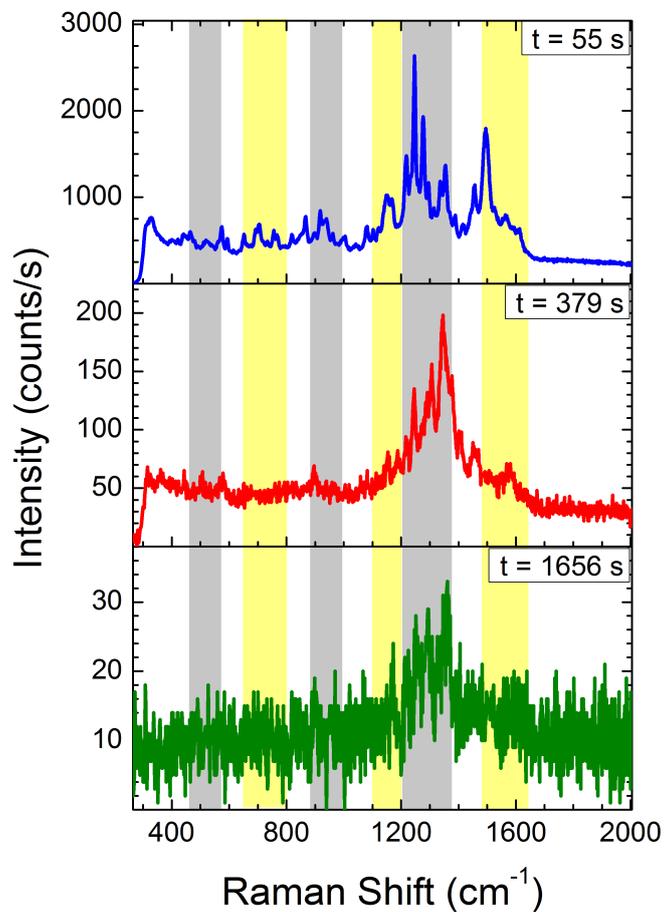
Figure 3.2 – SERS spectra recorded at different times of the first series. Times are precised in legend. Positions of the PDMS Raman bands from [59, 60] are grey areas, and positions of PS-SH Raman bands from [61] are yellow areas.

be paid in quantifying the signal, we can move to determine the LOD and the sensitivity of the sensor.

## 3.3 Quantification of the concentration

Measuring a concentration means being able to calibrate the hot spot and to determine a sensitivity dS/dC and the LOD.

Figure 3.3 – (a) Temporal fingerprints of MB recorded on the APD during 512 s with 1280 points/s for different concentrations of Methylene Blue on the spot 2. (b) Probability Density Function of the intensity distribution for different concentrations of Methylene Blue for spot 2.
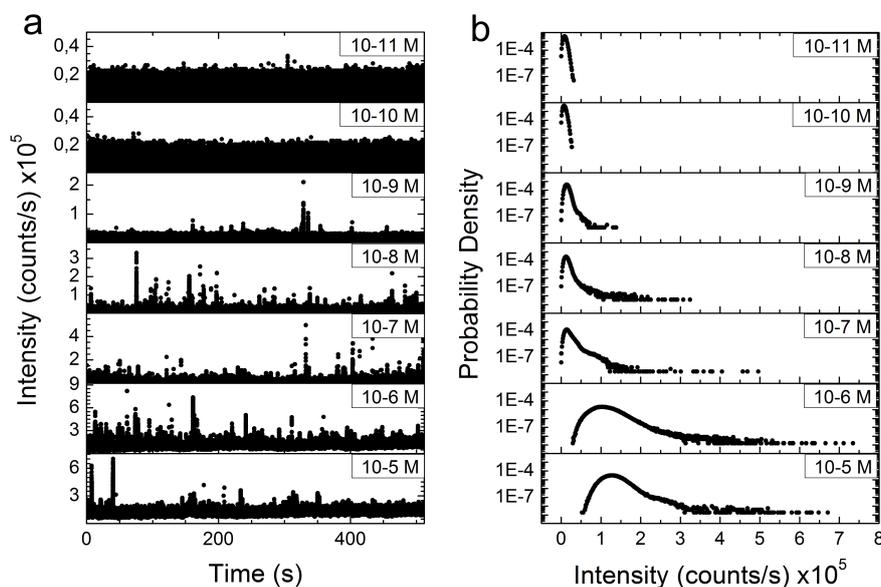
### 3.3.1 Limit Of Detection

We first investigate the signal as the intensity of Raman events. The time series are shown in Fig. 3.3 (a) by varying from concentration of the MB over 7 concentration decades. It was found that below a concentration of $10^{-10}$ M when the SERS substrate is cleaned and the microfluidics passivated by ageing, no Raman events could be detected which means that all Raman spectra are flat. The minimum LOD for our device is then $10^{-10}$ M of MB, namely around 30 pg/mL for 512 s acquisition. This limit can be considered as relatively high since the hot spot has the capability to detect a single molecule event. Indeed, the LOD is limited by the size of the hot spot (the smaller the hotspot, the high LOD) as well as the microfluidics device. Since the GNP are on the wall of the microfluidics, the parabolic profile of the flow rate imposes a zero speed of molecule at the surface thereby the molecules enter into the hot spot by a diffusion process. At a concentration of $10^{-10}$ M, the average distance between molecules estimated to 100 nm exceeds the cross section of our nanosensor.

### 3.3.2 Sensitivity

When increasing the concentration, the observation includes two major steps. First, between $10^{-10}$ M and $10^{-7}$ M, the number of events and as the intensity of the peaks
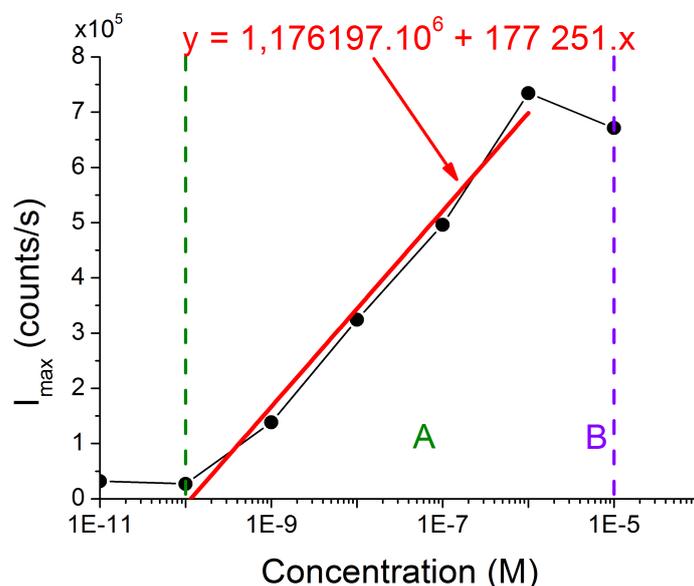
$$y = 1{,}176197.10^6 + 177\,251.x$$

Figure 3.4 – Maximal intensity recorded on APD versus concentration of Methylene Blue.

increase concurrently. The second step corresponds above $10^{-7}$ M to the slight drift of the signal baseline ; this background is no longer necessarily flat and hence the height of the peaks could appear as reduced.

Nevertheless, a more measurable, or quantitative, approach was needed. This was achieved through the statistical treatment by calculating the power density function of the time series Fig. 3.3 (b). Before sensing the MB, the Raman intensity is clearly distributed by a symmetric normal law centered close to the zero intensity. This is clearly the blank signal of the experiment including electronic and APD noise.

Once the MB is probed, the distribution changes shape, the first mode is always centered close to zero, but reveals a right tail having its length clearly dependent of the concentration. By simply plotting the maximum of intensity , namely the tail length versus the concentration, we can define a linear sensitivity of the nanosensor $S = 1.7x10^5\ Hz/M$ (Fig. 3.4) over 4 concentration decades between $10^{-10}$ M and $10^{-6}$ M. This great result demonstrate that a nanosensor can be used not only to characterize a single molecule but actually can be used as a counter of molecules: "a molecule meter".

The second LOD correspond to the saturation of the nanosensor: it is estimated here to $10^{-6}$ M. At this concentration, we assume that we switch from a single molecule

regime to a more complex configuration involving multiple molecules adsorbed or not. The average distance between molecules at $10^{-6}$ M is around 1 nm. This is also the concentration when the first mode of the distribution corresponding to the signal baseline starts to drift to higher intensity.

All the pdf distribution has been fitted using the Log-logistic function. The Log-logistic distribution, sometimes known as the Fisk distribution, is encountered in a variety of fields (economy, biology and physics) to analyze life time data [25]. The Log-logistic distribution was applied successfully to describe a process that is the product of a number of variables of small amplitude, namely we attempt to describe the process as a coupling between the substrate properties (EM field or electron tunneling) and the molecular Raman scattering. The two-parameter log logistic distribution is described by its theoretical power density function $f(x, \alpha, \beta)$:

$$f(x; \alpha, \beta) = \frac{(\beta/\alpha)(x/\alpha)^{\beta-1}}{[1 + (x/\alpha)^{\beta}]^2} \tag{3.1}$$

$\alpha$ is a scale parameter that corresponds to the median of the distribution. Note that the median is less sensitive to the extreme values compared to the mean. $\alpha$ is obtained by the following relations:

$$\alpha = \frac{sinb}{b}x = median(x) \tag{3.2}$$

Where b is obtained by solving the equation:

$$RSD^2 = \frac{2b}{sin2b} - \frac{b^2}{sinb} \tag{3.3}$$

The shape parameter $\beta$ can be deduced from b

$$\beta = \frac{\pi}{b} \tag{3.4}$$

When $\beta$ is high, $\alpha$ tends to mean(x). The Log-logistic function was found to well describe all the pdf for all concentrations only with two parameters.

The Fig. 3.5 shows that the median of the distribution does not move until $10^{-6}$. $\beta$ is then the unique parameter useful for measuring the concentration. We measure a linear sensitivity of $d\beta/dC$ = -0.26574. Above $10^{-6}$, alpha seems to be the most appropriate candidate to follow the highest concentration. However, the low number of data does not allowed determining correctly the linear slope.

To sum up the concentration analysis, we have demonstrated that the single molecule regime can be clearly determined: its range over 5 concentration decades is between two clearly defined LODs. We prove also that the concentration can be measured by two ways: a simple way using $I_{max}$ or a more statistical approach using $\beta$.
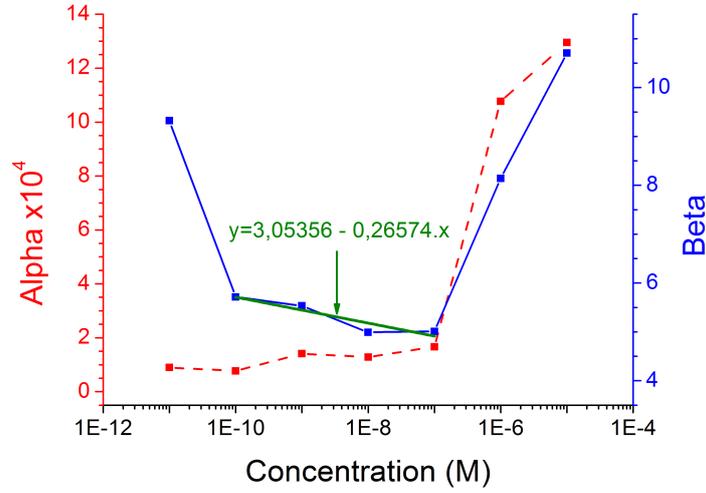
Figure 3.5 – Evolutions of the Log-logistic law parameters $\alpha$ and $\beta$ with the concentrations of Methylene Blue.

Our own preference is the statistical approach since it takes into account the whole set of measurements and not as isolated point as for I$_{max}$, which is laser power dependent.

## 3.4 Frequency analysis

The Log-logistic distribution used for survival studies motivates us to go further in the temporal analysis and to the supplement the intensity analysis. There could be two cooperative approaches: in fluorescence study, the calculation of the autocorrelation function in the time space was found particularly useful to estimate the live time of molecules. We implement this method and found in our case that this methods requires some model to fit our data, these models developed for fluorescence are not necessary valid in the Raman case. We turn then in a frequency space to treat the main noise sources. Fourier analysis converts time to frequency.

$$x_n = \frac{1}{N} \sum_{k=0}^{N-1} X_k . e^{i2\pi kn/N}, \ n \in \mathbb{Z} \tag{3.5}$$

Where N is the number of observation on the APD, n the frequency, X$_k$ the intensity at the moment k and x$_n$ the Point Spread Density for the frequency n.

The Fast Fourier Transform of the time series at each concentration was applied as shown in the Fig 3.6. The signal processing of our data has enabled us to identify the three key sources of noise.
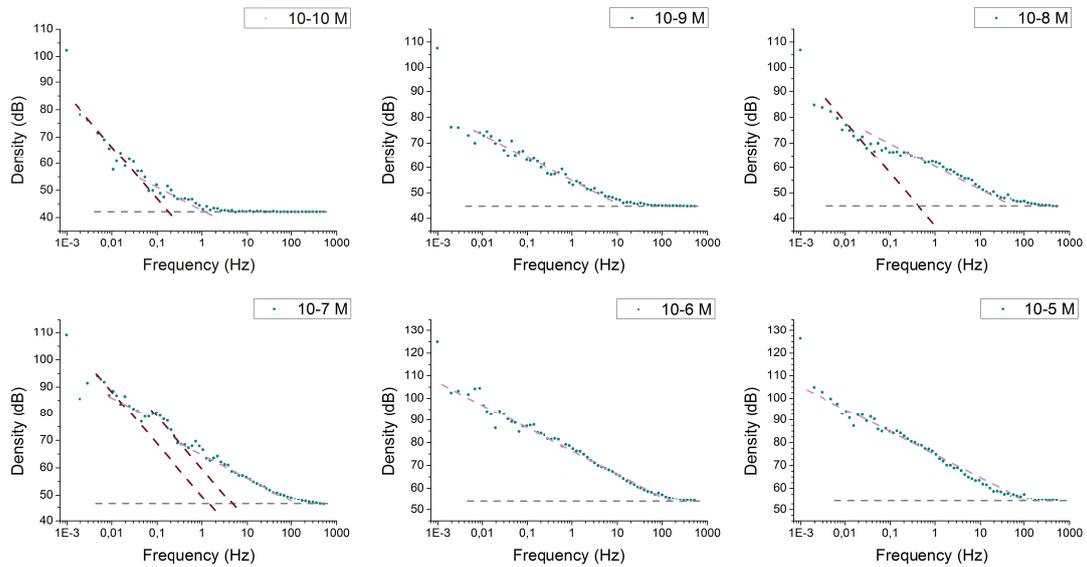
Figure 3.6 – Point Spread Density characteristics versus Frequencies obtained
for each concentration sounded.

The white noise is a random signal with a constant power spectral density independent of the frequency.

Pink noise or 1/f noise is a signal or process with a frequency spectrum such that the power spectral density (energy or power per Hz) is inversely proportional to the frequency of the signal. It decreases in power by 3 dB per octave (-10 dB per decade). It generally comes from condensed matter systems in quasi-equilibrium.

The Brown noise has its spectral density inversely proportional to $f^2$, meaning it has more energy at lower frequencies, even more so than pink noise. It decreases in power by 6 dB per octave (-20 dB per decade). It comes from the Brownian motion, namely the random motion of particles suspended in a liquid.

We have shown previously by analyzing the intensity fluctuation that at $10^{-10}$ M almost no events related to the MB was detected. The power density is distributed over frequency as follows: a white noise when f > 10 Hz (time < 50 ms) ; a pink noise easily identifiable with its slope of -10 dB/decade in the f range between 0.2 and 10 Hz ; below 0.2 Hz (time > 5 s ), the signal behaves as a random brown noise (-20 dB per decade).

Increasing the concentration shows a reduction of the white and Brownian noise and range to the profit the pink noise at high frequency. The most striking figure of this phenomenon is at $10^{-7}$ M when the shortest and highest peak are visible in the
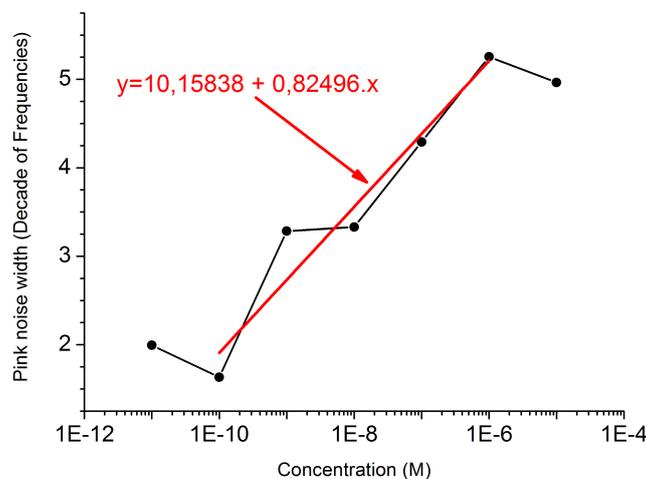
65

Figure 3.7 – Evolution of the frequency width from pink signal with concen-
tration.

time series. The pink noise is distributed over a large f range (0.1 up to 1000 Hz)
that the bright events can be as short as 1 ms and the average time between bright
peaks can reach 10 s. The nature of this pink noise gives information on the nature
of molecule. MB molecules should sense an equilibrium potential, which means the
molecules are either trapped by the hot spot or even adsorbed the GNPs surface.

All of this reinforces the idea that the nanosensor basically senses the molecules
adsorbed on the surface and is less sensitive to the molecules diffusing freely or ran-
domly in the hot spot space.

The frequency analysis can also be used to determine the concentration. Fig 3.7
shows the total extent of the pink noise as a function of the concentration. Based
on this curve, a linear relation between the width of the pink noise and the con-
centration is confirmed. Thus, this approach allows to determine a sensitivity of
$dF_{width}/dC = 0.82496$. So, the frequency analysis is not only a qualitative tool in-
forming about the nature of the interactions, but also a good approach to determine
and confirm a linear sensitivity with the concentration previously obtained with $I_{max}$
and $\beta$.

## 3.5 Spectral analysis

We have previously highlighted a good quantitative sensitivity however the ques-
tion of the spectral determination of the presence of Methylene Blue is necessary to
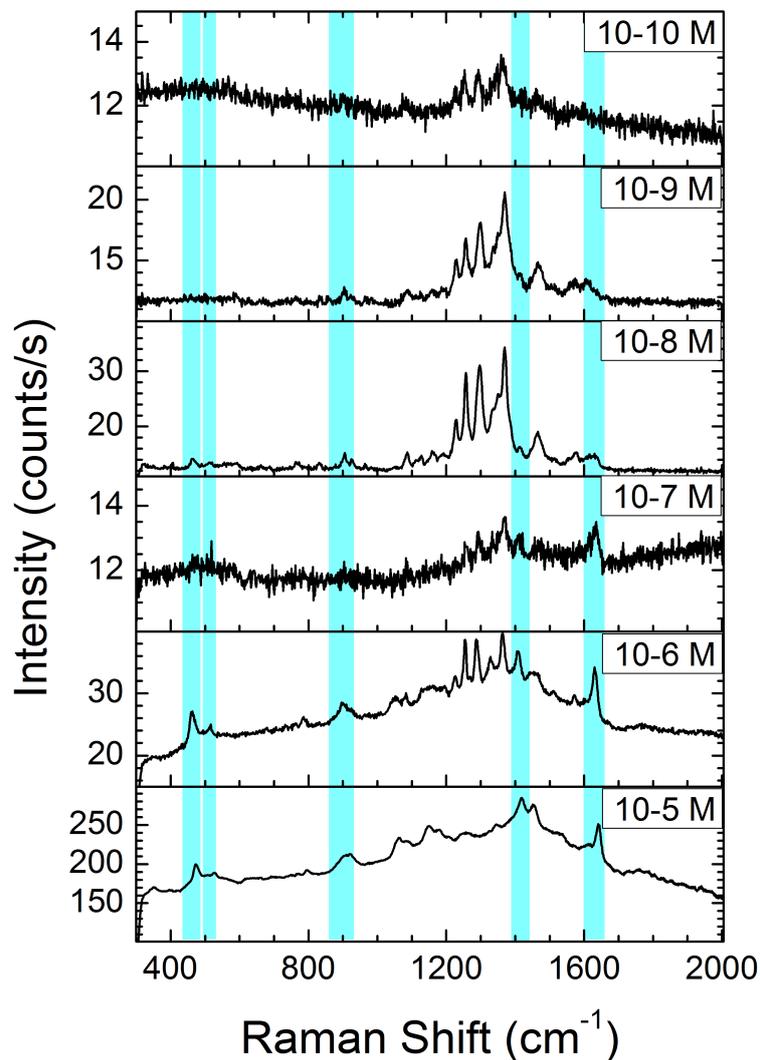
Figure 3.8 – Average SERS spectra recorded at different concentrations of Methylene Blue. Positions of the Methylene Blue Raman bands from [62, 63] are blue areas.

determine the spectral LOD of the system.

Thus, for each concentration has been recording simultaneously to the APD signal the spectral fingerprint each second on the spectrometer. Fig. 3.8 shows the average spectrum for each concentration sounded. The specific Raman bands of Methylene Blue fingerprint are blue areas on this figure.

It clearly appears that Methylene Blue is detected spectrally for at least $10^{-9}$ M

notably at 1400 cm$^{-1}$ and slightly around 450 and 850 cm$^{-1}$. The spectral fingerprint of Methylene Blue is then more and more marked with the increase of the concentration. However, pollutants previously observed in Fig. 3.2 are still presents and impact the spectra of the lowest concentration.

For the highest concentrations (10$^{-6}$ and 10$^{-5}$ M), the presence of Methylene Blue on the surface is confirmed by the high spectral resolution and spectral intensity observed. The spectral fingerprint is very high but superimposed to a high background. This background results from the fluorescence of the molecule. Methylene Blue is not very resonant at this wavelength but the high concentration implies non-negligible fluorescence from the solution. Moreover, the hot spot enhances the Raman signal and also the fluorescence signal. This effect is called Surface Enhanced Fluorescence (SEF).

## 3.6  Determination of an Enhancement Factor

We have previously stated that for a given hotspot, the intensity clearly varies as a function of the concentration and does fluctuate over time. The stability of GNPs immobilized on the surface must explain that the enhancement factor EF of one hotspot remains unchanged all along the measurement.

We then examined how the variability of the enhancement factor from one hotspot to another could influence the results. By analyzing a panel of 9 hotspots, at a given concentration, 10$^{-9}$ M in Fig. 3.9, it is obvious that no two hot spot are the same. Some (5 and 9) behaves as a concentration of 10$^{-10}$, spots 2,3,4 behaves as 10$^{-9}$; Spot 6 and 7 have the signature of 10$^{-7}$, Spot 8 and 1 act as a concentration 10$^{-6}$. Two conclusions can be drawn from this: i) all the hot spots are ruled by common statistical law, the Log logistic function; ii) the variations in the concentration sensing between 10$^{-10}$ and 10$^{-6}$ M indicate that the enhancement factor from one spot to another can vary at least over 4 orders of magnitude in EF.

The very notion of EF at the single hot spot is extremely difficult to circumscribe. First, the enhanced evanescent electromagnetic field decreases exponentially in a perpendicular direction to the surface. We guess that the molecules depending on where it sits in the EM field can sense statistically all the EM gradient; this should explain the right tail of the Log logistic function.

The EM gradient will be dependent of the steric hindrance of GNPs and of course we cannot guarantee the exact geometry from one spot to another. Even if we are able to estimate the variation of the EF, the exact quantification of EF is even more cumbersome. Using the simplest definition, the EF can be estimated by
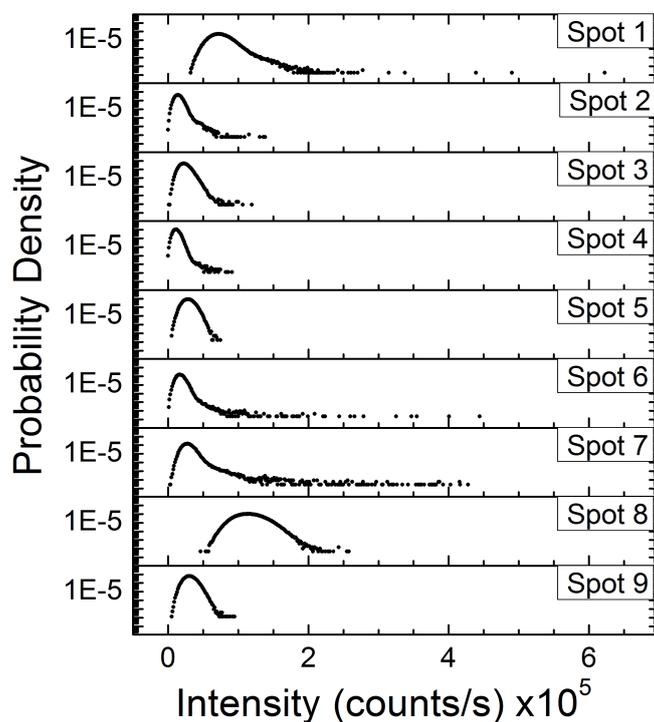
Figure 3.9 – Probability Density Functions of temporal acquisition on the APD for different hot spots at $10^{-9}$ M of Methylene Blue.

$$EF = \frac{C_{SERS-LOD}}{C_{Sol-LOD}} = \frac{10^{-10}\ M}{10^{-5}\ M} = 10^5 \tag{3.6}$$

Therefore, we state cautiously that the EF is within the range $[10^4$ - $10^9]$

Extreme EF of $10^{13}$ as pointed out in literature [41] may not always be necessary to work in a single molecule regime.

## 3.7 Conclusion

This chapter helps us to understand and set up a new strategy for measuring the concentration from the time series of the APD over 5 concentration decades. We found that a calibration of each hot spot is mandatory to assess the right concentration. The sensitivity as well as the enhancement factor is hot spot dependent. Each hotspot obeys to its own parameter of its statistics. Once the hot spot is identified and calibrated in concentration, we can ensure a reproducibility of the nanosensor. The log logistic function is the most appropriate statistical law to be used to predict the intensity fluctuation as a function of the concentration. The maximum intensity and

the $\beta$ parameter of the log logistic were found accurate to define the sensitivity.

The frequency analysis is instructive: the pink noise proved that molecules are in semi equilibrium within the hot spot having a live time dependent on the concentration. Spectroscopic analysis offers the definite advantage to identify the nature of the probed molecules while recognizing potential contaminants (surfactant or microfluidics release).

Selectivity of SERS towards molecular sorting

## 4.1 Introduction

As a very sensitive spectroscopic technique capable of providing precise chemical information of trace amount of molecules, surface enhanced Raman scattering (SERS) is a powerful analytical tool used in biochemistry and biomedical research [64, 65, 66]. In this context, SERS identification of proteins is a formidable task because of their complex structures and often dynamic nature. It was shown recently that the SERS spectrum of a protein [67] is essentially driven by all the amino acids in the vicinity of the protein-metal interface [68]. Experimentally, a spectrum is obtained from either an integrated response of a relative large amount of amino acids adsorbed as a mono-layer over millimeter-scale area or from a time-integrated response of a few molecules adsorbed on nanoparticle aggregates [69]. The tabulated SERS responses are thus an average of a collection of possible adsorption conformations of the amino-acids onto the surface weighted by the stochastic distribution of enhancement sites. Furthermore, vibration fingerprints strongly depend on the molecular environment such as the nominal pH value of the buffer solution[36, 70, 71], the presence of gauche and trans rotamers [72] or the chemical properties of the surface [73]. Orientation of amino acids as well as specific competitive interactions of their functional groups with the gold surface also provide specific SERS spectra [74]. A clear assignment of SERS lines is effectively mitigated when spectra result from the averaged contributions of various molecular conformations.

Distinguishing known spectra in a pool of signatures is usually made possible by statistical approaches. For instance, the bianalyte SERS method [75] uses a mixture of two resonant identifiable molecules to clearly and unambiguously demonstrate count-

ing of molecules at high concentration. Another powerful statistical method relies on principal component analysis (PCA) to extract all redundant information using a simple graphical representation [6, 49]. PCA requires also distinguishable resonant molecules with one distinct Raman fingerprint per species. In low concentration assay, the detection of a few non-resonant proteins highlighted the high temporal variability of the SERS spectra [76]. Thus, single proteins cannot be identified using a unique spectrum but rather by several vibrations of molecular groups near multiple possible anchoring points of the protein.

Our work provides a comprehensive SERS methodology developed both to detect unequivocally single amino acid molecules and to identify its principal spectra. Our approach based on a unique experimental strategy and statistical analysis enables to interpret the temporal fluctuations of spectra in terms of molecular identification. To illustrate the methodology described in this contribution, we detect the SERS response of a cysteine molecule which is a sulfur containing amino acid known for interacting with metallic surface [77, 78]. L-cysteine [73] adsorbs on silver surface through the sulfur atom [79] and also, in zwitterionic form, with the $COO^-$ groups close to the surface [80]. The electrostatic interaction between the positive charge of $NH_3^+$ group and the gold surface was also suggested in the literature [72]. Because of the adsorption specificities, several SERS spectra are thus expected.

I describe here how to analyze statistically the SERS spectra. The implementation of a principal component analysis (PCA) is particularly useful for sorting and identifying the principal spectra of the amino acid. In this context, we find extremely instructive the analysis of short Raman events. We finally assign the experimentally extracted spectra to all possible adsorption configurations of cysteine. The similarities and differences between the principal spectra and the Raman fingerprints simulated provide instructive elements on how cysteine binds to gold surface.

## 4.2 Statistical analyzes based on intensities distributions

The initial experimental step for achieving molecular recognition is to assure a good standard of repeatability of SERS spectra on the same substrate. A local detection combined with short acquisition times reveal the dynamic nature of the SERS spectra. For cysteine, these rapid fluctuations of the SERS response can be appreciated on the video [81]. The time trace is characterized by alternating flat spectra followed by highly unstable lines varying both in wavenumber and amplitude. By simultaneously recording SERS spectra every 1 s with the spectrometer and the total integrated intensity every 1 ms using the APD, the kinetics of the Raman events can be analyzed during long acquisition. An example of an APD time trace acquired over 2000 s is displayed
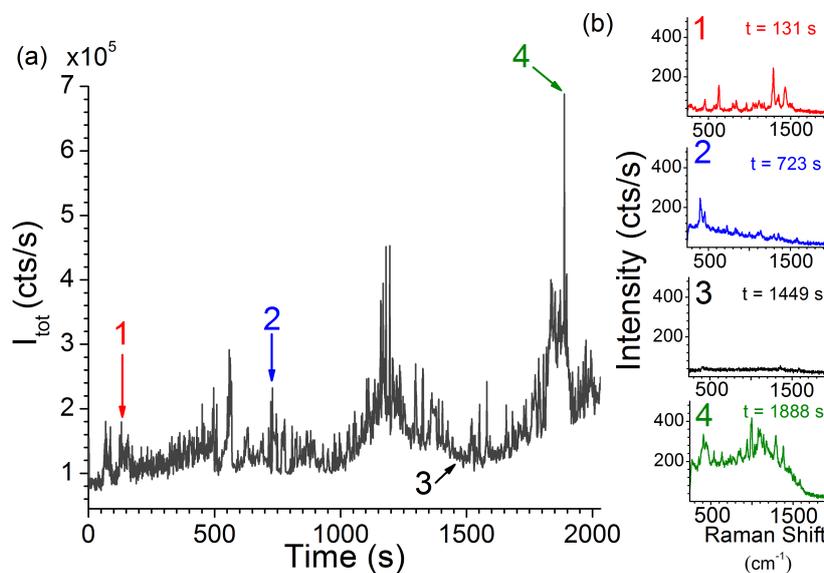
Figure 4.1 – (a) APD rate showing the temporal and amplitude fluctuations of the spectrally-integrated SERS response of the cysteine on raspberry-like GNPs during half an hour. (b) Examples of SERS spectra of cysteine acquired during 1 s at different times of the kinetics.

in Fig. 4.1 (a). The intensity of the SERS signal is punctuated by the occurrence of bright events with distributed amplitudes. The four insets in Fig. 4.1 (b) demonstrate that all the events are characterized by clearly differentiated Raman spectra. Because of the strong fluctuating nature of the cysteine SERS signal, the amino acid signature cannot be simply recognized by a single Raman spectrum.

## 4.2.1 Acquisition time impact on spectral fingerprint

The acquisition length depends on the number of samples and the selected sampling rate. The relevant integration time is obtained when at least two determinations carried out on separate portions of the series are self-similar. Figure 4.2 shows four spectra acquired at the same location with different integration times. Two consecutive spectra integrated for 500 s show distinct features. It becomes clear that repeatable spectra require very long time span as illustrated by the 1000 s and 2000 s integration time. For such long acquisitions, the spectra are inevitably broadened by all fluctuating lines. This may jeopardize the molecular recognition of analytes composed of similar vibrational modes. To circumvent the averaging of information for long acquisition times, we propose in the following a statistical analysis of the fluctuations in order to extract individual reference elements.
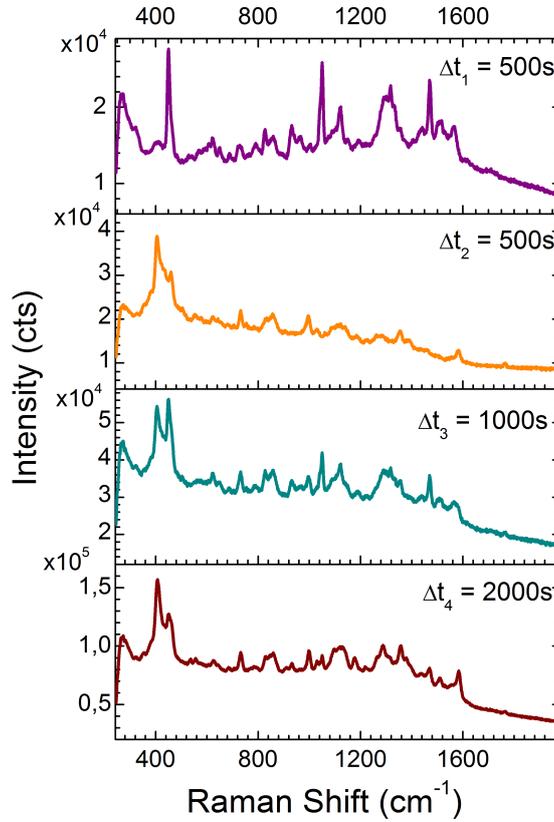
Figure 4.2 – Raman spectroscopy of cysteine adsorbed on GNPs and integrated for: (a) and (b) 500 s ; (c) 1000 s and (d) 2000 s.

### 4.2.2   Mandel Factor

The fluctuations of the spectra are first quantified by using the wavenumber-dependent Mandel parameter $M_\nu$ defined by:

$$M_\nu = -1 + \frac{\sigma_\nu^2}{< I_\nu >} \tag{4.1}$$

and calculated for each wavenumber $\nu$. $\sigma_\nu^2$ and $< I_\nu >$ are the variance and the mean of the Raman intensity, respectively. A null value of $M_\nu$ corresponds to a non-fluctuating system, *i.e.* a Poissonian distribution of intensity. A positive $M$ parameter indicates a super-Poissonian statistics. This is particularly useful to quantify how the distribution is spread compared to the mean rates [25]. From this M-factor calculation results a first spectral fingerprint of cysteine.

The spectral variation of $M_\nu$ calculated for a representative series of 2000 spectra is depicted in Fig. 4.3. The $M$-factor is a first simple indicator of the fluctuating character of the spectra. About 30 fluctuating Raman bands are thus identified between
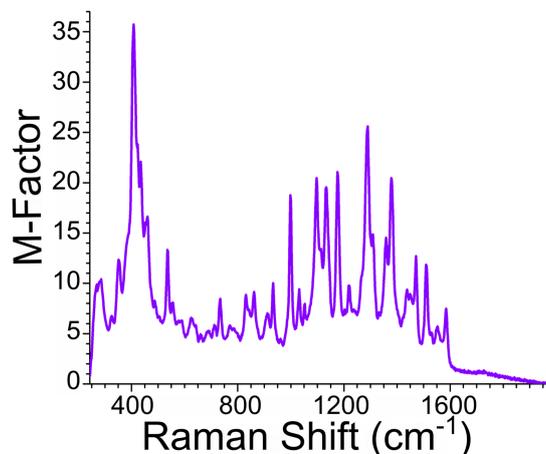
Figure 4.3 – Spectral evolution of the Mandel parameter analyzed for 2000 s.

the 350 cm$^{-1}$ and 1600 cm$^{-1}$ lines. Unfortunately, this method does not allowed to separate the different Raman lines occurring separately during the acquisition of all spectra.

### 4.2.3 Probability Density Function

The Probability Density Function (PDF) associated to the photon rate is then calculated from histograms created by plotting the occurrence of SERS amplitude versus the photon rate. The PDF is a convenient way to estimate the probability of occurrence of rare but bright events observed in the measurement [25].

The PDF of the spectral intensity is depicted by the colored dots (log-scale) in Fig. 4.4 for a set of around 1024 wavenumber histograms. Typically, ten very intense and narrow Raman bands marked in blue can be precisely determined in Fig. 4.4. An analysis of these selected bands shows a mono-modal intensity distribution except for the 450 and 1300 cm$^{-1}$ bands already identified by the $M$-analysis.

## 4.3 Multivariate statistical analysis based on shape of spectra: PCA

Beyond identifying the fluctuating lines with the M-factor or the PDF, a complementary approach is to sort the spectra into different categories according to their degree of similarity. The identification of the number of families of independent spectra provides a deeper insight into the source of the fluctuations. This decomposition is conducted through a statistical processing of the data by using principal component
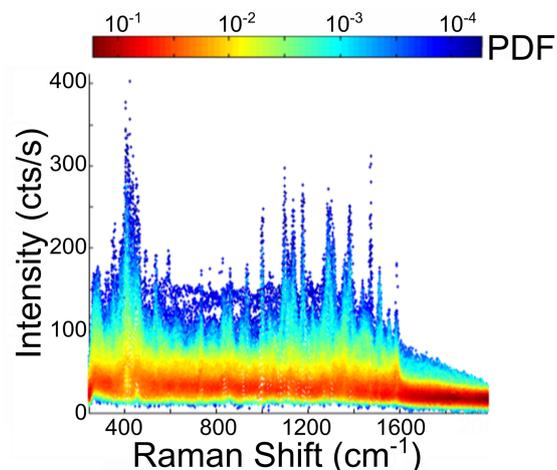
Figure 4.4 – Representation of the probability density distribution of the spectral signature. The red spectra are the most probable but contain very little information. Spectra containing intense characteristic line conversely have the lowest probability of occurrence.
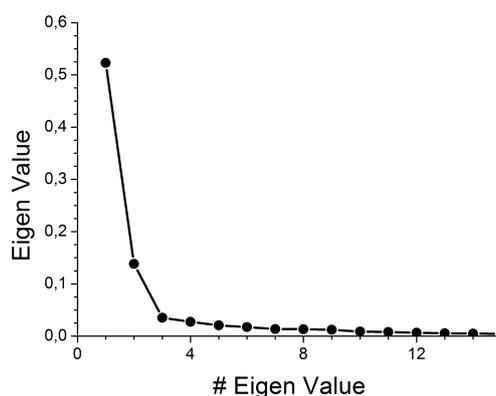


Figure 4.5 – Eigenvalue decomposition given by the principal component analysis of Cysteine.

analysis (PCA) [49].

The whole series of 2032 spectra is used to compute the covariance matrix and deduce the principal components (PCs). Orthogonalization of the covariance matrix shows that the experimental spectra of cysteine can reasonably be described by multiple eigenvectors, the eigenvalues give the weights of each eigenvector (or principal component PC). The amplitude of the eigenvalues plotted in Fig. 4.5 shows clearly that the first principal component represents 52% of the data and the second PC 15%, whereas the weight of the third PC is well below 5%. Two PC can be reasonably
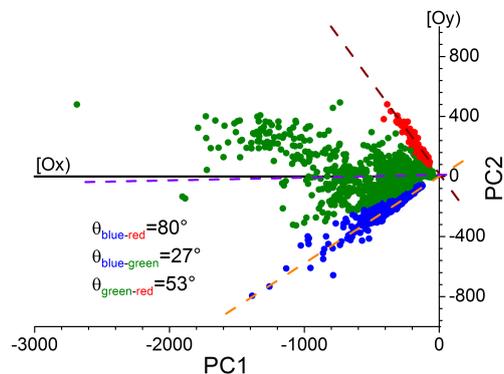
Figure 4.6 – Projection of the experimental spectra into the plane of the two main principal components PC1 and PC2. Spectra are sorted in three families colored in blue, red and green depending on their similarities in angle, as shown on Fig. 4.7.
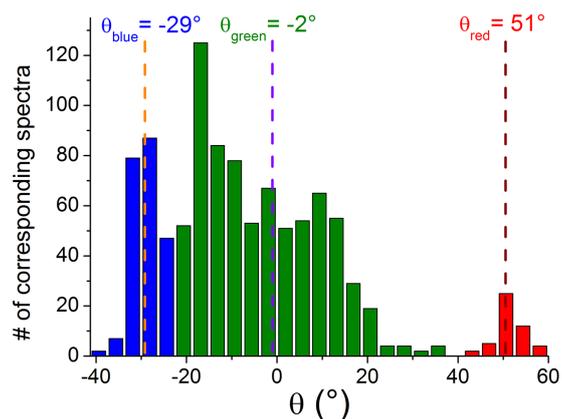


Figure 4.7 – Angular distribution $\theta$ of the spectra with respect to PC1 in the scatterplot shown on Fig. 4.6.

representative of most of the data.

The two-dimensional principal component space is represented in the scatterplot of Fig. 4.6. Each experimental spectrum is represented by a point on the plot. The coordinates of the points correspond to the projected contribution along the first principal component axis (PC1, abscissa) and the second principal component axis (PC2, ordinate). For this particular amino acid, the data are spread out into three separate branches passing through the origin O. Each point is then located in polar coordinate by the angle $\theta$ to the PC1 axis and a distance r from the origin O.

Figure 4.8 – (a) Coloring of Fig. 4.1 according to the outcome of the PCA sorting of Fig. 4.6. (b) Insets show selected time-lapse of blue, red and green events with a bin width of 1 s and a temporal resolution of 1 ms.

Fig. 4.7 shows a histogram of the distribution of angle. A color code is used for differentiating the three families of spectra. A clear distinction is made between the narrow peaks centered at $\theta_1 = -29°$ assigned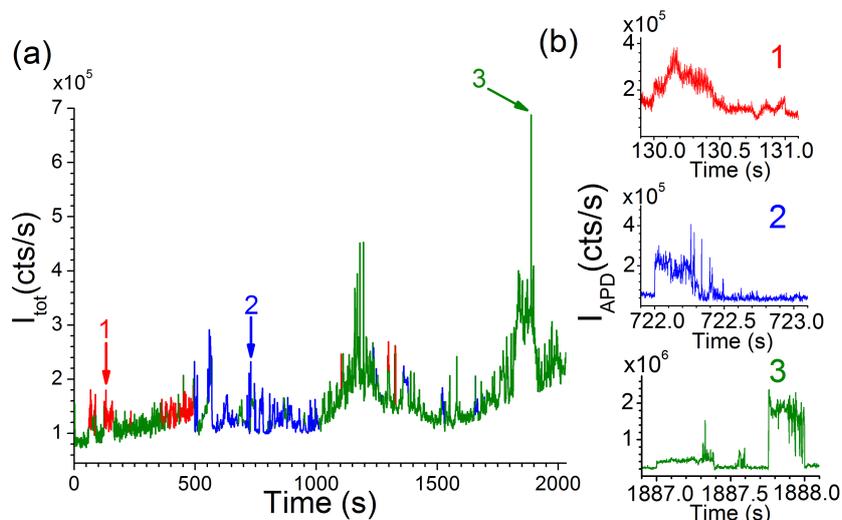 to the blue branch B and $\theta_2 = 51°$ related to the red branch R. A broader distribution located in between is shown in green (G). The angle between the blue and red distribution is close to the right angle, which means that families 1 and 2 are uncorrelated. Conversely, the statistics of the green family can be view as composed of three overlaid shifted distributions. The asymmetry of the histogram also reveals that the correlation with the blue family is more marked than that with the red one. The points aligned along the blue or red branches ($\theta$ fixed) provide information about the intensity fluctuation of the spectra. The Raman activity is here governed by both the local enhancement and the Raman modes of the amino acid.

## 4.4   Characteristic times determination

Having established that all spectra belong to one of the three RGB families (Fig. 4.6), the kinetics of Fig. 4.1 (a) is reconsidered by coloring each Raman event as shown in Fig. 4.8 (a). The graph demonstrates the importance of acquiring long kinetics. The colored events are typically occurring in bunches separated in time. The predominance of the green family with intense and long-lived events is confirmed compared to the punctuated occurrences attributed to the red and blue families. The insets of Fig. 4.8 (b) display representative kinetics of the three families spanning 1 s along the time trace.

These time-lapses of 1 s correspond to the acquisition length of the SERS spectra displayed in the insets of Fig. 4.1 (b). The red and blue traces share some similarities. In particular both graphs display a relatively long duration of the SERS signal. Conversely, the green trace is characterized by multiple events varying in both duration and intensity. The step-like nature of the time traces provides strong evidence of the detection of the single molecules. The residence time of single molecules on the active site is estimated between 300 and 400 ms. From Fig. 4.6, spectra were sorted into three families shown in blue, red, and green depending on their similarities. Once this first sorting is done, a reference spectrum for each family can be derived from averaging all experimental spectra of the same family. Panels b and d (red and blue states) of Fig. 4.9 represent typically the reference spectra of the red and blue states. The kinetics indicates also a mixture of multiple scattering events as seen on Fig. 4.8 (green family).

## 4.5   Spectral fingerprint interpretation

### 4.5.1   Molecular assignment of Raman peaks

The Raman vibration assignment remains the trickiest step with the classical tools of chemiometry for four main reasons: first due to the differences between Raman spectra in solution being less resolved and the SERS spectra that exhibit usually fewer bands but are better defined; second due to selection rules induced by the surface; third SERS spectra in the literature are not very similar to each other since all bands are not observed by the whole SERS community; and last, for now, freeware simulations based on DFT calculations considering the metal−molecules interactions provide only qualitative support with some possible similarities, nevertheless the assignment depends strongly on the chosen database.

Compared to the literature, we note that our identified spectra are the richest in content, as a result of the high enhancement factor of our SERS substrate. Our approach was first to find out the most probable conformation of the cysteine adsorption onto the gold surface. Eighty percent of Raman bands of the red spectra have been attributed to reported vibrations in the molecule (Table 4.1). If all peaks have been assigned chemically, the intensity ratios can differ from the literature due to the molecular orientation toward the electrical field $\vec{E}$ normal to the surface of the GNPs. The SERS selection rules correspond to the alignment of the resulting electric dipole of the molecular vibration and $\vec{E}$ at the single molecule level as first pointed out by the pioneer work of Moskovits [83], namely when the polarization states of the incident and scattered waves are perpendicular to the scattering plane.

Table 4.1 – Raman assignment of the characteristic spectra based on literature.

| Cysteine $(cm^{-1})$ | Cystine $(cm^{-1})$ | Assignment | | Reference | $I_{exp}$- Cysteine | $I_{exp}$- Cystine |
|---|---|---|---|---|---|---|
| 267 | | Au-S stretching | 1s | [73] | 47 | |
| | 406 | S-S bending | 2b | | | 88 |
| 449 | 460 | CCC bending | 5b | [72] | 52 | 69 |
| 524 | | | | [72, 73] | 31 | |
| | 552 | S-S stretching | 2s | [72, 73] | | 44 |
| 567 | | | | | 34 | |
| 623 | 623 | COO⁻ wagging | 10w | [82] | 44 | 42 |
| 667 | | C-S stretching | 3s | [72, 73, 82] | 34 | |
| 730 | 730 | | | [82] | 32 | 47 |
| 789 | | COO⁻ bending | 10b | [82] | 36 | |
| 832 | 849 | S-CH₂ bending | 3b | [72, 73] | 38 | 42 |
| 860 | | C-COO⁻ stretching | 9s | [72, 73] | 36 | |
| 931 | | CH₂ bending | 4b | [72, 73, 82] | 43 | |
| 960 | | | | | 40 | |
| 1002 | 1002 | NCH stretching | 7b | [72, 73] | 37 | 40 |
| 1058 | 1030 | C-N stretching | 7s | [72, 73] | 63 | 36 |
| 1078 | | | | [82] | 39 | |
| 1123 | 1120 | | | | 56 | 39 |
| 1140 | 1140 | NH$_3^+$ deformation | 8d | [82] | 40 | 37 |
| | 1184 | CH₂ bending | 4b | [72] | | 33 |
| | 1261 | | | | | 34 |
| 1297 | 1285 | CCH bending | 9b | [72, 73] | 63 | 32 |
| 1318 | | CCH deformation | 9d | [82] | 61 | |
| 1354 | | CH₂ bending | 4b | [72, 73] | 46 | |
| | 1357 | CCH bending | 6b | [72, 73] | | 36 |
| 1402 | 1386 | COO⁻ stretching | 10s | [72, 82] | 42 | 31 |
| 1443 | 1460 | CH₂ bending | 4b | [82] | 50 | 27 |
| 1472 | | CH₂ bending | 4b | [72, 73] | 58 | |
| 1516 | | | | | 48 | |
| 1568 | 1583 | NH$_3^+$ bending | 8b | [73] | 46 | 28 |
| | 1765 | | | | | 20 |

The Raman spectrum is composed of the association of different vibrations allowing the orientation of the global dipolar moment along the electromagnetic field. The sulfur atom is by far the most favorable anchor to gold. Each bound was considered and allocated in ascending number as it moves away from the sulfur anchor and is subject to vibrations which are composed of lateral modes (bending, b; wagging, w) or a longitudinal mode (stretching, s) aligned along $\overrightarrow{E}$. The most intense peaks 6s
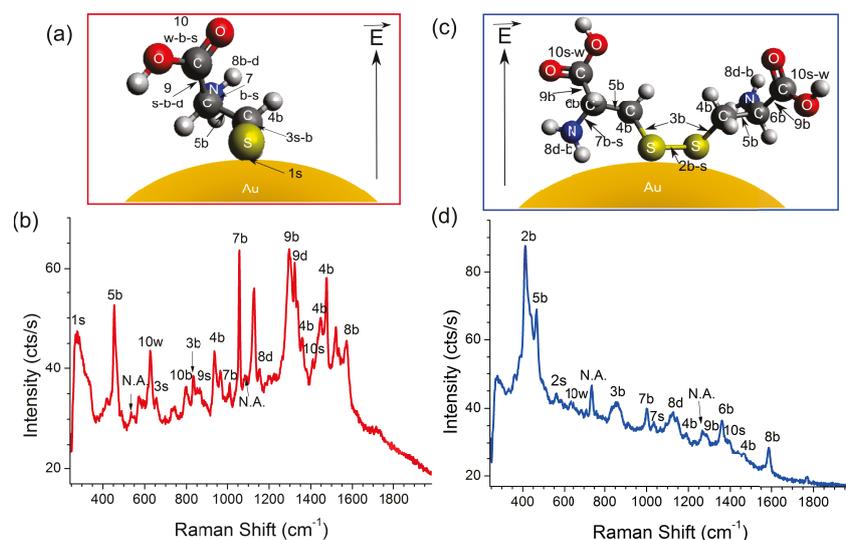
Figure 4.9 – Schemes of the molecular interactions with gold (a,c) and associated characteristic spectra (b,d) for the two molecular configurations : cysteine (a-b) and cystine (c-d). A localization of the different vibrations highlighted in spectra and associated to their references in the Table 4.1 is given by the peak assignments.

(1058 $cm^{-1}$), 5b (1297 $cm^{-1}$), 3b(1472 $cm^{-1}$) can be explained because the vibration is close to the surface exposed to a strong electric field gradient. A possible graphical representation for an educational point of view was obtained by using the Avogadro v.1.1.1 free software as shown in Fig. 4.9 (a). This optimization is based on the Universal Force Field, in which the atomic potential is calculated based on the element, the hybridization, and its connectivity.

Interpreting the blue spectrum has been a most challenging task. The different configurations of adsorption have been considered and compared to the experimental data depending if the amino acid is free or linked to the gold surface by its amine or carboxylic terminals. Anchoring of the cysteine by the amine group is unlikely since the characteristic lines at 497 and 1560 $cm^{-1}$ are absent in both blue and red spectra. Anchoring of the cysteine by the carboxylic group is not realistically feasible since the characteristic lines at 903 $cm^{-1}$ are missing. We found a very good correlation between the blue spectrum and another singular form of the cysteine. Two cysteines can indeed oxidize to form a dimer, named cystin, through a disulfide linkage. Experimentally, the broad peak centered at 425 $cm^{-1}$ may be attributed to S−S vibrations. Considering the SERS selection rules, a conceptual graphical representation plotted by using Avogadro is shown in Fig. 4.9 (c). The vibration modes at 450 and 1300 $cm^{-1}$ are common for all conformations and are therefore particularly strong on the green spectrum. This is consistent with the analysis of the Mandel factor and the probability
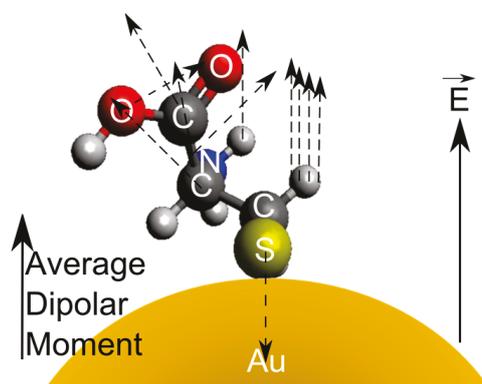
Figure 4.10 – Association of dipolar moments for each Raman band shown in the cysteine spectrum with the molecular scheme of the gold-cysteine complex.

density function shown in Figure 4 that identified these peaks as highly fluctuating.

According to this spectral analysis, we conclude that the blue family corresponds to a thiol-adsorbed cysteine configuration and the red family to the cystin configuration.

## 4.5.2   Molecular orientation vs spectral fingerprint

The orientation of the molecule then can be determined by assuming that except for the gold−sulfur bond the lateral modes are favored over normal modes. A possible orientation of the molecule was found. A preliminary interpretation is to associate an electric dipole moment to each vibration. For each band, the direction of this moment is determined by the angle between the vibration and the gold surface, and the norm is given by the square root of the intensity of the measured band. Applying this dipolar interpretation on the semi-empiric molecular orientation of the gold-cysteine complex, I obtain the Fig. 4.10. The average dipolar moment resulting from this model is finally aligned with the electromagnetic field coming from the gold substrate.

This explains also the spectral composition. Indeed, corresponding to a spectrum is composed by the Raman bands, global vibration of the molecule are in the axis of the electromagnetic field. This interpretation is in total agreement with the physical interpretation of the enhancement in SERS, first pointed out by Pr. Moskovits in 1985 [83]. The enhancement factor depends on the polarization of the electromagnetic field: the enhancement is higher along the electromagnetic field.

Nor can the chemical contribution in the enhancement in SERS be neglected since the distance between the adsorbed molecule sounded and the metallic substrate is crucial because of the close localization of the electromagnetic field around the nanoparticle [43].

## 4.6 Conclusion

As shown, the developed setup coupled with the highly SERS active substrates result in a very well method for SERS studies, and more particularly when statistical analyses are applied. The interest of PCA is clear in such work to differentiate the substrate component and the different species present in solution. Coupled with a fine spectral analysis, the cysteine orientation during its interaction with the gold substrate has been determined according to the physical origins of SERS effect. This has been able by the high concentration resolution reducing the number of molecule sounded.

However, all of this work has been done on small biomolecules. The main interest is to detect and identify without chemical functionalization the presence of specific proteins as HSP-70 in serum.

# CHAPTER 5

## Complex media / biomolecules / systems

## 5.1 Introduction

A complex medium is constituted by a large number of smaller non complex entities. This is especially the case as regards the structure of proteins built from twenty amino acids, or the blood constituted of many proteins ... The level of complexity is proportionate to the dependence criterion imposed between entities. Typically, in solid state physics, for a totally stiff material, only few variables (inertia, velocity) shall be sufficient to determine the motion of each entity.

Nevertheless, when one considers a vibrational analysis, it can get a lot more complicate since each entity can have a motion different from others. This also applies to fluids or soft matter where the number of state variables extends "ad infinitum". The entities in biology are usually independent which means that the knowledge of one entity tell us nothing about the neighboring parts.

At that point, there will always be some subjectivity or obscurities in the practical considerations in determining the level of independence of the entities. The same medium can be considered equally extremely complex because unexplained or very simple in merely representing "shallow" explanations.

In this chapter, four degrees of complexity for macromolecule analysis will be investigated.
First, a well-studied protein will be the target molecule: the Bovine Serum Albumin, the most widespread protein in bovine blood. This is an interesting protein since its structure in $\alpha$-helix is quite simple with very rare structural modifications. The key
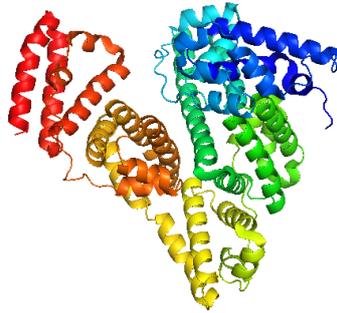
Figure 5.1 – BSA 3D representation.

issue for such protein is to determine if the characteristic spectrum is the sum of contributions from all amino acids or mainly from the most resonant one, and if we should obtain a specific spectrum for each molecular orientation on the surface. These questions immediately come to mind in the light of the large spectral variation found in literature for this single protein [6, 7, 8, 9, 10, 11, 12, 84, 85, 86] (Fig. 1.8).

Secondly, we will add difficulties by studying a multidomain protein, i.e. a protein with different parts (peptides) moving the one with the others thereby giving different conformations. The chosen protein is Heat Shock Protein-70, which is an interesting protein for early stage cancer detection. Indeed, this protein is overexpressed in blood when cancer attacks a cell.

The idea is to investigate the capability of our sorting method developed in Chapter 4 for retrieving molecular conformation in a highly complex medium.

## 5.2   Monodomain protein: BSA

First, we apply the method developed in Chapter 4 on a well-known protein in literature: BSA, or Bovine Serum Albumin. The BSA is the most widespread protein in the blood serum of bovine. Its 3D chemical representation is on Fig. 5.1.

The equivalent protein for human is the HSA. It is defined as a monodomain protein. Indeed, the BSA is composed of a single peptidic chain, i.e. a single chain of amino acids. Its structure is quite complex, as the main proteins, but is relatively compacted. BSA, which is a reference in the protein investigation by SERS in literature [6, 7, 8, 84, 85, 86].
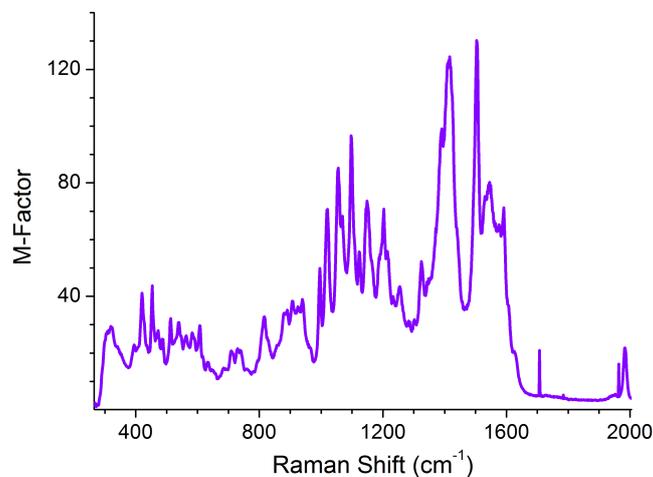
Figure 5.2 – Mandel M-Factor determination for a long kinetic of BSA study.

## 5.2.1 Experimental method

The experimental conditions for BSA and cysteine are quite close to the one of cysteine investigation. BSA from the INSERM of Dijon was diluted in PBS at a concentration of 50 $\mu$g/mL ($\approx$ 0.77 $\mu$mol/L) and was injected in a microfluidic channel with a flow of 2 $\mu$L/min driven by the peristaltic pump. Many records have been done on different hot spots with the same laser power and different times of record to test the reproducibility of the protein study with this system. It seems that the reproducibility is occurred for this molecule at this high concentration. The Fig 5.3 (a) shows the adsorption kinetics of BSA during one hour and half.

**An easy statistical approach: Mandel M-Factor determination**

The use of the Mandel M-Factor is interesting to highlight the principal Raman lines occurring during the BSA study on the different hot spots, i.e. with different enhancement configurations. The resulting spectrum is given on Fig 5.2. The comparison between its Raman lines decomposition and different works in literature for BSA [6, 7, 8, 84, 85, 86] implies the identification of the vibrational origins composing the Mandel spectrum. It principally results from Phenylalanine amino acid or from $\alpha$-helix proof (S-S bond, C-S stretching, $\alpha$-helix global vibration ...) (Table 5.1).

The Mandel spectrum is a global spectrum of BSA adsorbed on gold but without conformational modifications. A large number of Raman bands can be identified based on literature. This can be interpreted has a classical configuration of BSA observation in SERS. Another issue is the richness of the Mandel spectrum. This response provides

substantially much more information than in the literature. This is firstly explained by the size of the spectral window chosen. A second potential explanation is contributions of different conformations of single protein on the gold surface. The vibrations are somewhat different but the global fingerprint of the protein is still closely the same.

Some Raman bands present relatively high Mandel values. These fluctuations should provide from different spectra appearing in the long record.

**Spectral sorting by PCA**

To check the complexity of the different spectra and differentiate them in a long record, the application of PCA on the record seems to be a good solution. The interest of this method is to identify the similarities and differences between Raman bands when BSA changes its conformation on the gold surface.

This Principal Component Analysis begins by the eigenvalues plot to determine the number of principal components necessary to well describe the time series. Fig. 5.3 (b) shows that at least three eigenvalues are required. The third eigenvector represents indeed slightly more than 5 % of the variance of the signal.

Spectra have been then decomposed in 3D space on the three first principal components (Fig 5.3 (c)). The crosses represent the projection in the basis of three principal components. The 3D representation provides interesting separation between the different spectral families. Three branches are so observed and can be easily differentiated with a point of view close to the (Oxy) plane. However, it is difficult to sort the spectra on the basis of their angular distribution in the three-dimensional space.

A clearer and complete view was found when projecting the spectra in the 2D eigenspace formed by the two first principal components. The 2D points are represented by circles on Fig. 5.3 (c). The 2D projection is still efficient to differentiate the spectral families forming by the three branches. The angular distribution analysis, presented in Chapter 4 has been applied to sort the different spectral families and color them. The green case is a mix configuration between the blue and the red families.

The spectra are then localized in the time series and their associated families identified with the different blue, green and red colors. It results in the Fig. 5.3 (a). The resulting plot allows confirming that the different events described by two states blue and red are self-consistent since the blue and the red events repeat alternatively. Note that the duration of events is longer compared to the cysteine case. This is certainly due to the high affinity between the BSA protein and the gold surface. In fact, the a same BSA protein can stay more than hundreds of seconds in the hot spot as noticed by red events at 1000 and 2700 seconds and blue events at 0 and 1500 seconds.
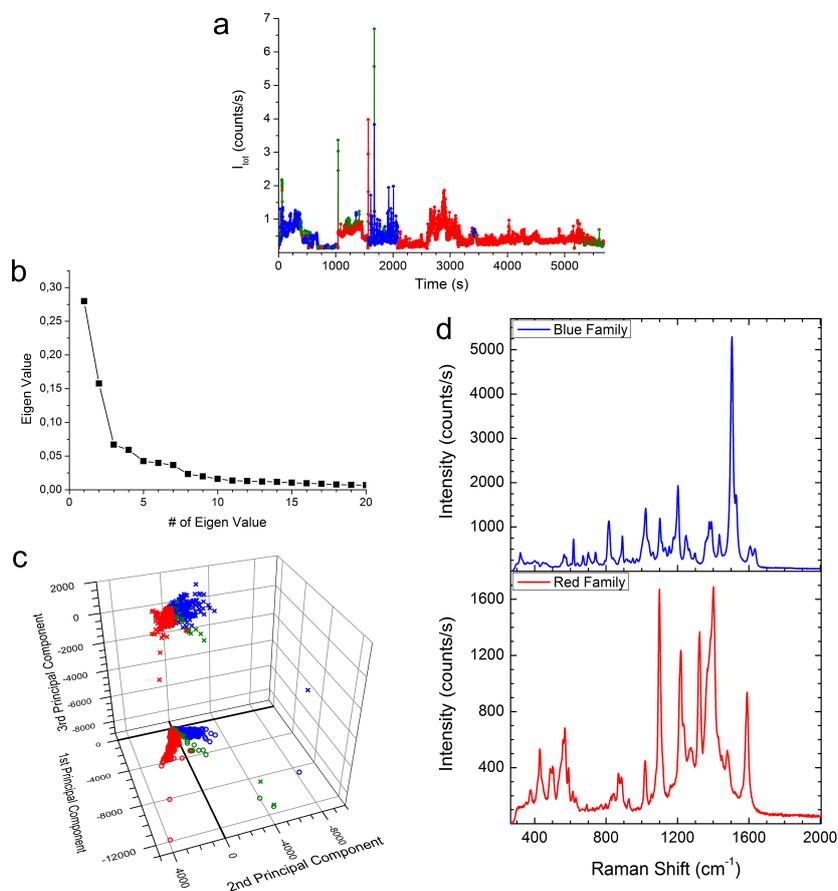
Figure 5.3 – (a) Kinetic of the BSA acquisition. Colors are given by the 2D projection of (c). (b) Eigenvalues distribution for the long BSA kinetic. (c) Spectral projection of BSA acquisition on its three first principal components. The projection on the two first principal components purposed in the (Oxy) plane allows identifying spectral families. Crosses represent the projection in the 3D basis and circles the projection in the 2D basis. (d) Characteristic spectra of BSA based on the average of the spectral response of each family identified by PCA.

**Spectral fingerprints analysis**

Now that two spectral families have been identified by PCA, their reference spectra were extracted and compared to the SERS literature [6, 7, 8, 84, 85, 86]. Each family can be associated to a specific spectrum as shown in Fig. 5.3 (d) resulting from the average of the spectra composing the corresponding family. Raman bands are then chemically assigned on the base of the literature [6, 7, 8, 84, 85, 86] (Table 5.1).

Table 5.1 – Raman assignments of the characteristic spectra of BSA based on literature.

| | M | Blue Family | Red Family | Reference | Identification |
|---|---|---|---|---|---|
| 319 | m | m | | | |
| 343 | | | w | | |
| 391 | w | m | | | |
| 420 | s | | s | | |
| 437 | | m | | | |
| 453 | s | | m | | |
| 469 | w | | w | | |
| 488 | w | | | | |
| 502 | | w | | | |
| 520 | m | w | m | [84] | S-S bond |
| 538 | m | | m | | |
| 552 | | w | | | |
| 563 | m | w | w | | |
| 578 | m | | w | | |
| 590 | | | w | | |
| 607 | m | | m | [85] | Phenylalanine |
| 622 | w | w | w | [85] | Phenylalanine |
| 655 | | w | w | [85] | C-S stretching |
| 678 | | w | w | | |
| 711 | w | w | w | [8, 84] | C-S stretching |
| 740 | w | m | w | [85] | $COO^-$ def. of backbone |
| 812 | w | s | w | [6, 85] | Tyrosine |
| 855 | | vw | m | [84] | Tyrosine |
| 886 | m | m | w | [6] | Tyrosine |
| 907 | | w | w | | |
| 928 | m | w | | [84] | Undefined |
| 964 | s | | w | [85] | C-C stretching |
| 1022 | s | s | s | [6, 7, 8, 84, 85] | Phenylalanine |
| 1057 | s | m | m | [6, 7, 8, 84, 85] | Phenylalanine |
| 1069 | | w | | [8, 84] | Undefined |
| 1098 | s | m | s | | |
| 1109 | | | w | | |
| 1123 | m | | w | | |
| 1153 | s | s | w | [86] | Tyrosine |
| 1203 | s | | s | [7, 84] | Phenylalanine |
| 1215 | | m | w | [84, 86] | Phenylalanine or Tyrosine |
| 1233 | | | w | [84] | Undefined |
| 1257 | m | w | | [7, 8, 85] | Random coil |
| 1292 | | w | w | [85, 86] | $\alpha$-helix |
| 1325 | m | | m | [7, 86] | Tryptophane |
| 1374 | s | s | | [85] | Tryptophane |
| 1410 | s | | s | [84, 86] | $COO^-$ |
| 1439 | | s | | [8, 84, 86] | $\delta(CH_2)$ |
| 1501 | s | s | m | [8] | Undefined |
| 1530 | s | w | | [8] | Undefined |
| 1595 | s | vw | s | [84, 85, 86] | Phenylalanine or $COO^-$ from backbone |
| 1626 | w | vw | | [7] | Amide I |

These spectral identifications are very interesting to understand the origins of the spectra and the different amino acid contributions. Also, it is clear that the amino acid Phenylalanine is the major contributor to the spectra. This is due to its resonant radical principally formed by a benzene cycle. The question is why does the other benzenoic amino acid Tryptophane remain quietly invisible in the BSA spectra? In
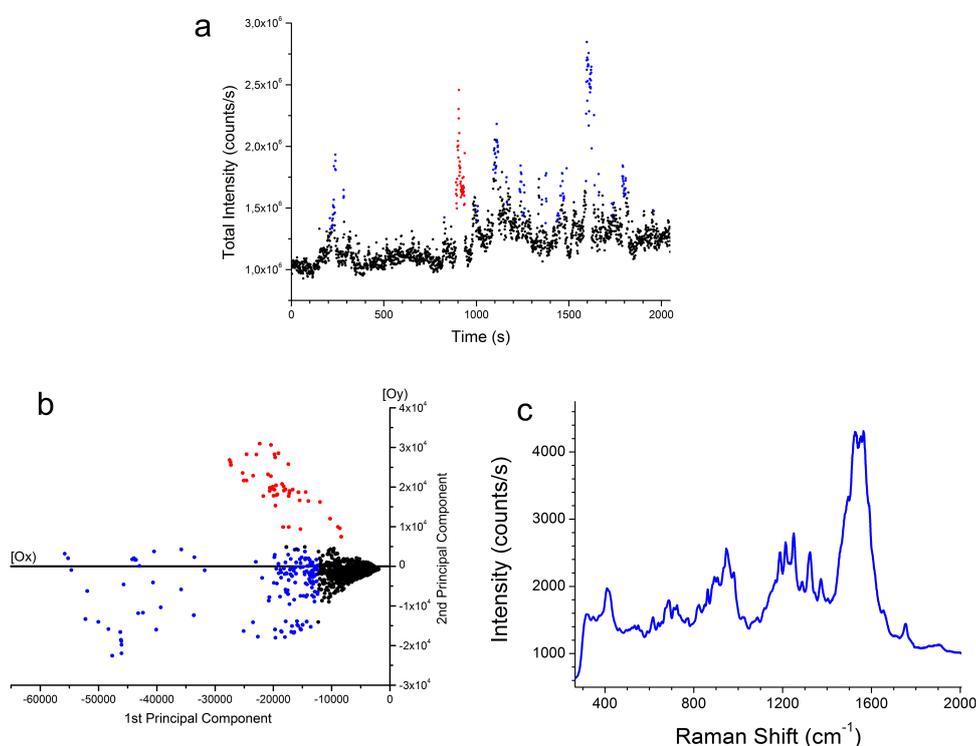
Figure 5.4 – (a) Total SERS intensity of BSA recording during 2048 seconds versus the time of acquisition and colored on the base of (b).(b) Projection of each spectrum recorded for BSA in the eigenspace formed by the two first principal components of this record. Colors are affected by position in this space, i.e. by spectral similarities. (c) Reference spectrum of blue family identified by PCA for BSA protein.

fact, there are only two Tryptophanes in BSA structure, and one of them is blocked in the heart of the protein. One other specific spectral origin is the secondary structure of BSA: the $\alpha$-helix. In fact, BSA has principally a $\alpha$-helix secondary structure. The differences between spectra are thus not really decisive and only depend on the different orientations of the protein towards the gold surface.

## 5.2.2   Spectral proof of the unfolded form of the protein

In the previous section BSA has been characterized by a specific Raman fingerprint corresponding to a well-known form of the protein at this pH of 7. However, another form described by another SERS spectrum has been isolated among the many time series (Fig. 5.4 (a)). We apply so the PCA on this record of 2048 seconds.

Table 5.2 – Raman assignments of the Blue family characteristic spectra based on BSA spectral fingerprint.

| Raman Shift (cm$^{-1}$) | Relative Intensity | BSA identification |
|:---:|:---:|:---:|
| 319 | w | ✓ |
| 347 | w | ✓ |
| 408 | m | ✓ |
| 462 | w | ✓ |
| 540 | w | ✓ |
| 616 | w | ✓ |
| 640 | w | |
| 685 | w | ✓ |
| 722 | w | ✓ |
| 768 | w | |
| 822 | w | ✓ |
| 861 | w | ✓ |
| 900 | w | ✓ |
| 946 | m | |
| 978 | w | ✓ |
| 1018 | w | ✓ |
| 1084 | w | ✓ |
| 1119 | w | ✓ |
| 1187 | m | |
| 1212 | m | ✓ |
| 1240 | m | ✓ |
| 1288 | w | ✓ |
| 1323 | m | ✓ |
| 1374 | w | ✓ |
| 1403 | w | ✓ |
| 1495 | w | ✓ |
| 1525 | s | ✓ |
| 1585 | s | ✓ |
| 1653 | w | |
| 1751 | m | |

Two eigenvectors were sufficient to describe the record. The eigenspace of projection is based on the two first eigenvectors, i.e. the two first Principal Components of the system as shown in Fig. 5.4 (b). Two spectral families have been then so identified applying the PCA: one in blue and one in red. The time series was then colored according by Fig. 5.4 (b). The duration of the blue events is shorter than the red one and are repeated before and after the red event. This blue family seems to be the family of a natural BSA at pH 7 adsorbed on gold because of the time of residence of the protein on the surface. Its reference spectrum (Fig. 5.4 (c)) has to be compared with the BSA spectra of the literature and the one previously identified. The results are summarized in Table 5.2. Thus, it appears clearly that the reference spectrum of blue family corresponds to a natural form of BSA as previously observed in the previous section.
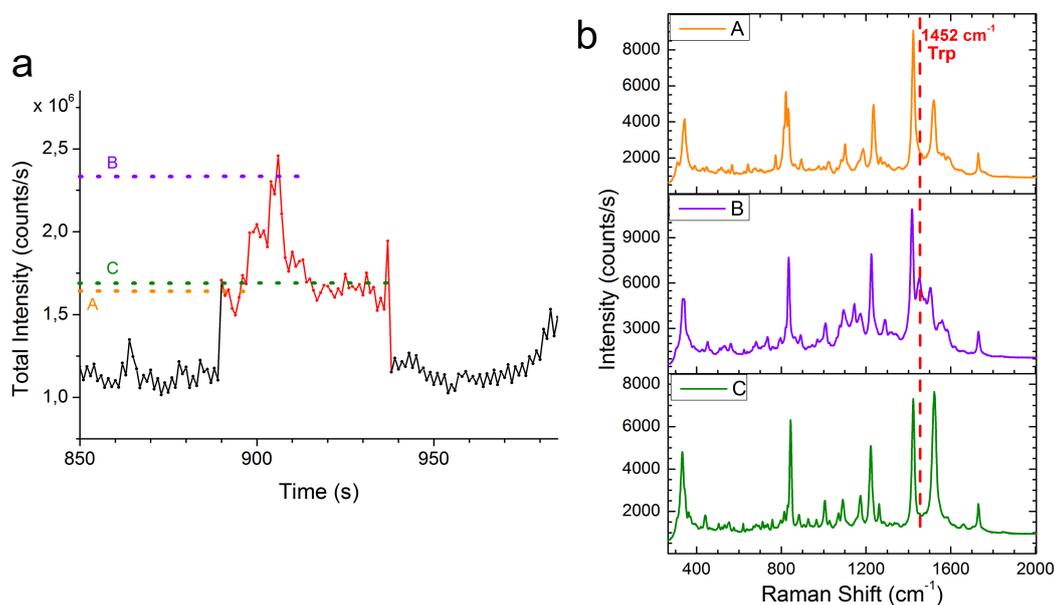
Figure 5.5 – (a) Zoom on the temporal signature of the red event of BSA.
(b) Spectral signatures of different steps of the red event of BSA. Each
spectrum is associated to a form happening in time at different moments
localized on (a).

Concerning the red event, it is a long event (more than 40 s) associated to a completely different reference spectrum than the blue one, and exhibiting an attractive time signature especially in zooming on the red event (Fig. 5.5 (a)). It clearly appears a stepwise process with two stairs. The first stair is characterized by two Raman spectra (Fig. 5.5 (b)) A and C which appear similar. The topmost stair is associated with the spectrum B. Spectra A and B differ very little except the band 1452 cm$^{-1}$. This peak corresponds to a Tryptophane band. Our interpretation was connected to others observations in literature.

In this aim, we looked for such step signature of BSA signal in literature. This research results on one publication of Pang and Gordon in Nanoletters in 2012 [3]. They present the study of the trapping of BSA with a specific nanotweezer based on a double nanohole with an enhancement site in the middle of the structure. They investigated the optical signal transmitted through the structure. The signal is transmitted when the protein of BSA is trapped in the spot because it scattered the laser signal. In kinetic, they obtained the Fig. 5.6 (a).

They saw a step function of the optical transmission signal. They attributed these
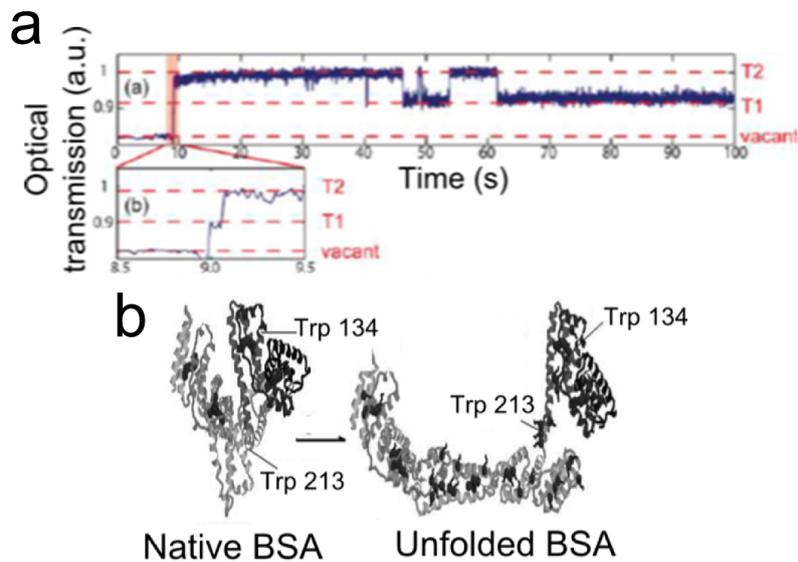
95

Figure 5.6 – (a) Temporal signature of the optical transmission during BSA trapping. Adapted from [3]. (b) 3D representations of the two conformations of BSA: folded and unfolded, and localization of the two Tryptophanes of BSA. Adapted from [87].

steps to two conformations of BSA. The highest transmission corresponds to the natural form of BSA, also called the folded form. The lower step is attributed to the unfolded form of BSA corresponding to the open form of the protein. These two forms are presented on Fig. 5.6 (b). The change of conformation explains the optical transmission modification simply based on the steric hindering modification in the nanohole between the folded and the unfolded forms.

Thus in our case, the step distribution is interpreted as the change of conformation of the BSA adsorbed on gold: a folded form adsorbed on the nanoparticle (case A), then this protein opens itself and enables to spectrally observe the peak coming from the inside Tryptophane (case B) and finally the protein came back in its natural form staying in interaction in the hot spot (case C) before a total desorption.

This unfolded form of BSA is very rare at pH 7. This explains why this red event appears only one time during the 2048 seconds of the acquisition. However, we show for the first time in the literature the two forms of BSA by SERS, using the intrinsic biomarker of BSA: its second Tryptophane that can be sounded only in the unfolded conformation.
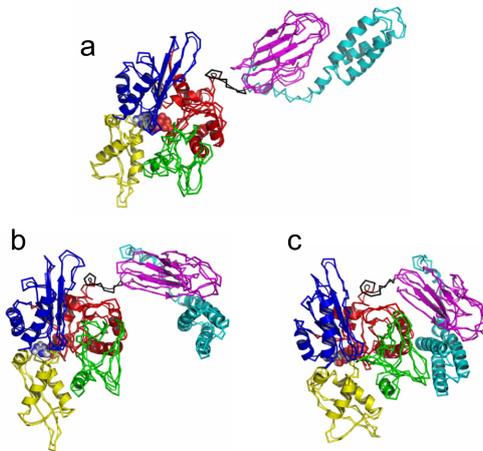
96

Figure 5.7 – 3-Dimensional models of different HSP conformations. Each peptidic chain is modelized with a specific color. Adapted from [1]

## 5.3 Multidomain and complex protein: HSP-70

I have shown that the using setup enables investigating the different protein conformations in the case of BSA. The idea is now to see if, using the same approach, it is able to better understand a protein of unknown structure. The chosen protein is the human Heat Shock Protein-70.

### 5.3.1 Interest of the chosen protein

This biomolecule is a specific one with many interests. Thus, it is a good cancer biomarker, but also a chaperone protein, i.e. a protein that helps another protein or molecule to be modified or transferred through a cell membrane. Such protein is complex with different peptidic domains. These domain arrangements induce conformational modifications due to the motion of each domain against the others. Some 3D models of these conformations have been theoretically calculated and purposed by Nicolai *et al.* [1]. These models are shown in Fig. 5.7.

This molecule is interesting to test the limit of using the developed method for protein identification with very low differences between the different molecules of a same protein in different conformations. In other words, we will look at a pool of different conformations of the same protein.

### 5.3.2 Conformation identification by spectral analysis

To identify different conformations on the base of spectral analysis, a solution of HSP diluted in PBS at a concentration of 50 $\mu$g/mL is passed through a microfluidic
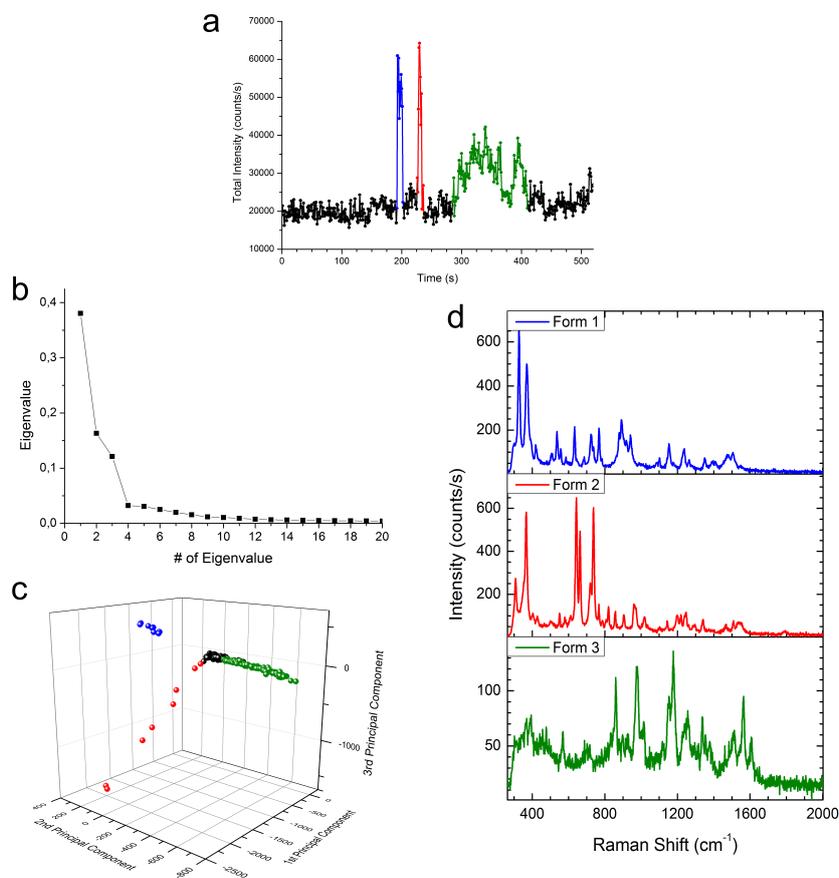
Figure 5.8 – (a) Colored kinetic of HSP-70 study by SERS. (b) Eigenvalues decomposition for HSP-70 study by SERS. (c) Principal Component decomposition of HSP-70 study by SERS. (d) Characteristic spectra identified for HSP-70.

channel with a flow rate of 2 $\mu$L/min.

The developed experimental methodology is applied for HSP study. A global record of 520 seconds long is so obtained and analyzed statistically (Fig. 5.8 (a)). The application of the Mandel analysis is unusual because of the mix of conformations composing the HSP-70 solution. It is the same case for the PDF analysis. That is why PCA is the statistical method directly chosen to analyze this dynamic record.

The application of PCA results in the eigenvalues decomposition (Fig. 5.8 (b)). Three eigenvectors are necessary to fully describe the series. Each spectrum was decomposed on a three dimensional eigenspace defined by the three first Principal Components of the record (Fig. 5.8 (c)). Three well-defined branches can be so dis-

tinguished. Specifically, two types of branches are present. First, the green branch is composed by a large number of spectra, at the opposite of the red and blue branches, which are composed by only few spectra. To check the lifetime of each form, the associated kinetic trace is shown on Fig. 5.8 (a).

The two shorter families (blue and red) are each composed by only one event very intense. The third one (green) is also composed by one event, but a very long one larger than one hundred seconds. In terms of physical effect, these different spectra should be the consequence of the interaction between different forms of HSP-70 with the metallic substrate. As still shown on Fig. 5.7, three conformations exist for this protein. The same number of forms identified in the experiment equal to the number of conformations of the protein should not be a coincidence. Each spectral family may be associated to a specific 3D conformation of the protein.

Each spectral family can be characterized by its own spectral fingerprint. These different fingerprints are summarized in Fig. 5.8 (d) and Table 5.3. They are obtained by the calculation of the average spectrum for each family. Each reference spectrum is clearly different to another as shown by Table 5.3. Unfortunately, the exact peptidic structure of this protein is no totally known. It becomes so impossible to ensure a correct assignment of the different Raman bands of each characteristic spectrum. However, these fingerprints can be used as the fingerprint of HSP-70 in a mix of proteins as blood serum.

If we go further in the reference spectra interpretation, we can identify similarities between the HSP-70 spectra and BSA fingerprint. Indeed, these two molecules are proteins and it is important to remember that proteins are composed by only twenty different amino acids. These similarities between fingerprints are symbolized in Table 5.3 by blue lines. Thus, a lot of similarities appear probably directly due to the peptidic combination, i.e. the nature of amino acid composing the protein. In a so big protein (around 620 amino acids), each amino acid is present at least one time. It explains so the similarities. In the opposite, some large part of HSP-70 spectra cannot be assigned based on BSA fingerprint. This is notably the case of Raman bands for high Raman shifts (more than 1450 cm$^{-1}$). This can be easily explained by the global form of the protein. Indeed, it is known that in this range of Raman shifts appear Raman bands coming from the secondary of the protein structure. However, BSA is only composed by $\alpha$-helix, what is not the case of HSP-70 composed by $\alpha$-helix but also $\beta$-sheets, $\beta$-turns... It is so possible to distinguished two different proteins simply based on their SERS fingerprints in spite of the large spectral response of amino acids inducing many similarities in fingerprints.

Table 5.3 – Raman assignments of the characteristic spectra of HSP-70. In blue are colored Raman bands occurring also in BSA fingerprint determined previously in Table 5.1.

| Raman Shift (cm$^{-1}$) | Form 1 | Form 2 | Form 3 |
|---|---|---|---|
| 308 | | m | |
| 315 | | | m |
| 326 | s | m | |
| 345 | | | m |
| 368 | s | s | |
| 395 | | | m |
| 418 | m | | w |
| 459 | | w | |
| 478 | | | w |
| 504 | m | w | w |
| 536 | m | | w |
| 550 | | w | |
| 560 | m | | |
| 567 | | | m |
| 580 | | w | |
| 588 | w | | |
| 603 | | | w |
| 610 | | w | |
| 640 | m | s | w |
| 663 | | w | |
| 684 | w | | |
| 715 | | m | m |
| 725 | m | | |
| 738 | | s | |
| 768 | m | w | |
| 786 | w | w | w |
| 819 | | m | |
| 833 | | | w |
| 859 | | m | s |
| 887 | s | | |
| 905 | | m | m |
| 918 | w | | |
| 943 | m | | |
| 972 | | m | m |
| 1005 | w | | m |
| 1019 | | w | |
| 1057 | | | w |
| 1102 | w | | |
| 1123 | | | w |
| 1143 | | m | |
| 1154 | m | | |
| 1178 | w | | s |
| 1198 | | w | |
| 1221 | | w | |
| 1240 | m | w | s |
| 1267 | w | | |
| 1292 | | w | |
| 1338 | | w | |
| 1349 | m | | |
| 1377 | | w | m |
| 1392 | w | | |
| 1412 | w | | |
| 1469 | m | w | |
| 1509 | m | w | w |
| 1545 | w | | |
| 1562 | | | s |
| 1607 | | | w |
| 1781 | | s | |
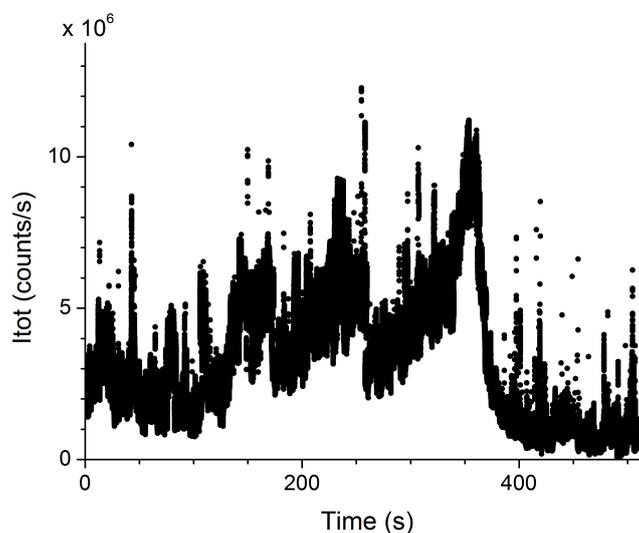
Figure 5.9 – Kinetic of human blood serum studied by SERS.

## 5.4  Complex medium: case of human blood serum

Before investigating the presence of HSP-70 in blood serum, it is important to understand how a complex medium as blood serum can spectrally react in our experimental case, i.e. which types of information can be extracted from a specific experiment on this complex solution.

In this aim, a solution of human blood serum has been injected in one microfluidic channel and the signal coming from one SERS hot spot in this channel has been recorded during 520 seconds. In terms of temporal evolution, it results in Fig. 5.9. This temporal evolution shows hundreds Raman events more or less long. This dynamic fingerprint is so very rich and complex.

As in each previous experiments, spectra are recorded simultaneously on a spectrometer with a time resolution of 1 spectrum per second. It is so possible to image the evolution of the Raman spectra during the record (Fig. 5.10). Based on this figure, it is clear that there are important fluctuations in terms of intensity but also of Raman bands. It seems so very important to applied PCA to sort the different spectral fingerprint acquired during the record.

The use of the PCA on this record firstly results in the eigenvalues decomposition shown in Fig. 5.11 (a). Simply on the basis of this figure, it is clear that PCA is not really efficient to sort spectra in this complex case. Indeed, it seems that at least 6 eigenvalues are necessaries to well describe the complexity of the record, i.e. that 6 eigenvectors are needed to correctly decompose each spectrum. A 6-dimensional

101
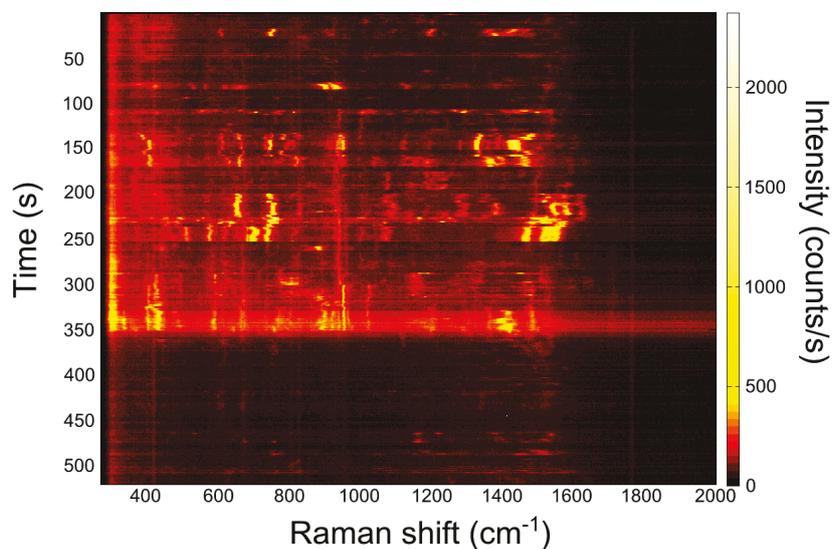
Figure 5.10 – Time evolution of the Raman spectra during the record of SERS of serum.
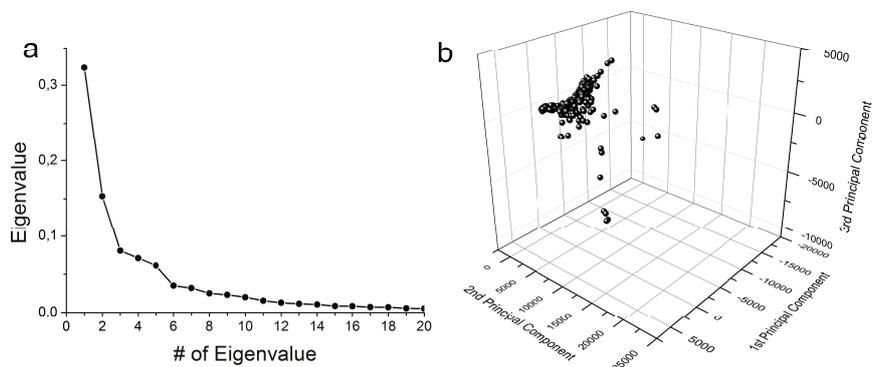


Figure 5.11 – (a) Eigenvalues decomposition for human blood serum studied by SERS. (b) Principal Component decomposition of human blood serum studied by SERS.

projection is so necessary but unable.

However, a projection based on the three first eigenvectors is shown in Fig. 5.11 (b). Some branches seem to be highlighted but a complex cloud of spectra appears for lowest coefficient without any specific angular distribution. This decomposition confirms that the application of PCA on a raw complex medium like human blood serum is inappropriate in very complex solution.

Figure 5.12 – (a) Eigenvalues decomposition for the mix serum-HSP-70 studied by SERS. (b) Principal Component decomposition of the mix serum-HSP-70 studied by SERS.

## 5.5 Complex protein in complex medium: case of HSP in serum

The detection of HSP-70 at a very low concentration (around few nanogramms per milliter) in blood serum is the main goal of the Eu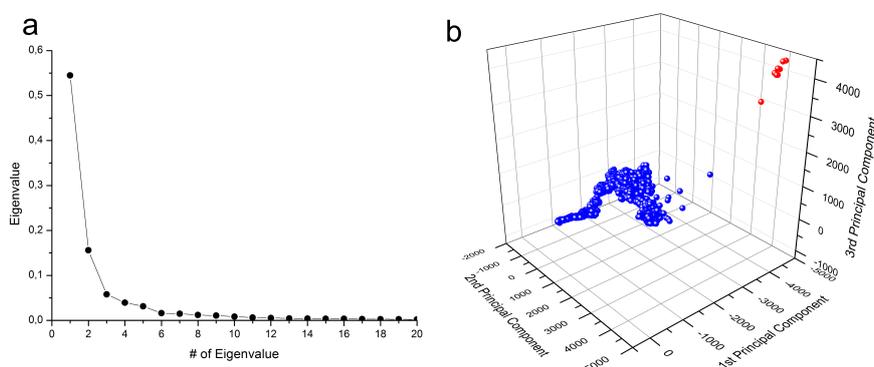ropean project SPEDOC in which the origins of this thesis are inscribed. In order to test the ability of our technique to solve this sensing problematic, a low concentrated solution of HSP-70 (50 $\mu$g/mL) was mixed with an equivalent volume of blood serum. The global concentration of HSP is clearly higher than the goal concentration but the test is more qualitative than quantitative. We have previously shown that the use of our method on raw blood serum is unusual, however in this specific case a hope is allowed because this can be considered as a simple two species mix.

Thus, the solution is injected through a microfluidic channel with a flow of 2 $\mu$L/min. A record of the SERS signal from a single hot spot is acquired during 2048 seconds, i.e. the record is composed by 2048 spectra of 1 seconds.

In this specific case, two approaches are gathered: the detection of a complex protein like HSP-70 and the separation of different species in a mix. This combination totally justifies the use of a multivariate statistical method to clearly identify the HSP-70 fingerprint in this complex medium purposed by blood serum.

The application of PCA on this record results in the eigenvalue decomposition shown on Fig. 5.12 (a). From this decomposition brings out the number of Principal Components necessary to well-describe the complexity of the record. Thus the projec-
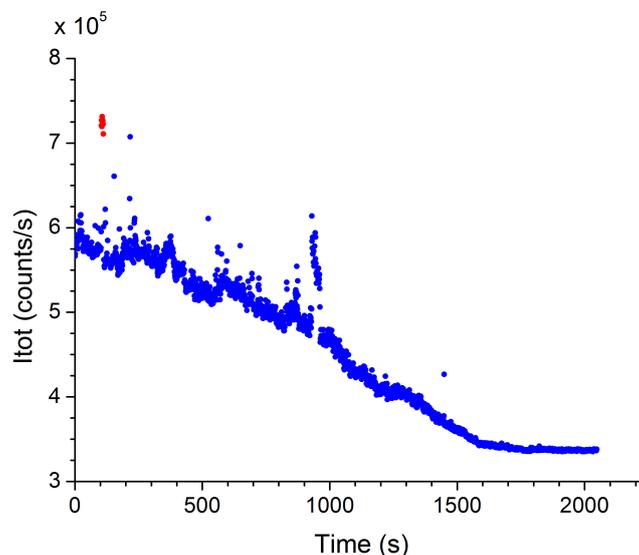
Figure 5.13 – Colored kinetic of the mix serum-HSP-70 studied by SERS. Colors result from Fig. 5.12 (b).

tion in a three dimensional space governed by the three first Principal Component is a good solution. This projection is represented in Fig. 5.12 (b).

The Principal Component decomposition thus obtained shows that the majority of recorded spectra can be described by the two first Principal Components, what is completely different for a low number of them that are clearly localized along the third Principal Component. That is why we chose to look precisely at these spectra firstly by tag them in red and thus localized them in the kinetic signature of the record shown in Fig. 5.13.

From this kinetic, it results that these spectra are localized at one specific event ($\approx 120$ s) during the record. The uniqueness of this event informs about the difficulty to obtain such event and implies an important caution. This caution is confirmed by the global form of the kinetic that can be interpreted as a defocusing of the laser or a problem in the microfluidic system (e.g. leakage). The question is to know if this event should be the signature of HSP-70. Indeed, we previously knew that it would be difficult to isolate HSP in this mix due to the high affinity of Human Serum Albumin, the equivalent of BSA for human the principal protein in blood serum, for GNPs.

In this aim, we compare the spectral fingerprint of this event (Fig. 5.14) with the spectra obtained for HSP-70. Just with a fast comparison purpose with Table 5.4, it is clear that the extracted spectrum is a spectrum of the HSP-70 present in the

104
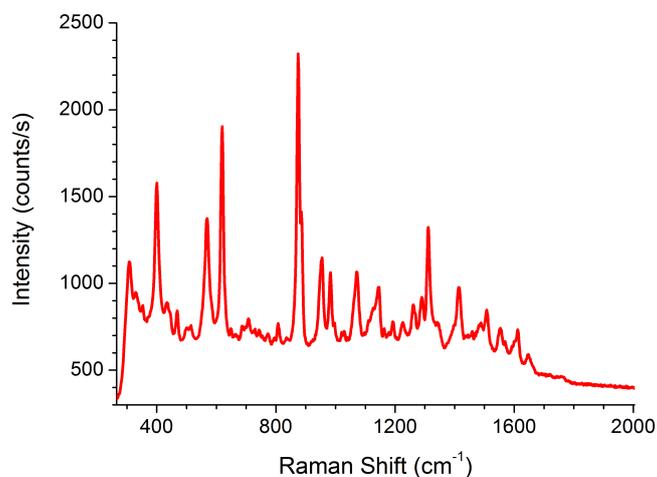
Figure 5.14 – Characteristic spectra identified of HSP-70 in the serum-HSP-70.

solution. Looking for the specific form of HSP-7à that should have been sounded, it seems that it should be the Form 2, which is the most present in HSP-70 assignment for this spectrum (Table 5.4).

Thus, our method enables to identify the specific spectral fingerprint of HSP-70 in a so complex medium as human blood serum based on the Principal Component decomposition. However, this result has to be confirmed by multiple experiments in the same condition. Indeed, a clear effect of defocusing or other appears looking on the kinetic and has to be solved. A repetition of an HSP-70 event at different moment of the studied record should be a good point to confirm the method.

Table 5.4 – Raman assignments of the Red family characteristic spectra based on HSP-70 spectral fingerprints.

| Raman Shift (cm$^{-1}$) | Relative Intensity | HSP-70 identification |
| --- | --- | --- |
| 309 | m | Form 2 |
| 329 | w | Form 1 or 2 |
| 358 | w | Form 1 or 2 |
| 400 | s | Form 3 |
| 436 | w | |
| 468 | w | Form 2 or 3 |
| 515 | w | Form 1, 2 or 3 |
| 573 | s | Form 2 or 3 |
| 620 | s | Form 2 |
| 708 | w | Form 2 or 3 |
| 806 | w | |
| 882 | s | Form 1 |
| 952 | m | Form 1 |
| 983 | m | Form 2 or 3 |
| 1029 | w | Form 2 |
| 1073 | m | |
| 1143 | m | Form 2 |
| 1193 | w | Form 2 |
| 1228 | w | Form 2 |
| 1263 | w | Form 1 |
| 1290 | w | Form 2 |
| 1312 | m | |
| 1343 | w | Form 1 |
| 1413 | m | Form 1 |
| 1475 | w | Form 1 or 2 |
| 1509 | w | Form 1, 2 or 3 |
| 1557 | w | Form 3 |
| 1615 | w | Form 3 |
| 1649 | w | |

## 5.6   Conclusion

After working on single amino acids, the question was to know if our approach should be efficient on protein investigations which are more specific in the living than amino acids. In this aim, we first investigated one of the most abundant proteins: Bovine Serum Albumin (BSA), which as a direct human equivalent, the HSA. BSA is a classic target for SERS protein sensing. Our experiments result in result close to SERS fingerprints found in literature, with a better spectral definition. Further investigations have also proved different rare conformations of this protein just based on an intrinsic biomarker spectral fingerprint, Tryptophane.

Then, the protein conformation identification has been confirmed in the study of the complexity of a multi-domain protein: Heat Shock Protein-70, which is the specific stress cancer biomarker in human blood investigated in the European project FP7 SPE-DOC and the complexity induced by a complex medium such as serum demonstrates the limit of our analytical method.

Finally, the identification of the protein HSP-70 was retrieved in human blood serum at a concentration of 50 $\mu$g/mL.

# Conclusion and Perspectives

At the beginning of this thesis, the question of the definition of SERS as a biosensor was asked. To answer to this, a new experimental approach has been developed for SERS studies based on a specific GNPs nanoraspberry form without any chemical functionalization on the surface during its use. This was associated to a microfluidic chip allowing working in solution and native biological conformations. All of these were mounted on a confocal microscope with a very high Numerical Aperture of 1.20 and two parallel detectors: a spectral detection on a spectrometer and a temporal detection on an Avalanche PhotoDiode. Also, the dynamic approach enables improving the reproducibility of the sensor. However, the large amount of spectra and APD values and the complexity of these results bring us to push the development of the statistical approach developed.

Then, the quantification by SERS was tested studying a large amount of concentrations for a localized hot spot. This study does not allowed to well define an Enhancement Factor for the experimental system. Indeed, an intrinsic definition of an Enhancement Factor in single molecule conditions seems difficult. However, this study enables determining the Limit Of Detection of the setup for a 512 seconds long acquisition. The temporal recording has been also of high interest to define a new approach to argue the single molecule regime based on the Fourier Transform of the temporal signal to work in the frequency domain and also highlight the different types of signal observed, and more precisely, the signal coming from the adsorption of the targeted molecule on the surface. Two types of LODs have been defined. First in frequency space, the LOD is so around 1 fM. Then, in time domain, we show that LOD depends of the sounding spot. In the case presented, it was around $10^{-10}$ M.

The question of the main gain from statistics has been developed in the context of a small non resonant molecule study, the case of cysteine. The spectral sorting

allowed by the Principal Components decomposition of each spectrum implies sorting in different spectral families. The study of cysteine also argued the single molecule regime by its temporal fingerprint study. Spectrally, the application of SERS on a such molecule found out about the gold-molecule interaction and enables sorting two similar molecules in a same solution.

In the last part, the interests were to understand and test how SERS can be efficient for biological macromolecules studies and selective for complex protein and/or complex media detection. Thus, different complexity levels induce different performance level of SERS. First, concerning a standard protein such as BSA, our approach implies a characteristic spectrum determination for the protein which finally is not simply an amino acid fingerprints composition. For this protein, the chosen method enabled one very rare conformation of BSA but highlighted by the statistical analysis, the acquisition speed and the single molecule resolution. Secondly, a complex protein, HSP-70, has been probed and spectrally characterized with our system, the different conformations given by the protein has been sorted and each spectral specific fingerprints determined. Thirdly, the application of the method to a complex medium as serum highlighted the limits of the chosen method in terms of statistical analysis but also concerning the real interest of SERS in such case. Finally, the mix of HSP-70 with serum has been tested. The idea was to try to detect the presence of a complex but important molecule in a complex medium to determine the ability of SERS in terms of specific identification without any previous functionalization. Hence, we have been able to detect the HSP-70 based on our statistical approach and comparing with the characteristic fingerprints of HSP-70 previously determined with the same system.

To summarize, in this thesis, the definition of SERS as a biosensor has been tested and a new approach developed for. Also, in terms of quantification, it has been shown that SERS can be an efficient tool. Concerning the selectivity, the spectral quality was improved. A low limit of detection associated to the statistical and dynamic approach allows a very good sensitivity (under the nanomolar). This approach also enables a high reproducibility in time of the sensor. Thus, as low as SERS does not well answer to the sensor capabilities in a classical approach, in our case the coupling between a non-functionalized GNPs substrate coupled with a microfluidic chip, all mounted on a confocal microscope for temporal dynamic studies statistically analyzed has contributed to define SERS as an efficient biosensor.

## Perspectives

SERS is so a very selective sensor that can be a good tool for molecular or protein recognition. However, we have shown that our approach is limited in terms of complexity of the system. Thus, a maximum of three different "species" has been separated during a same acquisition in the case of the HSP-70 study. The development of a new

statistical approach for the spectral sorting seems necessary. One of the ideas is to use a derivative application of the Phasor still used in fluorescence studies. The interest of this method is that it converts the spectra in two parameters in the Fourier space. There is also no limitations in terms of complexity as we should have with PCA and the number of necessaries eigenvalues.

All results of this thesis are preliminaries concerning the dynamical approach of SERS and should be easily used to develop an experimental and analytical systematic method for the investigation of specific molecules as well in biomedical than in energetic, environmental or other domains. However, the quantification aspects have to be developed and better understand to result in a SERS not only as a good sensor but also as a concrete and efficient diagnostic tool.

# Bibliography

[1] A. Nicolai, P. Senet, P. Delarue, and D. R. Ripoll, "Human Inducible Hsp70 : Structures , Dynamics , and Interdomain Communication from All-Atom Molecular Dynamics Simulations," *Journal of Chemical Theory and Computation*, vol. 6, no. 8, pp. 2501–2519, 2010.

[2] D. C. Carter and J. X. Ho, "Structure of Serum Albumin," in *Lipoproteins, Apolipoproteins, and Lipases* (F. M. R. C.B. Anfinsen John T. Edsall and D. S. Eisenberg, eds.), vol. 45 of *Advances in Protein Chemistry*, pp. 153–203, Academic Press, 1994.

[3] Y. Pang and R. Gordon, "Optical trapping of a single protein.," *Nano letters*, vol. 12, pp. 402–6, Jan. 2012.

[4] A.-L. Rérole, G. Jego, and C. Garrido, "Hsp70: Anti-apoptotic and Tumorigenic Protein," in *Molecular Chaperones SE - 16* (S. K. Calderwood and T. L. Prince, eds.), vol. 787 of *Methods in Molecular Biology*, pp. 205–230, Humana Press, 2011.

[5] K. L. Norrod, L. M. Sudnik, D. Rousell, and K. L. Rowlen, "Quantitative Comparison of Five SERS Substrates: Sensitivity and Limit of Detection," *Appl. Spectrosc.*, vol. 51, no. 7, pp. 994–1001, 1997.

[6] G. Das, F. Gentile, M. L. Coluccio, A. M. Perri, A. Nicastri, F. Mecarini, G. Cojoc, P. Candeloro, C. Liberale, F. De Angelis, and E. Di Fabrizio, "Principal component analysis based methodology to distinguish protein SERS spectra," *JOURNAL OF MOLECULAR STRUCTURE*, vol. 993, no. 1-3, SI, pp. 500–505, 2011.

[7] C. David, N. Guillot, H. Shen, T. Toury, and M. L. de la Chapelle, "SERS detection of biomolecules using lithographed nanoparticles towards a reproducible SERS biosensor," *NANOTECHNOLOGY*, vol. 21, Nov. 2010.

[8] M. Iosin, F. Toderas, P. L. Baldeck, and S. Astilean, "Study of protein–gold nanoparticle conjugates by fluorescence and surface-enhanced Raman scattering," *Journal of Molecular Structure*, vol. 924–926, no. 0, pp. 196–200, 2009.

[9] M. Kahraman, I. Sur, and M. Culha, "Label-Free Detection of Proteins from Self-Assembled Protein-Silver Nanoparticle Structures using Surface-Enhanced Raman Scattering," *ANALYTICAL CHEMISTRY*, vol. 82, pp. 7596–7602, Sept. 2010.

[10] C.-K. Yao, J.-D. Liao, C.-W. Chang, and J.-R. Lin, "Spatially reinforced nano-cavity array as the SERS-active substrate for detecting hepatitis virus core antigen at low concentrations," *Sensors and Actuators B: Chemical*, vol. 174, pp. 478–484, Nov. 2012.

[11] M. Zhang, A. Zhao, H. Sun, H. Guo, D. Wang, D. Li, Z. Gan, and W. Tao, "Rapid, large-scale, sonochemical synthesis of 3D nanotextured silver microflowers as highly efficient SERS substrates," *Journal of Materials Chemistry*, vol. 21, no. 46, p. 18817, 2011.

[12] J. Zhou, K. Ren, Y. Zhao, W. Dai, and H. Wu, "Convenient formation of nanoparticle aggregates on microfluidic chips for highly sensitive SERS detection of biomolecules," *Analytical and Bioanalytical Chemistry*, vol. 402, no. 4, pp. 1601–1609, 2012.

[13] C. Eliasson, A. Lorén, J. Engelbrektsson, M. Josefson, J. Abrahamsson, and K. Abrahamsson, "Surface-enhanced Raman scattering imaging of single living lymphocytes with multivariate evaluation.," *Spectrochimica acta. Part A, Molecular and biomolecular spectroscopy*, vol. 61, pp. 755–60, Feb. 2005.

[14] J. Kneipp, H. Kneipp, A. Rajadurai, R. W. Redmond, and K. Kneipp, "Optical probing and imaging of live cells using SERS labels," *Journal of Raman Spectroscopy*, vol. 40, pp. 1–5, Jan. 2009.

[15] S. Kawata, Y. Inouye, and P. Verma, "Plasmonics for near-field nano-imaging and superlensing," *Nat Photon*, vol. 3, pp. 388–394, July 2009.

[16] C. J. Murphy, A. M. Gole, S. E. Hunyadi, J. W. Stone, P. N. Sisco, A. Alkilany, B. E. Kinard, and P. Hankins, "Chemical sensing and imaging with metallic nanorods.," *Chemical communications (Cambridge, England)*, pp. 544–57, Feb. 2008.

[17] L. Qin, S. Zou, C. Xue, A. Atkinson, G. C. Schatz, and C. a. Mirkin, "Designing, fabricating, and imaging Raman hot spots.," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 103, pp. 13300–3, Sept. 2006.

[18] C. De Bleye, P.-Y. Sacré, E. Dumont, L. Netchacovitch, P.-F. Chavez, G. Piel, P. Lebrun, P. Hubert, and E. Ziemons, "Development of a quantitative approach using surface-enhanced Raman chemical imaging: first step for the determination of an impurity in a pharmaceutical model.," *Journal of pharmaceutical and biomedical analysis*, vol. 90, pp. 111–8, Mar. 2014.

[19] I. Romero, J. Aizpurua, G. W. Bryant, and F. J. García De Abajo, "Plasmons in nearly touching metallic nanoparticles: singular response in the limit of touching dimers.," *Optics express*, vol. 14, pp. 9988–99, Oct. 2006.

[20] D. L. Jeanmaire and R. P. Van Duyne, "Surface Raman Spectroelectrochemistry," *Journal of Electroanalytical Chemistry*, vol. 84, pp. 1–20, 1977.

[21] P. Hildebrandt and M. Stockburger, "Surface-enhanced resonance Raman spectroscopy of Rhodamine 6G adsorbed on colloidal silver," *The Journal of Physical Chemistry*, vol. 88, no. 24, pp. 5935–5944, 1984.

[22] E. C. Le Ru, E. Blackie, M. Meyer, and P. G. Etchegoin, "Surface Enhanced Raman Scattering Enhancement Factors : A Comprehensive Study," *The Journal of Physical Chemistry C*, vol. 111, no. 24, pp. 13794–13803, 2007.

[23] E. H. Ismail, M. M. H. Khalil, F. A. A. Seif, F. El-magdoub, A.-a. N. Bent, A. Rahman, and U.-f. S.-c. D.-h. Al, "Biosynthesis of Gold Nanoparticles Using Extract of Grape ( Vitis Vinifera ) Leaves and Seeds," *Progress in Nanotechnology and Nanomaterials*, vol. 3, pp. 1–12, 2014.

[24] T. Wang, X. Hu, and S. Dong, "Surfactantless synthesis of multiple shapes of gold nanostructures and their shape-dependent SERS spectroscopy.," *The Journal of Physical Chemistry B.*, vol. 110, pp. 16930–6, Aug. 2006.

[25] J. Margueritat, A. Bouhelier, L. Markey, G. Colas des Francs, A. Dereux, S. Lau-Truong, J. Grand, G. Lévi, N. Félidj, J. Aubard, and E. Finot, "Discerning the Origins of the Amplitude Fluctuations in Dynamic Raman Nanospectroscopy," *The Journal of Physical Chemistry C*, vol. 116, pp. 26919–26923, Dec. 2012.

[26] B. L. Darby and E. C. Le Ru, "Competition between Molecular Adsorption and Diffusion: Dramatic Consequences for SERS in Colloidal Solutions.," *Journal of the American Chemical Society*, vol. 136, pp. 10965–73, Aug. 2014.

[27] L. Chen and J. Choo, "Recent advances in surface-enhanced Raman scattering detection technology for microfluidic chips.," *Electrophoresis*, vol. 29, pp. 1815–28, May 2008.

[28] W. Zhu, M. G. Banaee, D. Wang, Y. Chu, and K. B. Crozier, "Lithographically fabricated optical antennas with gaps well below 10 nm.," *Small (Weinheim an der Bergstrasse, Germany)*, vol. 7, pp. 1761–6, July 2011.

[29] W.-C. Lin, S.-H. Huang, C.-L. Chen, C.-C. Chen, D. P. Tsai, and H.-P. Chiang, "Controlling SERS intensity by tuning the size and height of a silver nanoparticle array," *Applied Physics A*, vol. 101, pp. 185–189, June 2010.

[30] E. M. Hicks, O. Lyandres, W. P. Hall, S. Zou, M. R. Glucksberg, and R. P. Van Duyne, "Plasmonic Properties of Anchored Nanoparticles Fabricated by Reactive Ion Etching and Nanosphere Lithography," *The Journal of Physical Chemistry C*, vol. 111, pp. 4116–4124, Mar. 2007.

[31] M. Kahl, E. Voges, S. Kostrewa, C. Viets, and W. Hill, "Periodically structured metallic substrates for SERS," *Sensors and Actuators B: Chemical*, vol. 51, pp. 285–291, Aug. 1998.

[32] W. Ahn, Y. Qiu, and B. M. Reinhard, "Generation of scalable quasi-3D metallo-dielectric SERS substrates through orthogonal reactive ion etching," *Journal of Materials Chemistry C*, vol. 1, no. 18, p. 3110, 2013.

[33] R. Tan, a. Agarwal, N. Balasubramanian, D. Kwong, Y. Jiang, E. Widjaja, and M. Garland, "3D arrays of SERS substrate for ultrasensitive molecular detection," *Sensors and Actuators A: Physical*, vol. 139, pp. 36–41, Sept. 2007.

[34] M. Mabuchi, T. Takenaka, Y. Fujiyoshi, and N. Uyeda, "Surface enhanced Raman scattering of citrate ions adsorbed on gold sol particles," *Surface Science*, vol. 119, pp. 150–158, 1982.

[35] C. D. Keating, K. M. Kovaleski, and M. J. Natan, "Protein:Colloid Conjugates for Surface Enhanced Raman Scattering: Stability and Control of Protein Orientation," *The Journal of Physical Chemistry B*, vol. 102, pp. 9404–9413, Nov. 1998.

[36] C. E. Talley, L. Jusinski, C. W. Hollars, S. M. Lane, and T. Huser, "Intracellular pH sensors based on surface-enhanced raman scattering.," *Analytical chemistry*, vol. 76, pp. 7064–8, Dec. 2004.

[37] N. Félidj, J. Aubard, G. Lévi, J. Krenn, M. Salerno, G. Schider, B. Lamprecht, a. Leitner, and F. Aussenegg, "Controlling the optical response of regular arrays of gold particles for surface-enhanced Raman scattering," *Physical Review B*, vol. 65, p. 075419, Feb. 2002.

[38] K. Kneipp, H. Kneipp, I. Itzkan, R. R. Dasari, and M. S. Feld, "Surface-enhanced Raman scattering and biophysics," *Journal of Physics: Condensed Matter*, vol. 14, pp. R597–R624, May 2002.

[39] N. Guarrotxena, B. Liu, L. Fabris, and G. C. Bazan, "Antitags: nanostructured tools for developing SERS-based ELISA analogs.," *Advanced materials (Deerfield Beach, Fla.)*, vol. 22, pp. 4954–8, Nov. 2010.

[40] Z. Wang and L. J. Rothberg, "Origins of blinking in single-molecule Raman spectroscopy.," *The Journal of Physical Chemistry B*, vol. 109, pp. 3387–91, Mar. 2005.

[41] K. Kneipp, Y. Wang, H. Kneipp, L. T. Perelman, I. Itzkan, R. Dasari, and M. S. Feld, "Single molecule detection using surface-enhanced Raman scattering (SERS)," *PHYSICAL REVIEW LETTERS*, vol. 78, pp. 1667–1670, Mar. 1997.

[42] S. M. Nie and S. R. Emery, "Probing single molecules and single nanoparticles by surface-enhanced Raman scattering," *SCIENCE*, vol. 275, no. 5303, pp. 1102–1106, 1997.

[43] E. C. Le Ru and P. G. Etchegoin, "Quantifying SERS enhancements," *MRS Bulletin*, vol. 38, pp. 631–640, Aug. 2013.

[44] E. C. Le Ru, J. Grand, I. Sow, W. R. C. Somerville, P. G. Etchegoin, M. Treguer-Delapierre, G. Charron, N. Félidj, G. Lévi, and J. Aubard, "A scheme for detecting every single target molecule with surface-enhanced Raman spectroscopy.," *Nano letters*, vol. 11, pp. 5013–9, Nov. 2011.

[45] S. Lee, S. Joo, S. Park, S. Kim, H. C. Kim, and T. D. Chung, "SERS decoding of micro gold shells moving in microfluidic systems.," *Electrophoresis*, vol. 31, pp. 1623–9, May 2010.

[46] C. Delhaye, J.-L. Bruneel, D. Talaga, M. Guirardel, S. Lecomte, and L. Servant, "Tailoring Surface-Enhanced Raman Scattering Effect Using Microfluidics," *JOURNAL OF PHYSICAL CHEMISTRY C*, vol. 116, pp. 5327–5332, Mar. 2012.

[47] C. Lim, J. Hong, B. G. Chung, A. J. DeMello, and J. Choo, "Optofluidic platforms based on surface-enhanced Raman scattering.," *The Analyst*, vol. 135, pp. 837–44, May 2010.

[48] E. Kamholz and P. Yager, "Theoretical analysis of molecular diffusion in pressure-driven laminar flow in microfluidic channels.," *Biophysical journal*, vol. 80, pp. 155–60, Jan. 2001.

[49] P. G. Etchegoin, M. Meyer, E. Blackie, and E. C. Le Ru, "Statistics of single-molecule surface enhanced Raman scattering signals: fluctuation analysis with multiple analyte techniques.," *Analytical chemistry*, vol. 79, pp. 8411–5, Nov. 2007.

[50] A. Otto, "Raman scattering from adsorbates on silver," *Surface Science*, vol. 92, no. June 1978, pp. 145–152, 1980.

[51] J. Billmann, G. Kovacs, and A. Otto, "Enhanced Raman effect from cyanide adsorbed on a silver electrode," *Surface Science*, vol. 92, pp. 153–173, 1980.

117

[52] D.-Y. Wu, X.-M. Liu, S. Duan, X. Xu, B. Ren, S.-H. Lin, and Z.-Q. Tian, "Chemical Enhancement Effects in SERS Spectra: A Quantum Chemical Study of Pyridine Interacting with Copper, Silver, Gold and Platinum Metals," *The Journal of Physical Chemistry C*, vol. 112, pp. 4195–4204, Mar. 2008.

[53] J. Xie, Q. Zhang, J. Y. Lee, and D. I. C. Wang, "The Synthesis of SERS-Active Gold Nanoflower Tags for In Vivo Applications," *ACS nano*, vol. 2, no. 12, pp. 2473–2480, 2008.

[54] H. Yockell-Lelièvre, J. Desbiens, and A. M. Ritcey, "Two-dimensional self-organization of polystyrene-capped gold nanoparticles.," *Langmuir : the ACS journal of surfaces and colloids*, vol. 23, pp. 2843–50, Feb. 2007.

[55] H. J. Kimble, M. Dagenais, and L. Mandel, "Photon antibunching in resonance fluorescence," *Phys. Rev. Lett.*, vol. 39, pp. 691–695, Sep 1977.

[56] L. Mandel, "Sub-Poissonian photon statistics in resonance fluorescence," *Opt. Lett.*, vol. 4, no. 7, pp. 205–207, 1979.

[57] F. Treussart, R. Alléaume, V. Le Floc'h, L. Xiao, J.-M. Courty, and J.-F. Roch, "Direct Measurement of the Photon Statistics of a Triggered Single Photon Source," *Physical Review Letters*, vol. 89, p. 093601, Aug. 2002.

[58] G. Margolin and E. Barkai, "Single-molecule chemical reactions: Reexamination of the Kramers approach," *Physical Review E*, vol. 72, p. 025101, Aug. 2005.

[59] M. R. Gonçalves, F. Enderle, and O. Marti, "Surface-Enhanced Raman Spectroscopy of Dye and Thiol Molecules Adsorbed on Triangular Silver Nanostructures: A Study of Near-Field Enhancement, Localization of Hot-Spots, and Passivation of Adsorbed Carbonaceous Species," *Journal of Nanotechnology*, vol. 2012, pp. 1–15, 2012.

[60] F. Paquet-Mercier, N. B. Aznaveh, M. Safdar, and J. Greener, "A microfluidic bioreactor with in situ SERS imaging for the study of controlled flow patterns of biofilm precursor materials.," *Sensors*, vol. 13, pp. 14714–27, Jan. 2013.

[61] O. Péron, E. Rinnert, T. Toury, M. Lamy de la Chapelle, and C. Compère, "Quantitative SERS sensors for environmental analysis of naphthalene.," *The Analyst*, vol. 136, pp. 1018–22, Mar. 2011.

[62] M. H. Harpster, H. Zhang, A. K. Sankara-Warrier, B. H. Ray, T. R. Ward, J. P. Kollmar, K. T. Carron, J. O. Mecham, R. C. Corcoran, W. C. Wilson, and P. a. Johnson, "SERS detection of indirect viral DNA capture using colloidal gold and methylene blue as a Raman label.," *Biosensors & bioelectronics*, vol. 25, pp. 674–81, Dec. 2009.

[63] C. D'Andrea, A. Toma, C. Huck, F. Neubrech, E. Messina, E. D. Fabrizio, M. Lamy, D. L. Chapelle, P. G. Gucciardi, B. Fazio, and O. M. Marago, "Optical Nanoantennas for Multiband Surface-enhanced Infrared and Raman Spectroscopy," *ACS nano*, vol. 7, no. 4, pp. 3522–3531, 2013.

[64] M. Nidya, M. Umadevi, and B. J. M. Rajkumar, "Optical and morphological studies of L-histidine functionalised silver nanoparticles synthesised by two different methods," *Journal of Experimental Nanoscience*, vol. 0, no. 0, pp. 1–14, 2013.

[65] C. Andreou, M. R. Hoonejani, M. R. Barmi, M. Moskovits, and C. D. Meinhart, "Rapid Detection of Drugs of Abuse in Saliva Using Surface Enhanced Raman Spectroscopy and Microfluidics," *ACS Nano*, no. 8, pp. 7157–7164, 2013.

[66] A. Matschulat, D. Drescher, and J. Kneipp, "Surface-enhanced Raman scattering hybrid nanoprobe multiplexing and imaging in biological systems.," *ACS nano*, vol. 4, pp. 3259–69, June 2010.

[67] A. Barth and C. Zscherp, "What vibrations tell about proteins," *Quarterly Reviews of Biophysics*, vol. 35, no. 04, pp. 369–430, 2002.

[68] a.E. Aliaga, I. Osorio-Roman, C. Garrido, P. Leyton, J. Cárcamo, E. Clavijo, J. Gómez-Jeria, G. D. F., and M. Campos-Vallette, "Surface enhanced Raman scattering study of l-lysine," *Vibrational Spectroscopy*, vol. 50, pp. 131–135, May 2009.

[69] A. M. Schwartzberg, C. D. Grant, A. Wolcott, C. E. Talley, T. R. Huser, R. Bogomolni, and J. Z. Zhang, "Unique Gold Nanoparticle Aggregates as a Highly Active Surface-Enhanced Raman Scattering Substrate," *The Journal of Physical Chemistry B*, vol. 108, pp. 19191–19197, Dec. 2004.

[70] S. W. Bishnoi, C. J. Rozell, C. S. Levin, M. K. Gheith, B. R. Johnson, D. H. Johnson, and N. J. Halas, "All-optical nanoscale pH meter.," *Nano letters*, vol. 6, pp. 1687–92, Aug. 2006.

[71] a. Michota and J. Bukowska, "Surface-enhanced Raman scattering (SERS) of 4-mercaptobenzoic acid on silver and gold substrates," *Journal of Raman Spectroscopy*, vol. 34, pp. 21–25, Jan. 2003.

[72] G. Diaz Fleming, J. J. Finnerty, M. Campos-Vallette, F. Célis, A. E. Aliaga, C. Fredes, and R. Koch, "Experimental and theoretical Raman and surface-enhanced Raman scattering study of cysteine," *Journal of Raman Spectroscopy*, vol. 40, pp. 632–638, June 2009.

[73] C. Jing and Y. Fang, "Experimental (SERS) and theoretical (DFT) studies on the adsorption behaviors of l-cysteine on gold/silver nanoparticles," *Chemical Physics*, vol. 332, pp. 27–32, Jan. 2007.

[74] E. Podstawka, Y. Ozaki, and L. M. Proniewicz, "Part III: Surface-Enhanced Raman Scattering of Amino Acids and Their Homodipeptide Monolayers Deposited onto Colloidal Gold Surface," *Appl. Spectrosc.*, vol. 59, no. 12, pp. 1516–1526, 2005.

[75] E. C. Le Ru, M. Meyer, and P. G. Etchegoin, "Proof of single-molecule sensitivity in surface enhanced Raman scattering (SERS) by means of a two-analyte technique.," *The journal of physical chemistry. B*, vol. 110, pp. 1944–8, Feb. 2006.

[76] I. Pavel, E. McCarney, A. Elkhaled, A. Morrill, K. Plaxco, and M. Moskovits, "Label-Free SERS Detection of Small Proteins Modified to Act as Bifunctional Linkers.," *The journal of physical chemistry C Nanomaterials and interfaces*, vol. 112, pp. 4880–4883, 2008.

[77] S. Stewart and P. Fredericks, "Surface-enhanced Raman spectroscopy of amino acids adsorbed on an electrochemically prepared silver surface," *Spectrochimica Acta Part A: Molecular and Biomolecular Spectroscopy*, vol. 55, pp. 1641–1660, July 1999.

[78] a. Culka, J. Jehlička, and H. G. M. Edwards, "Acquisition of Raman spectra of amino acids using portable instruments: outdoor measurements and comparison.," *Spectrochimica acta. Part A, Molecular and biomolecular spectroscopy*, vol. 77, pp. 978–83, Dec. 2010.

[79] H. Lee, M. S. Kim, and S. W. Suh, "Raman spectroscopy of sulphur-containing amino acids and their derivatives adsorbed on silver," *Journal of Raman Spectroscopy*, vol. 22, no. 2, pp. 91–96, 1991.

[80] M. Graff and J. Bukowska, "Adsorption of enantiomeric and racemic cysteine on a silver electrode—SERS sensitivity to chirality of adsorbed molecules.," *The journal of physical chemistry. B*, vol. 109, pp. 9567–74, May 2005.

[81] "See on website: http://pubs.acs.org/doi/suppl/10.1021/jp504395c/suppl_file/jp504395c_si_002.avi."

[82] N. R. Yaffe and E. W. Blanch, "Effects and anomalies that can occur in SERS spectra of biological molecules when using a wide range of aggregating agents for hydroxylamine-reduced and citrate-reduced silver colloids," *Vibrational Spectroscopy*, vol. 48, pp. 196–201, Nov. 2008.

[83] M. Moskovits, "Surface-enhanced spectroscopy," *Review of Modern Physics*, vol. 57, no. 3, pp. 785–828, 1985.

[84] M. Iosin, V. Canpean, and S. Astilean, "Spectroscopic studies on pH- and thermally induced conformational changes of Bovine Serum Albumin adsorbed onto

gold nanoparticles," *Journal of Photochemistry and Photobiology A: Chemistry*, vol. 217, pp. 395–401, Jan. 2011.

[85] M. Wang, M. Benford, N. Jing, G. Coté, and J. Kameoka, "Optofluidic device for ultra-sensitive detection of proteins using surface-enhanced Raman spectroscopy," *Microfluidics and Nanofluidics*, vol. 6, pp. 411–417, Jan. 2009.

[86] X. X. Han, H. Y. Jia, Y. F. Wang, Z. C. Lu, C. X. Wang, W. Q. Xu, B. Zhao, and Y. Ozaki, "Analytical Technique for Label-Free Multi-Protein Detection Based on Western Blot and Surface-Enhanced Raman Scattering," *Analytical Chemistry*, vol. 80, no. 8, pp. 2799–2804, 2008.

[87] C. Sun, J. Yang, X. Wu, X. Huang, F. Wang, and S. Liu, "Unfolding and refolding of bovine serum albumin induced by cetylpyridinium bromide.," *Biophysical journal*, vol. 88, pp. 3518–24, May 2005.