Master's Degree in Proteomics and Bioinformatics

Imaging Shotgun IPG-IEF

Jovan Simicevic
Biomedical Proteomics Research Group
Department of Structural Biology and Bioinformatics
University of Geneva

Project Director:	Prof. Denis Hochstrasser
Group Leader:	Dr. Pierre Lescuyer
Supervisors:	Alireza Vaezzadeh
	Prof. Jacques Deshusses

Abstract

In the past few years, the combination of high-throughput identification of proteins via whole proteome digestion with multidimensional liquid chromatography (LC) tandem mass spectrometry (MS/MS), referred to as Shotgun proteomics, is used in a routine manner in the proteomics field, and referred as to. Recently, Cargile *et al.* [1] presented an alternative to the classical liquid chromatography as first dimension separation technique, by using immobilized pH gradient isoelectric focusing (IPG-IEF). The separation in the second dimension is performed by Reverse phase liquid chromatography. Among the advantages of this technique, one can mention: high loading capacity, high resolving power, broad dynamic range and high reproducibility.

However, the LC-MS/MS step of the pipeline is time-consuming. In order to rapidly obtain a preview of the sample, a shortcut approach was developed based on the transfer of peptides from the IPG strip a capture membrane. The membrane is then scanned in the MS instrument to create virtual MS images. A set of fluorescent markers was developed to be used as reference points for the differential comparison. The markers permit a better alignment between the images during superimposition. Moreover, the markers can be used to normalize the pI gradient irregularities and differences in focusing.In this study, we present the advantages and the weak points of the developed pipeline by applying it to the analysis of *S.aurues* S_{30} helicase knock-out gene mutant H(-), and its helicase gene complemented mutant H(+) to test this newly developed technique.

Acknowledgments

I would like to express my most profound gratitude to Alireza Vaezzadeh for guiding me throughout all the different aspects of my internship. He showed the highest interest in my scientific formation, and demonstrated an innate ability in the discipline of teaching. I certainly learned a lot from him not only from a scientific point of view, but also from a human perspective. I will certainly carry the memory of these six months spent working together for a long time to come.

My most profound gratitude also goes to Professor D. Hochstrasser for his supervision in my work, and for his generosity in accepting me as a member of his laboratory. I would like to thank Professor J. Deshusses and Dr. Pierre Lescuyer for their precious advices and supervision. I am also grateful to Dr. Patrice François and his laboratory for the preparation of the biological samples and the interest expressed in our work.

A special thanks goes to all the members of the BPRG group and the members of the Proteomic Platform for their help and support. Also, I am grateful to all the members of the Swiss institute of bioinformatics for the developing of the software. Finally, I would like to thank Dr. J.-C. Sanchez and Dr. P. Palagi for their counseling.

Table of contents

List of abbreviations

AcN	Acetonitrile
BA	Ammonium bicarbonate
BSA	Bovine serum albumin
CA	Carrier ampholytes
CHCA	α-cyano-4-hydroxycinnamic acid
Da	Dalton
DTE	1, 4-dithioerythritol
ESI	Electrospray ionization
eV	Electron volt
FA	Formic acid
HPLC	High-performance liquid chromatography
Hz	Hertz
IEF	Isoelectric focusing
IPG	Immobilised pH gradient
kV	kilovolt
kVh	kilovolt x hour
LC	Liquid chromatography
m/z	Mass-to-charge ratio
MALDI	Matrix-assisted laser desorption/ionisation
MS	Mass spectrometry
MS/MS	Tandem mass spectrometry
MSI	Mass spectrometric imaging
PCR	Polymerase chain reaction
pI	Isoelectric point
PMF	Peptide mass fingerprinting
PTM	Post-translational modifications
PVDF	PolyVinylidene DiFluoride

1. Introduction

The history of life sciences traces the study of the living world from ancient to modern times. Although the concept of life science as we intend it today arose in the 19th century, biological sciences emerged from traditions of medicine and natural history reaching back to Galen and Aristotle in ancient Greece. During the Renaissance and early modern period, biological thought was revolutionized by an increasing interest in empiricism and the discovery of many novel organisms. Over the 18th and 19th centuries, biological sciences such as botany and zoology became increasingly professional scientific disciplines. But it wasn't until the early 20th century, with the rapid development of genetics, that we were able to make an incredible amount of discoveries and technical innovations, especially after Watson and Crick proposed the structure of DNA. Following the establishment of the Central Dogma and the cracking of the genetic code, an increasing interest was devoted to the fields of cellular and molecular biology. By the late 20th century, new fields such as genomics and proteomics came into place. It was the beginning of a new era in scientific research.

1.1 Proteomics

Whereas genomics aims to understand the structure of the genome, examining the molecular mechanisms and the interplay of genetic and environmental factors, proteomics is directed toward the study of the "proteome", meaning the complete set of proteins produced by a species. The word Proteomics literally means the proteomic complement to a genome. The main characteristic of this newly developed field is the use of large-scale protein separation and identification technologies. The term proteomics was coined in 1994 by Marc Wilkins [1] who defined it as "the study of proteins, how they're modified, when and where they're expressed, how they're involved in metabolic pathways and how they interact with one another".

Proteomics is considered an essential step in the study of biological systems, just like genomics. Its study presents quite a few challenges though: first, the level of transcription of a gene gives only a bare estimate of its level of expression into protein. Second, many proteins experience post-translational modifications. A large number of proteins are actually not active unless modified. Third, many transcripts give rise to more than one protein due to those post-translational modifications or alternative splicing. Finally, many proteins function only if in presence of other specific proteins or molecules. While an organism's genome is rather constant, its proteome can differ drastically from cell to cell and at different stages of the cell's life cycle, and constantly changes in response to its environment. Therefore, proteomics is a good complement to genomics providing a better understanding of biological processes at a given moment in time.

Even though proteomics complements genomic based approaches, it presents a few technical challenges. Proteins are expressed in a wide range of detection and one of the major difficulties is represented by the fact that there is no protein equivalent of PCR [2]

(polymerase chain reaction) for the amplification of low-abundance proteins. Furthermore, proteins are folded in very specific structures, making generic methods hard to apply. The analysis of post-translational modification represents another challenge.

1.1.1 Protein separation

A crucial step in proteomics is the resolution of the separation. Protein separation is procedure that aims to isolate proteins of interest from contaminants of a biological mixture, to separate desired proteins from all other proteins or to reduce the complexity and enrich low abundant species. Separation techniques generally exploit the physico-chemical properties and binding affinities of proteins. Because the concentration of proteins in biological samples may vary by different orders of magnitude (7 for cells more than 12 for body fluids) [3], it is extremely important to design tailor-made separation techniques. It becomes a crucial aspect when dealing with low abundant proteins, due to the fact that no protein amplification technique exists. The efficiency of fractionation and separation techniques determines the quality of the analysis as a whole.

Very often protein separation protocols include one or more chromatographic steps. The basic concept of liquid chromatography (LC) is to flow a solution containing proteins or peptides through a column packed with various materials. Different proteins interact differently with the column material, and can thus be separated by the time required to pass the column, or by the conditions required to elute the protein from the column.

Size exclusion chromatography [4] employs the size of proteins as separating criterion. The principle is that smaller molecules have to traverse a larger volume in a porous matrix before being collected. Other separation methods are based on charge or hydrophobicity. Ion Exchange Chromatography [5] separates compounds according to the degree of their ionic charge. Affinity chromatography [6] separates compounds upon molecular conformation. It utilizes ligands that capture the target with high specificity. The separation technique that is most widely used in proteomics is high pressure liquid chromatography (HPLC). This method uses high pressure to flow a solute the column in a rapid manner, limiting diffusion and therefore improving resolution. It is often used in reverse phase mode (RPLC) [7] where the stationary phase is non-polar and the mobile phase is polar. RPLC exploits the differences in hydrophobicity of proteins and peptides.
The two techniques most widely used in proteomics for protein and peptide separation exploit two specific physico-chemical properties of proteins: molecular weight and isoelectric point.

Isoelectric focusing (IEF) is a technique for separating protein and peptides based on their electric charge differences [8].
There are different types of electrophoresis commonly used in proteomics. Some of them are run in a supporting media, such as papers, films or gels, whereas others are run in a liquid solution. The most common method that used a supporting media is gel electrophoresis. Gel electrophoresis separates compounds using an electric tension applied to a gel matrix. In most cases the gel is a cross-linked polymer whose composition and porosity is chosen based on the specific weight and size of the target to

be analyzed. For proteins and peptides the gel is usually composed of different concentrations of acrylamide and a cross-linker, producing a network of polyacrylamide. Paper and thin-layer electrophoresis have been abandoned in profit of gel electrophoresis, because of improved separation and the high loading capacity of agarose and polyacrylamide gels.

Free Flow Electrophoresis (FFE) [9] is a highly versatile technology for the separation of proteins and peptides. It has good sample recovery, high sample loading capacity, and high resolution power. However, the buffer constituents may interfere with MS measurements.

Capillary electrophoresis (CE) [10] requires of needing very low sample amounts, usually not more than 2-4 nl. Moreover, it is prone to automation and can be easily coupled with other analytical instruments, such as HPLC. On the other hand, it is a costly technique.

SDS-PAGE (sodium dodecyl sulfate polyacrylamide gel electrophoresis) is a technique used to separate proteins according to their electrophoretic mobility [11] (a function of length of polypeptide chain or molecular weight as well as folding and post-translational modifications). The solution of proteins to be analyzed is initially mixed with SDS, an anionic detergent which disrupts secondary and tertiary structures, and applies a negative charge to each protein in proportion to its mass. SDS linearizes the proteins so that they may be separated strictly by molecular weight. By binding in a ratio of approximately 1.4 g of SDS per 1.0 g of protein, it gives a uniform mass to charge ratio, so that the distance of migration through the gel can be assumed to be directly related to only the size of the protein. The gel, as its name implies, is made of polyacrylamide, a compound obtained by copolymerization of acrylamide and a cross-linker, usually piperazine diacrylyl. The size of the pores is determined by the concentration of the mixed solution. Protein will move differently through the gel matrix: short proteins will more easily fit through the pores in the gel, while larger ones will encounter more resistance, thus separated according to their size.

As stated above, IEF and SDS-PAGE are the two most commonly used techniques for protein separation. It is possible to combine the two techniques into what is referred to **Two-dimensional Gel Electrophoresis (2-DE).** In 2-D electrophoresis [12] proteins are separated by net surface charge at first, separated then by molecular weight as a second dimension. The result is a gel with proteins trapped in it. This protein map is very specific to the biological sample in question. It is very rare that two different proteins have precisely the same net surface charge and the same molecular weight.

1.1.2 Mass spectrometry

Mass spectrometry is an analytical technique in which molecules are ionized in order to measure their mass-to-charge ratio. It is generally used to find the composition of a sample by generating a spectrum representing the masses of sample components. Mass spectrometry is an incredibly valuable tool in the field of proteomics. It can be used to identify proteins through variations of the mass of the analytes. The most common approach to proteomics is a bottom-up approach in which the protein is digested by a

protease, such as trypsin, and the newly formed peptides are then analyzed in order to find out their masses.

A mass spectrometer consists of three basic parts: an ion source, a mass analyzer, and a detector. The ion source is the part of the mass spectrometer that ionizes the analyte. The newly formed ions are transported by magnetic or electric fields to the mass analyzer. Two techniques often used with liquid and solid biological samples include electrospray ionization (ESI) and matrix-assisted laser desorption ionization (MALDI). The mass analyzer separates the ions according to their mass-to-charge ratio. There are many types of mass analyzers, and each type has strengths and weaknesses. Many mass spectrometers use two or more mass analyzers for tandem mass spectrometry (MS/MS). For instance, in a TOF-TOF analyzer two time-of-flight mass spectrometers are used consecutively. The first TOF-MS is used to separate the precursor ions, and the second TOF-MS analyzes the product ions after fragmentation. Between the first and second analyzer there is usually an ion gate (for selecting the precursor ion) and an ion fragmentation region. Finally, a detector records the current produced when an ion hits a surface. The sequence of signals acquired in the detector will produce a mass spectrum. As stated above, there are many ways of coupling the three parts of a mass spectrometer together, with very few limitations. The coupling depends greatly on its final utilization.

Once the peptides masses have been determined the mass list can be sent to a database, where the list is compared to the masses of known peptides, in order to clearly identify the protein. The protein is therefore identified by a number of matching peptides. If the masses of the peptides do not match a known protein, there is the possibility to sequence peptides by "de novo sequencing" in MSMS mode. Another use of mass spectrometry in proteomics is for protein quantification. By labeling proteins with stable heavier isotopes it is possible to determine the relative abundance of proteins [13].

1.1.3 Protein identification

Two approaches are widely in use for high throughput protein identification: Peptide Mass Fingerprinting (PMF) and Peptide Fragmentation Fingerprinting (PFF). Both methods rely on the abundance of sequence information available in gene and protein databases.

Peptide mass fingerprinting (PMF) is an analytical technique for protein identification that was developed in the early 1990's by several groups independently [14]. It utilizes sophisticated algorithms for sequence correlation and databases that contain sequence information. In PMF proteins are first cleaved into smaller peptides using specific cleavage reagents. In a second time the absolute masses of these peptides are accurately measured using a mass spectrometer, such as MALDI-TOF. The same masses are then compared to theoretical peptide masses calculated from a database containing known protein sequences. These theoretical peptide masses are extracted by using algorithms that translate the known genome of the organism into proteins, which are theoretically cut into peptides, in order to calculate their absolute masses. The masses of these peptides are then compared to the experimental ones. The results are statistically analyzed to find the

best match. The mass spectrum acts as the signature of a protein, and it is often sufficient to identify the protein.

However, a single mass value is sometimes not sufficient for unequivocal protein identification. **Peptide fragmentation fingerprint** (PFF, also called MS/MS or tandem mass spectrometry), utilizes additional information, such as internal ion fragment masses, to confirm the identity of a protein or to identify the site of post-translational modifications [14]. Algorithms similar to those of PMF are used for MSMS ion search. Following the same methodology all the proteins contained in the database are theoretically digested to find the matching parents peaks. These same parent peaks are subsequently fragmented in order to be compared to the experimental results. Once again, the correlation of the two determines the outcome. The advantage of MSMS ion search compared to PMF is that fewer, but more precise fragmentation spectra can uniquely identify the protein.

Often, mass spectrometry is combined with **liquid chromatography** (LC) due to its high physical separation capabilities. LC-MS (or LC-MSMS) is a powerful technique used for many applications and presents a very high sensitivity and specificity. Therefore, complicated mixtures can be analyzed directly, despite high differences in concentration magnitude. In most instances the method of choice is RPLC.

2. Shotgun IPG-IEF

2.1 Introduction

2.1.1 Shotgun proteomics

Proteomics deals with highly complex biological compounds. In a traditional separation technique such as Two-dimensional polyacrylamide gel electrophoresis (2D-PAGE), two orthogonal separation methods are interfaced in order to obtain a higher degree of separation. It has been reported that 2D-PAGE has a high enough resolution to be able to detect on a single run over 10,000 protein spots [15]. Nevertheless, an additional analytical step such as mass spectrometry is generally required for the identification of individual spots. Dynamic range and protein solubility remain major issues in 2D-PAGE [16]. Hence, the need for a novel approach to detect, identify and quantify every protein in a sample. Several groups therefore examined the eventuality of replacing one or both gel electrophoretic dimensions with alternative separation methods. The most interesting relates to the use of multidimensional liquid chromatography (LC). Giddings *et al.* [17] demonstrated the overall improvement in peak resolution by means of orthogonal chromatographic separation techniques.

Shotgun proteomics pertains to the bulk proteome digestion followed by multidimensional separation. In most shotgun proteomics analysis the second dimension is performed by RPLC, due to the fact that the mobile phase is compatible with mass spectrometry. Multidimensional chromatography coupled to mass spectrometry has rapidly grown in use and is now routinely part of the shotgun proteomics approach. This methodology was first introduced by Yates *et al*. in 2001[18]. Proteins digested into peptides by proteases are analyzed by multidimensional chromatography coupled to tandem mass spectrometry (MSMS). A single dimension separation does not have enough peak capacity to handle the thousands of peptides originated by the digestion of proteomes coming from complex samples. The inability of one-dimensional separation techniques to resolve complex biological samples for shotgun proteomics has required the development of multidimensional separation methods, which allow for enhanced resolution and peak capacity. Multidimensional separation includes two or more independent separation techniques coupled together for the analysis of a single sample. The vast amount of mass spectra thus generated are then compared to theoretical tandem mass spectra using database search algorithms for the identification of proteins. The acquisitions of these enormous datasets lead to the development of powerful bioinformatics tools capable of quickly and effectively handle the rapidly growing flow of data [19].

There are two main approaches for the application of multidimensional separation methods, offline and online [20]. In an offline approach, the first dimension is not directly coupled to the second one. Fractions from the first column are collected and later submitted to the second column. On the other hand, an online approach employs the

coupling of the two separation methods by means of automation. The fractions from the first dimension are directly eluted onto the second dimension, thus avoiding the need for fraction collection. Online approaches are usually faster than off-line approaches, and sample loss is minimized.

A technique largely employed nowadays relies on strong cation exchange-reversed phase liquid chromatography (SCX/RPLC), better known as MudPIT [21][22][23]. MudPIT has proven to be a useful technique for the quantitative analysis of proteomes and multi-protein complexes [24][25]. It is a fully automated, coupled SCX/RP MSMS approach designed for the analysis of complex peptide mixtures. It has proven to have sufficient separation capacity, when coupled with digestion strategies to generate high sequence coverage of proteins.

Multidimensional separation methods are mostly coupled to an electrospray ionization source (ESI), because they both deal with liquid phase solutions. However, in the recent years an increasing interest has been devoted to the incorporation of liquid chromatography with MALDI mass spectrometers [26]. The LC fractions are deposited onto a MALDI target. The first dimension separation is performed offline and fractions are collected, desalted, and loaded onto a RP column. Eluents from the RP column are then mixed with a matrix online or deposited on the target prior to the matrix spotting phase. An advantage of interfacing LC with a MALDI source is that the rate of collection of MSMS data is decoupled from the chromatographic separation, allowing as much or as little time as necessary for acquiring spectra without fear of missing components as they elute from the LC column.

To sum up, the use of multidimensional separations in the field of shotgun proteomics, as well as advances in mass spectrometry have permitted to gain an enormous amount of information previously impossible to unravel.

2.1.2. IEF

As stated in section 1.1**,** isoelectric focusing (IEF) is a technique for separating proteins and peptides based on their electrical charge differences. The molecules to be focused are incorporated in a gel having a pH gradient. When an electrical field is applied through the medium, molecules migrate until they reach a pH point where their net surface charge is zero, thus not moving any further within the gel. The proteins are focused into sharp bands, positioned at points where the pH gradient corresponding to their pI. In IEF the pH gradient is established by including a mixture of low molecular weight aliphatic ampholytes. These molecules are designed to have specific pKs. By adding the ampholytes it is possible to establish a well defined pH gradient. Even though it was an efficient method, it suffered from a major drawback: the ampholytes would drift toward the cathode during focusing [27]. Also, the gradient was extremely difficult to reproduce. Ultimately, immobilized pH gradients (IPG) were developed, in order to overcome the problems encountered with carrier ampholyte based approaches. The carrier ampholytes are immobilized in the polymerized acrylamide in order to form a fixed pH gradient [28]. Besides preventing the drift the linking ensures that the gels can be cast in an efficient and reproducible manner. In addition, the polymerized ampholytes stay within the gel and

do not contaminate the sample. Agarose gels are preferred to polyacrylamide gels for the separation of larger proteins because of their larger pore size. On the other hand, polyacrylamide gels are preferred for the focusing of smaller proteins and peptides due to their smaller pore size.

IEF as a method of separation presents numerous advantages. First of all, it benefits from high load capacity. Second, large sample volumes do not lower resolution and third, the separation does not require the denaturation of proteins, thus any kind of subsequent investigation (e.g. antibody detection) is not hindered. Not to mention that IEF is capable of very high resolution even with proteins differing by only a single charge.

Other advantages of using IPGs include: high resolution and high sample loading, excellent control over the pH range, ionic strength, buffering capacity, and flexibility regarding the choice of the pH gradient.

2.1.3 Shotgun IPG-IEF

The development of more efficient multidimensional protein separation techniques, followed by an extensive improvement in mass spectrometry instrumentation spurred the development of faster and more accurate methods for protein analysis in complex mixtures [29].

In 2004 Cargile and his team proposed an alternative method based on the utilization of IPG isoelectric focusing as first dimension separation in shotgun proteomics, replacing strong cation exchange (SCX), considered until then a milestone in the Shotgun methodology [30]. Correspondingly, peptides are separated by isoelectric point at first, and by retention time thereafter. Cargile and his team went a step further, developing an accurate pI prediction algorithm that efficiently filters data for peptide-protein identification, lowering thus the false positive rate for peptide identification [31]. They clearly showed that Shotgun IPG IEF aside from being a relatively simple "modus operandi", also leads to a noticeable increase in resolution and sensitivity.

In a comparative study regarding the use of SCX and IPG-IEF as a first dimension separation protocol, that same group demonstrated how narrow range IPG-IEF leads to an increased number of protein identifications in *R. Norvegicus* tissue samples [32]. Moreover, they reported in their studies that IEF appears to have a much greater sensitivity when compared with SCX.

In the Shotgun IPG IEF workflow, after the standard steps of protein purification and digestion, the separation of peptides is carried out by isoelectric focusing using IPG strips that had previously been rehydrated overnight. Once the focusing has ended, the strips are dissected into a number of fractions. The extraction of the peptides from the fractions is accomplished using a series of washes. The peptides are then further separated by RPLC. Directly after, peptides are eluted onto a MALDI target by a spotting robot. Identification data is obtained with the aid of a MALDI TOF-TOF instrument.

In order to speed up the time frame of the analysis and to leave as little variability between experiments, it is imperative to automate the pipeline as much as possible. Moreover, the fractionation process remains the weakest link in the chain. Due to the fact that the repartition of peptides along the strip is everything but equitable, a great improvement would be to fractionate the strips so to have an equal number of peptides between fractions.

A critical step is the interpretation of the enormous amount of datasets generated. The identification of peptides is carried out using their intrinsic properties, by matching empirical values with theoretical ones, obtained from protein sequence databases. Due to the increasing discovery of new proteins, and to the improvement in detection limits by shotgun technologies, a powerful bioinformatics software is needed in order to analyze rapidly and precisely this huge amount of data. Phenyx (Genebio, Geneva, Switzerland) and Mascot (Matrix Science Ltd., Boston, U.S.A.) proved to be user-friendly and extremely powerful search engines.

As shown in Figure 1, the workflow starts with the purification of the protein mixture and digestion with endoprotease, generally Trypsin. Consequently, the newly formed peptides are loaded onto an IPG strip and are separated by IEF. Once the focusing has ended, the strips are cut into a predetermined number of fractions and placed each in an Eppendorf tube. The peptides are extracted from the gels and loaded on a RPLC column for the second dimension separation. Each fraction is successively eluted onto a MALDI target with the aid of a spotting robot. The MSMS acquisition is performed using a 4800 MALDI TOF-TOF mass spectrometer. Any ESI platform can also be used in combination with the Shotgun IPG-IEF platform. Finally, the MSMS data obtained are analyzed with the aid of bioinformatics. Software such as Phenyx (Genebio, Geneva, Switzerland) or Mascot (Matrix Science, Boston, U.S.A.) are used for the protein identification.

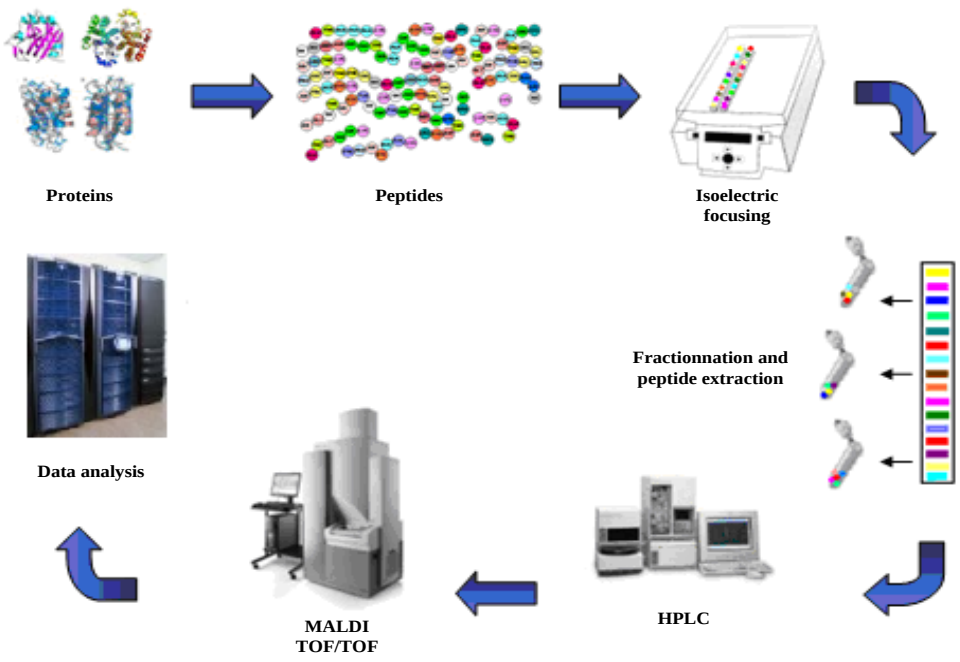


Figure 1. The Shotgun IPG-IEF workflow. The proteins are first digested into peptides before being focused by IEF on IPG strips. Once the peptides have focused, the strips are cut into a predetermined number of fractions and the peptides are extracted. Subsequently, the peptides are eluted onto a MALDI target by RPLC using a spotting robot. The samples are scanned in a 4800 MALDI TOF-TOF instrument. The last step involves the analysis of the acquired data.

2.2 Imaging Shotgun IPG IEF

In the mid 1990's Caprioli and co-workers introduced a novel method for tissue imaging using MALDI mass spectrometry [33]. What was new in this technique was the possibility of localizing or following changes in organisms at the molecular level by imaging component distribution of specific tissues. The potential of this methodology is obvious [34][35]. MS Imaging is a promising field in proteomics because it provides information regarding the spatial arrangement of molecules within a defined region.

The MS image is created by rastering sequentially the surface of a tissue section while acquiring a mass spectrum from every point. Usually, each mass spectrum is the average of the number of shots taken. The result is a molecular weight specific map of the distribution of proteins along the tissue section. For this specific method of acquisition, the MALDI-MS instrument has to be equipped with specific software, capable of not only creating, but also storing the resulting data. The conversion of the data into an image has to be performed by specific bioinformatics tools.

An interesting alternative to tissue imaging is referred to as the molecular scanner, which is aimed at visualizing and characterizing biological samples at a molecular level using

1D or 2D-PAGE [36]. The proteins in the gel are transferred through an enzymatic membrane onto a collecting membrane. The enzymes present in the first membrane digests the protein into peptides. The collecting membrane containing the digested peptides is subsequently coated with matrix and scanned in a MALDI TOF mass spectrometer. A software reconstructs the image of the original gel and provides identification of the proteins presented in the gel [37].The advantage of such a technology is that minimal sample handling is required. Furthermore, the sample present on the collecting membrane is stable and can be reused for further analysis if necessary, for the rate of peptide diffusion is very low.
The molecular scanner has proved to be a powerful tool for molecular imaging and protein identification and has the advantage of relying on the specificity and sensitivity of mass spectrometry.

A novel methodology has been recently developed by the Biomedical Proteomics Research Group at the University of Geneva. The so-called "Imaging Shotgun IPG IEF" (ISII) is based on blotting the IPG strip onto a porous support, such as a PVDF (Polyvinylidene Difluoride) membrane. The peptides pass from the gel onto the membrane by capillarity. The membrane is then dried at room temperature for a few minutes and adhered onto a MALDI target using a double face adhesive tape. CHCA matrix is deposited on the surface of the PVDF membrane by a spotting robot. The acquisition step is performed by shooting the laser directly onto the membrane in a MALDI mass spectrometer. The obtained spectra are concatenated to form an image of the distribution of peptides with respects to their pI and mass to charge ratio. Such an image allows the visualization of the peptide distribution along the membrane.

This method allows the identification of areas of high and low peptide density. The most interesting aspect is the possibility of rapidly obtaining an image of the distribution of peptides in a sample, which could be used for the detection of differences. The bypassing of time consuming chromatographic separation shortens the pipeline time frame to exactly one working day.

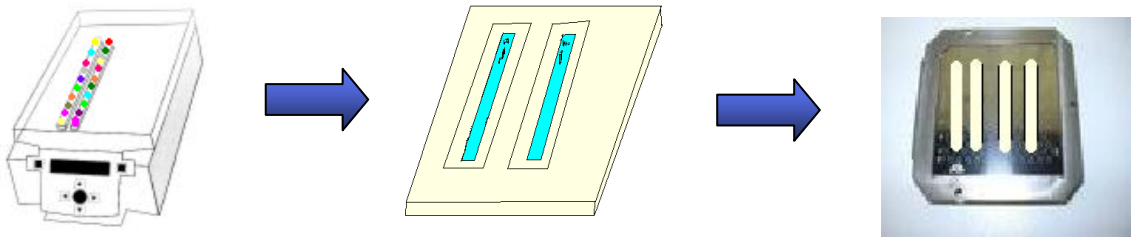


Figure 2. The Imaging Shotgun IPG-IEF workflow. After focusing, peptides are transblotted from the IPG strip to a PVDF membrane which is attached on a MALDI target, covered with matrix and scanned in the MALDI instrument.

2.2.1. MSight

MSight is a tool specifically developed for the representation of mass spectra along with data from the LC step. MSight was created by the Proteome Informatics Group (PIG) at the Swiss Institute of Bioinformatics (SIB) [38]. Images obtained from high-throughput mass spectrometry contain information that remains hidden when looking at a single mass spectrum at a time. By concatenating the spectra one after the other in a visual presentation, a clear image of peak distribution is generated. The importance of imaging in differential analysis of proteomic experiments has already been established through 2D gels and can now be foreseen with two dimensional MS.

The Key futures of MSight are:

- Advanced zooming for the display of data at various resolutions without information loss.

- Simultaneous multiple image display. This facilitates comparison of data from various experiments or experimental conditions.

- Multiple images alignment through the use of landmarks to compensate for differences in elution time or migration distance.

- Individual pixel annotation. Comments related to the experimental protocol can also be added to the images.

Originally developed for image processing of LC-MS datasets, MSight became quickly useful for other applications. In Imaging Shotgun IPG-IEF, the resulting spectra are brought together in one single image. On the horizontal axis the mass over charge and on the vertical axis the pI are represented.

The most interesting future of MSight is a comparison function “Differential Display”, which allows the direct comparison of the generated images. It also facilitates the comparison of data from different samples or various experimental conditions.

2.2.2 Biomap

Biomap is an image processing application for data imaging (www.maldi-msi.org). The software was developed by M. Rausch and based on IDL (Research Systems, Boulder, CO) provides specific tools for MS image analysis. One advantage of Biomap is that it allows to select single points or regions of interests (ROIs) on the generated image and to display the corresponding mass spectrum. Another advantage is the possibility of selecting specific masses of interest and to calculate by integration over the corresponding peak its distribution on the scanned area. Hence, the resulting image gives at the same time the location and the intensity of the corresponding MS signal.

In the Imaging Shotgun IPG-IEF workflow, Biomap is used to visualize the distribution of peptides at the surface of the PVDF membrane. A key feature is the visualization of m/z values of choice, allowing a more in depth comparison analysis.

To sum up, MALDI MSI is a very promising analytical tool for biomedical research. From a technical point of view, the technique benefited from considerable improvements (e.g. the reduction of the laser spot diameter, which could end up with a higher lateral resolution [39]). Matrix coating is the step which requires improvement. For this reason alternative matrix deposition methods are currently under development [40].

2.3 Peptidic fluorescent markers

In isoelectric focusing the pI of a specific protein can be established by measuring the pH of the focused area in the gradient, or buy the utilization of standard molecules with a stable well-established isoelectric point. Originally proteins were used as pI markers, but problems encountered in their use limited their applicability. First, the instability of proteins led to changes in their pI. The causes of the pI shift can be related to the hydrolysis of side chain amides of asparagine or glutamine residues [41], as well as to the hydrolysis of peptide bonds. Moreover, protein denaturation and the three-dimensional structure loss can further amplify the pI shift. The result is that not only the marker may shift from its original position, but it may degrade and form multiple bands, rendering the subsequent analysis quite ambiguous.

An attempt to improve the stability of pI markers was made with the introduction of low molecular mass amphoteric dyes. Problems encountered with this compounds ranged from marker precipitation to covalent interaction with other molecules present in the gel [42][43]. Stastna *et al.* presented a new set of improved dyes that overcame the limitations of previous color pI markers [44]. Nevertheless, the limitation of colored pI markers lies in the pH range. It is quite hard to find suitable markers that cover a wide pH range, especially in the basic region.

Shimura *et al.* developed synthetic oligopeptides to be used as pI markers [45]. Peptides with any desired amino acid composition can be promptly synthesized. The composition of their ionic side chain ensures that a wide range of pIs can technically be attained. Due to the fact that each ionic group ionizes independently, the resulting pI value of the peptide as well as the sharpness of its focusing can be precisely predicted. Peptide markers overcame most of the limitations that predecessors suffered from, rendering them extremely practical for isoelectric focusing.

In order to be used as reference points for the determination of peptides and proteins in IEF, the markers have to be detectable. Therefore the peptides have to be labeled with a molecule that reveals their presence. Shimura also reported the use of fluorescence-labeled peptides in CIEF [46]. The principle is that a fluorescent molecule is covalently linked to the peptide. The drawback of this methodology is that the fluorophore attached to the peptide changes its pI, since the labeling can potentially change the acid-base properties of the peptide in question. Nevertheless, an estimate can be computed and a confirmation obtained by experimental procedures.

Fluorescent labeling is a sensitive and quantitative technique. It is widely used in molecular biology and biochemistry for analytical applications. Nucleic acid and protein quantification, as well as blotting techniques (e.g. Western), take advantage of fluorescence based detection methods [47] [48].

Compared to other detection methods, fluorescent detection offers some non negligible advantages, such as:

- **High sensitivity** (allows the detection of low abundance molecules)
- **Multiple label possibility** (multiple fluorochromes can be detected separately)
- **Stability** (compared to other labeling techniques, such as radiolabeling, fluorescently labeled reagents can be stored for long periods)
- **Low hazard**
- **Low cost**

A commonly used fluorescent molecule is fluorescein. Fluorescein is an organic fluorophore commonly used in microscopy [49]. It is easily available, simple to detect and can be measured using its strong fluorescence and highly absorptive character. However, the molecule appears to be photochemically instable. For this reason, it is recommended to avoid long exposure of the molecule to direct sun light.
For the detection, a laser scanner can be used to measure the intensity of the fluorescent light, and can consequently create an image of the sample in question. In IEF the fluorescein labeled peptides focalize at their pI forming a horizontal band. If the marker is present in the sample in high enough concentration, it is clearly visible with the human eye. Otherwise a laser scanner can help in the detection process.

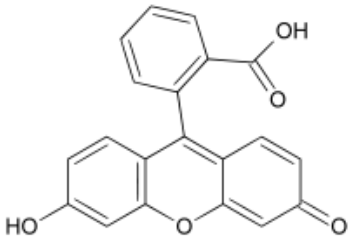


Figure 3. The fluorescein molecule

3. Materials and methods

3.1 Reagents and chemicals

The reagents used are of standard quality, except for the acetonitrile (Fluka), which has HPLC quality. Water was purified by the Millipore's MilliQ system or LiChrosolv® water was used (Merck). Products were purchased from the following companies:

- Applied Biosystems (Framingham, MA, USA)
- Applied Microbiology Inc (Tarrytown, MA, USA)
- BioRad (Hercules, CA, USA)
- Difco (Detroit, MI, USA)
- Fluka (Buch, Switzerland)
- GE Healthcare (Piscataway, NJ, USA)
- Merck (Darmstadt, Germany)
- Millipore (Bedford, MA, USA)
- Schleicher & Schuell (Dassel, Germany)
- Sigma Aldrich (St. Louis, MO, USA)
- Waters (Milford, MA, USA)

Bovine serum albumin (BSA), porcine trypsin, bovine carbonic anhydrase, bovine β-casein, bovine β -lactoglobulin, rabbit phosphorylase b, α-cyano-4-hydroxycinnamic acid (CHCA), acetonitrile (AcN), formaldehyde (37%), trifluoroacetic acid (TFA), 1,4-dithioerythritol (DTE), ammonium bicarbonate (BA), iodoacetamide and Tris were purchased from Sigma-Aldrich. SDS-PAGE precast gels 4-20% Tris-HCl, ampholines (4-7 and 3-10), Sequi-Blot™ 0.2 µm pore size PVDF membranes and molecular mass markers came from BioRad. Ethanol, formic acid (FA), high boiling-point petroleum ether, acetic acid, glycine and SDS came from Fluka. Chloroform, methanol, saccharose and urea were provided by Merck. Immobiline™ DryStrips and PlusOne DryStrip Cover Fluid paraffin oil were purchased from GE Healthcare. Mueller Hinton broth came from Difco and hydrolytic enzyme lysostaphin (Ambicin) was purchased from Applied Microbiology Inc.

3.2 Sample preparation

3.2.1 Growth conditions and time point

To obtain protein extracts, Staph. *aureus* strain S_{30} was grown in Mueller Hinton broth (MHB; 200 ml in 1000-ml flask) with agitation at 37°C, as previously described [50]. When the post-exponential phase was reached (OD_{540nm}=6 corresponding to 2-3 x 10^9 cells/ml), cells were chilled on ice and harvested by centrifugation at 8'000 x *g* for 5 minutes at 4°C. For the preparation of total protein extracts, 20 ml culture aliquots were

washed in 1.1 M saccharosecontaining buffer [51] and then suspended in 2 ml aliquots of the same buffer containing 50 µg/ml of the hydrolytic enzyme lysostaphin for 10 minutes at 37°C. For preparation of membrane extracts, protoplasts were recovered after centrifugation (30 minutes at 8'000 x *g*) and hypo-osmotic shock was applied in the presence of 10 µg/ml DNase I (Fluka)to decrease the viscosity of the medium. Membrane pellets were obtained after ultracentrifugation at 110'000 x *g* for 50 minutes in a Beckman Optima TLX (Beckman Coulter Int'l S.A., Nyon, Switzerland).

3.2.2 Chloroform precipitation

In the delipidation process protein extracts were evaporated by speed-vac and resolubilized in 100 µl 50mM BA pH 8.5 per mg of crude protein extract. 1 ml of a chloroform/methanol (2:1,v:v) solvent was added, thoroughly vortexed and placed on ice, before being centrifuged at 4°C for 15 minutes at 14'000 rpm. The lower phase containing the $CHCl_3$ was carefully extracted and the supernatant was resuspended in 300 µl of MeOH, thoroughly vortexed and centrifuged at 4°C for 20 minutes at 14'000 rpm. The supernatant was then extracted and the pellet placed in the speed-vac to discard the remaining MeOH. Then the proteins were resuspended in 300 µl BA 50 mM and 2 µl were diluted in 8 µl of H_2O for the SDS-PAGE control gel.

3.2.3 Digestion protocol

Reduction, alkylation and digestion took place in a domestic microwave oven (FUNAI, Hamburg, Germany) with a maximum output power of 850 W and a frequency of 50 Hz but the oven was set on reduced power (~175 W) for all steps. The samples were placed in 1.5 ml Eppendorf tubes in a home made holder placed in a beaker containing 500 ml of water at 25°C and the irradiation was done during 6 minutes each time, which resulted in a gradient of temperature from 25 to ~55°C. For a mg of proteins, the reduction was done by adding 40 µl of DTE 45 mM and the alkylation by addition of 90 µl of iodoacetamide 100 mM. After the alkylation the samples were set on ice to cool down for better digestion. The trypsin enzyme was added for digestion at a protease-to-protein ratio of 1:10. When deemed necessary, a double digestion was performed, with a ratio of around 1:15 the first time and of 1:25 for the second. 8 µl of the peptides were taken and diluted in 2 µl of H_2O for the control gel.

3.2.4 Purification

After digestion peptides were concentrated and desalted using an Oasis HLB 1 cc 10 mg solid phase extraction cartridge (Waters). 800 µl of 0.1 % FA were added to the sample and the pH was verified with pH paper (pH 0-14, Merck). If the pH was not around 2-3, 1 to 5 µl of pure FA were added. The column was first equilibrated with 1 ml of 0.1 % TFA 60 % AcN and then equilibrated with 1 ml 0.1 % FA. The sample was passed slowly, washed with 1 ml 0.1 % FA and eluted in 700 µl 0.1 % TFA 60 % AcN. The samples were then evaporated to dryness using a speedvac and then re-suspended in 50 µl of H_2O and re-evaporated to ensure all FA was discarded.

3.3 Fluorescein coupled peptides

Peptides were prepared manually by Solid Phase Peptide Synthesis using standard Fmoc/ tBu strategy. On 0.05 mmol Rink amide 4-methylbenzhydrylamine resin (Fluka) were coupled 0.25 mmol Fmoc-protected amino acids activated with 0.24 mmol HBTU in the presence of 0.3 mmol DIEA. Couplings were allowed to react for 60 min with occasional stirring. Fmoc-amino acids were protected by the following groups: Arg(Pbf), Asp(OtBu), Cys(Trt), Glu(OtBu), His(Trt), Lys(Boc). Removal of Fmoc protecting group was done with a 20% piperidine solution in DMF for 5 and 15 min, followed by DMF wash. Acylation of peptide was done with 0.5 mmol acetic anhydride in the presence of 0.6 mmol DIEA. Peptides were cleaved from the resin with 4 ml of TFA solution containing 3% water and 3% triisoporpylsilane (Aldrich, Switzerland) as scavengers. After 4 hour, reaction mixture was filtered and the resin was rinsed twice with 2 ml TFA. Solution was concentrated by TFA evaporation and peptides were precipitated and washed with cold diethyl ether before lyophilization. Peptide sequences after deprotection were: DDEHACG-NH_2, Ac-DHHACG-NH_2 and RKHGCA-NH_2 for respectively peptides 1, 2 and 3. Each peptide was checked for purity by analytical HPLC and MALDI-TOF MS.

The peptides were dissolved in water at concentrations of 2 or 3 mg/mL. pH was controlled by addition of 3 µL of 1 M triethanolamine-bicarbonate. A 10 mg/mL solution of iodoacetamido fluorescein in DMF was added in a 20% molar excess. The solution was left at room temperature for 10 min and subjected to microwave heating. The tube was placed in a beaker with 500 mL water and subjected to irradiation in a kitchen microwave oven for 6 min at 175 watts. The temperature of the bath rose to 57-59°C. After cooling the solution were subjected to purification.

Purification of fluorescent peptides was done by reverse-phase HPLC on Waters equipment using a Macherey-Nagel C8 column (4 x 250 mm 300 Å 5 µm particle size) at 0.6 ml/min. Solvent A was 0.1% TFA in HPLC grade water. Solvent B was 90% acetonitrile with 0.1% TFA. Elution was done with a 60 min linear gradient 20-80%B.

Preparative TLC was performed on Silicagel 60 devoid of fluorescent indicator. The solution of fluorescent mixture corresponding to 150 µg of peptide was distributed on a 9 cm line and a first migration was obtained with a 2:1 mixture of $CHCl_3$ / Methanol in order to remove unreacted fluorescein derivative, the peptide remaining at the origin. . A second migration was obtained with the following solvent mixtures: AcN/Me_2CO/AcOH/H_2O: 30:10:2:20 for marker 1, Me_2CO/H_2O/NH_4OH_{con}: 60:12:1.5 for marker 2 and Me_2CO/H_2O/NH_4OH_{con}: 51:22:7.5 for marker 3. The silica containing the fluorescent peptide was scraped and extracted with 50% trifluoroethanol supplemented according to the peptide nature with acetic acid for markers 1 and 2 and with ammonia for marker 3

3.4 Electrophoresis

3.4.1 IPG-IEF

After purification, samples were re-suspended in 100 µl of rehydration buffer. Immobiline™ DryStrip 7, 18, or 24 cm, pH 3-10 or, pH 4-7 strips (GE Healthcare) were rehydrated overnight using the Reswelling Tray (GE Healthcare). Isoelectric focusing was performed on an Ettan IPGphor II system (GE Healthcare). Paper wicks (GE Healthcare) soaked in 145 µl H_2O were used for the connection between the strips and the electrodes and the whole was covered in 100 ml of DryStrip Cover Fluid paraffin oil (GE Healthcare). The focusing was done with the following conditions (for the 7cm strips): 5 minutes step at 100 V, 30 minutes linear gradient from 100 V to 500 V, 30 minutes linear gradient from 500 V to 1000 V, 30 minutes step at 1000 V, 30 minutes linear gradient from 1000 V to 5000 V and step at 5000 V up to a total of 9 kVh.

The temperature was set at 15^{o}C and the current to 60 µA per strip.

Rehydration Buffer:
- 4M or 8MUrea (Merck)
- 0.2% Pharmalyte 3-10 or 4-7 (Amersham)
- 10 µl bromophenol blue (Fluka)
- LiChrosolv® H_2O to 10 ml

3.4.2 SDS

The proteins and peptides were solubilised in 10 µl of Laemmli buffer and reduced by heating at 95°C for 5 minutes. A volume of 20 µl was loaded in each well of a precast 4-20 % Tris-HCl gradient gel (BioRad). All SDS-PAGE gels were done on a Miniprotean II BioRad SDS-PAGE System and the separation took place at a constant voltage of 200 V for about 30 minutes in 1l of running buffer. Once the run was terminated, MS compatible silver staining was done as described by Allard *et al.* [52].

Laemmli Buffer:
- 2% SDS (Fluka)
- 0.025% bromophenol blue
- 10% Glycerol (Merck)
- Trizma Base 50mM , pH 6.8 (Sigma)
- 0.5%v/v _-mercaptoethanol (Millipore)
- MilliQ H_2O to 100ml

Running Buffer:
- Trizma Base 100mM
- Glycine 100mM (Fluka)
- SDS 1% (v/v)
- MilliQ H_2O to 1 l

3.5 Transblot

The transfer was performed by capillarity. Filter papers (Schleicher & Schuell) and 0.2 µm PVDF membranes (BioRad) were used.

Transfer Buffer (10X): Trizma Base pH 8.3, 125mM
Glycine 960mM

Three filter papers of 10 x 8 cm were cut and one was soaked in the transfer buffer for 10 minutes and then thoroughly blotted. For each IPG strip, an 8 x 1 cm PVDF membrane was first soaked in methanol for 10 minutes and then rehydrated by total immersion in the transfer buffer. Two 100 x 50 cm pieces of commercially available cellophane film were cut and placed flat in a cross shape. A 20 x 20 cm glass plate with the 2 dry filter papers in the middle was placed in the centre of the cross. As soon as the focusing was finished, the soaked filter paper was placed on top of the two dry ones, the PVDF membrane was placed onto the filter papers (up to 4 membranes per paper) and the IPG strip was placed in the centre of the latter. The sandwich was closed by carefully placing a second 20 x 20 cm glass plate on top and the whole was made hermetic by folding each arm of the cross. A 1.25 kg weight was very carefully placed on top in the middle and the transfer was done during 120 minutes at room temperature. The pressure on the strips was about 16 grams per cm^2.

3.6 Peptide extraction and purification

3.6.1 From the IPG strip

Once focusing accomplished, the strip was washed 3 times 20 seconds in high boiling-point petroleum ether to remove the paraffin oil. During the fractionation each fraction was put into 0.5 ml Eppendorfs containing already 100 µl of 0.1% TFA and put onto the agitator for 30 minutes after having been vortexed. The 100 µls were removed and placed in a clean Eppendorf. The process was repeated twice for 20 minutes and the final volume, i.e. 300 µl, was frozen. The peptides were then purified on Oasis 96-Well µElution Plate (Waters). The plate was washed and equilibrated with 200 µl of 0.1% FA 60% AcN and then with 200 µl of 0.1% FA. Samples were slowly passed, washed with 200 µl of 0.1% FA 5% AcN and eluted in 2 x 50 µl of 0.1% FA 60% AcN.

3.6.2 From the PVDF membrane

After the transfer, the membranes were scanned in a Voyager DE-STR MALDI-TOF (Applied Biosystems) and the fractionation of the membrane was done in regards to the desired peptide position, calculated from the markers. Each fraction was put into 0.5 ml Eppendorfs containing 100 µl of 0.1% TFA 50% AcN and put onto the agitator for 20 minutes after having been vortexed during 5 minutes. The 100 µl were removed and

placed in a clean Eppendorf. This was repeated twice for 20 minutes, only the second time the extraction was done without AcN. The final volume, i.e. 300 µl, was evaporated and the peptides were resuspended in 300 µl 0.1% TFA. The peptides were then purified on Oasis 96-Well µElution Plate (Waters) using the same protocol as for the IPG strip fractions described above.

3.7 Mass spectrometry

3.7.1 MS imaging

At the end of the transfer step, the capture membrane was cut in two and applied on a MALDI target containing no wells (modified by Applied Biosystems) with double sided tape (3M). The MALDI-TOF matrix was applied with a home-made spotting robot. Mass spectra were acquired on a voyager DE-STR MALDI-TOF mass spectrometer (Applied Biosystems) equipped with a 337 nm UV nitrogen laser, a delayed extraction device and an acquisition rate of 20 Hz. The acquisition was performed with an acceleration voltage of 20 kV, a grid of 63% and a delay extraction time of 180 nanoseconds. The mass range was defined from 800 to 3000 Da with a low-mass gate fixed at 800 Da. A blank target was selected as the plate file on the Voyager 4.3 acquisition software. The exact position of the membranes on the plate was determined by defining the margins of each one. A spot set was created that defines the area to be scanned. In order to obtain a good representation of the repartition of the peptides without too much data and as the diameter of the laser on the membrane was about 50 µm, each acquisition was spaced by 250 µm (or 150 in some instances). For each spot, 100 spectra were accumulated.

In the "Automated" section of the method set-up, the spot file was selected and the number of spectra to be acquired was equaled to the number of points in the spot set file and the "Save All Spectra" option was selected, saving all the spectra from points defined in the spot file into a unique ".dat" file. Once the data was collected, they were transferred to MSight (SIB, Switzerland) for visualization.

MALDI-TOF matrix:	10 mg/ml of CHCA in 70% MeOH-1% TFA 10 mM $NH_4H_2PO_4$

The MALDI target was introduced into the 4700 MALDI-TOF/TOF (Applied Biosystems) and a template was created in regards to the position of the membranes on the plate. After tuning MS spectra were acquired each 250 microns (or 150). Each MS spectrum was individually visualized and the precursor masses were manually selected at random positions for MS/MS. After the MS analysis, a peak list was created for each fraction with the 4700 explorer 2.0 peak-to-Mascot embedded software with these settings: peptide mass range from 60-toprecursor minus 20, minimum S/N 0.5 and maximum 200 peaks per precursor.

3.7.2 LC-MS/MS

After extraction and purification, samples were resuspended in 20 µl of solution A and a volume of 5 µl of peptide solution of each fraction was loaded on a 10 cm long home-made column with an ID of 100 µm, packed with C_{18} reverse phase (YMS-ODS-AQ200, Michrom Biosource, CA, USA). The elution gradient of the LC ranged from 4% to 38% solvent B (Solvent A: 3% AcN, 0.1% FA, Solvent B: 95% AcN, 0.1% FA) was developed in 40 minutes and samples were eluted directly onto a MALDI target plate using a home-made spotting robot.

MALDI-TOF/TOF Matrix was then applied and allowed to dry in a speed-vac. Peptides were analyzed in MS and MS/MS mode using a 4800 MALDI-TOF/TOF, with a Nd:YAG laser at 355 nm operating at 200 Hz repetition. 800 consecutive laser shots were accumulated for MS and 1500 for MS/MS. For the CID, Argon gas was used, at a gas pressure of 4-8 x 10^{-7} torr. Data-dependent MS/MS analysis was performed automatically on the 10 most intense ions from MS spectra. External calibration with lysozyme C was done in MS and MS/MS (*m/z* 1753.6) when judged necessary. After the MS/MS analysis, a peak list was created as explained above.

MALDI-TOF/TOF matrix: 5 mg/ml of CHCA in 50% ACN-0.1% TFA
10 mM $NH_4H_2PO_4$

3.8 Data analysis

3.8.1 Protein identification

In the LC-MS/MS analysis with *S. aureus* samples, peak lists of all fractions of the same strip or membrane were merged before database searching with Phenyx (GeneBio, Switzerland). The searching was performed against a home-made database containing non-redundant predicted ORFs from genome-sequenced strain N_{30} with 90% homology with other strains [53]. On the Phenyx submission webpage MALDI-TOF/TOF was selected as instrument type. The taxonomy selected was "other Firmicutes". Two search rounds were selected, both with trypsin as the proteolytic enzyme, oxidized methionine as variable modification and carbamidomethylation of cysteine as fixed modification. In the second round deamidation was also selected as variable modification. In the first round, one missed cleavage with normal cleavage mode was selected whereas in the second round three missed cleavages with half-cleaved node were selected.

"Turbo" was selected only in the first round, with a tolerance of 0.4 Da, a coverage of more than 0.2 and *a*, *b* and *y* ion series. The minimum peptide length allowed was 5 amino acids. Parent ion tolerance was 1 Da in the first round and 0.4 Da in the second. The acceptance criteria were slightly lowered in the second round search (1st round: AC score 8.0, peptide Z.score 6.5 and pvalue 1.0E-7; 2nd round: AC score 8.0, peptide Z-score 6.0 and p-value 1.0E-7). For direct MS/MS on protein standards, the peak lists obtained from the pool were submitted to Mascot (MatrixScience, USA). Searching was

performed against UniProtSP database. On Mascot submission webpage MALDI-TOF/TOF was selected as instrument type. The taxonomy selected was other Mammalia. Trypsin as selected as the proteolytic enzyme, oxidized methionine as variable modification and carbamidomethylation of cysteine as fixed modification. Two missed cleavages were selected, as well as monoisotopic mass values. The peptide mass tolerance was set to ± 2 Da and the fragment mass tolerance to ± 1 Da.

3.8.2 Visualisation

Regarding the data obtained from the voyager DE-STR MALDI-TOF mass spectrometer, a unique ".dat" file containing all spectra was imported into the MSight software. Each spectrum is automatically concatenated to its neighbours using the "Concatenate Images" function, thus creating an image with the *m/z* ratio on the *x* axis and the number of the spectra, corresponding to its position on the membrane and therefore to the pI, on the *y* axis. Once the image was obtained, it was "cleaned" by using the "Remove Background from Image" function, as well as the "Normalise Images using the TIC" for a harmonisation of the spectra. Once these functions were used, the contrast of the image was also adapted. The Differential Display future was used for the superimposition of the images. Up to 6 landmark points were used as reference.

For BioMap imaging, the MALDI MS Imaging software (Novartis) was used on the Voyager DE-STR MALDI-TOF. Once the membrane area was defined, the instrument was set to acquire every 250 µm vertically and horizontally, thus creating an image of the whole membrane (about 2500 spectra). The MS parameters were the same as for normal MS imaging. The data was then imported on the BioMap 3.7.5.2 software to visualize the total ion image of the membrane and to select various peptides for localization.

4. Results and Discussion

Initially, for the focusing of the markers fluorescent peptidic we decided to use BSA peptides. BSA is easy to obtain, and presents a fairly simple spectrum in mass spectrometry. It was possible to unambiguously identify BSA peaks and fluorescent peptidic markers peaks. Subsequently, in order to obtain a more homogeneous gradient in the IPG strip, we substitutes BSA peptides with *E.Coli* peptides. For the comparison of ISIEF images it was important to find samples that presented a very similar protein profile, for the identification of differences at the protein expression level by comparing the images. We obtained *S.aureus* S_{30} proteins from Dr. Patrice François at Genomic Research Laboratory (Service of Infectious Diseases, University of Geneva Hospitals). The total protein extracts comprised the Wild Type, a mutant strain lacking the helicase gene H(-), and that same mutant strain complemented for the helicase gene H(+). Prior to the utilization of the bacterial strains, an SDS-PAGE gel was cast to determine protein profile differences (figure 4).

Figure 4. SDS-PAGE of *S.aureus* S_{30} bacterial strains: (A) Wild Type, (B) H(-), and (C) H(+). The staining was performed using MS-compatible silver.

The three strains showed a very similar protein profile. From the SDS-PAGE gel alone, clear differences were not noticeable. We therefore decided to use the samples with the hope of finding out differences in imaging experiments.

4.1. Sensitivity test

4.1.1 Results

A simple test was designed to compare the sensitivity of the two techniques: Shotgun IPG-IEF (SIEF) and Imaging Shotgun IPG-IEF (ISIEF). It was important to discover the limit of peptide detection of each approach in order to maximize peptide recovery. For the sensitivity test BSA peptide 927 m/z was chosen. Four samples were prepared (figure 5). Fluorescent marker at pI 4.25 (GE-Healthcare) was added. From an initial BSA concentration of 50 pmol/µl, a series of dilutions was performed to obtain the lowest concentration of 50 fmol/µl. Thus, the concentration of BSA in the samples was the following: 50 pmol/µl, 10 pmol/µl, 1 pmol/µl, 100 fmol/µl. The fluorescent marker was used to find the positioning of the peptide on the membrane and on the gel. The distance between the marker and the peptides was calculated according to the pICarver software (www.expasy.org/tools/pICarver). The fraction including peptide 927 m/z was excised from the membrane and from the gel. The two pipelines were separately resumed.

	pI	Mass (da)	Position (cm)
pI marker	4.25	N/A	0.5
Peptide 927	5.59	927	4.4

Figure 5. BSA peptides and the 4.25 pI GE marker used for the sensitivity test.

The MALDI-TOF results showed that in both techniques the signal generated by BSA peptide 927 m/z was observed in the targeted fraction, up to the 10 pmol level (figure 6); up to the 1 pmol level in SIEF. Therefore, it appeared that by extracting the peptides directly from the gel, an increased sensitivity can be obtained.

Considering that the sample was dispersed along the whole strip (7 cm) and that the gels and membranes were cut in exactly 10 fractions, the real amount of analyzed material in 10 mm fraction was 10-30 fmol, which is an acceptable sensitivity level for such a technology.

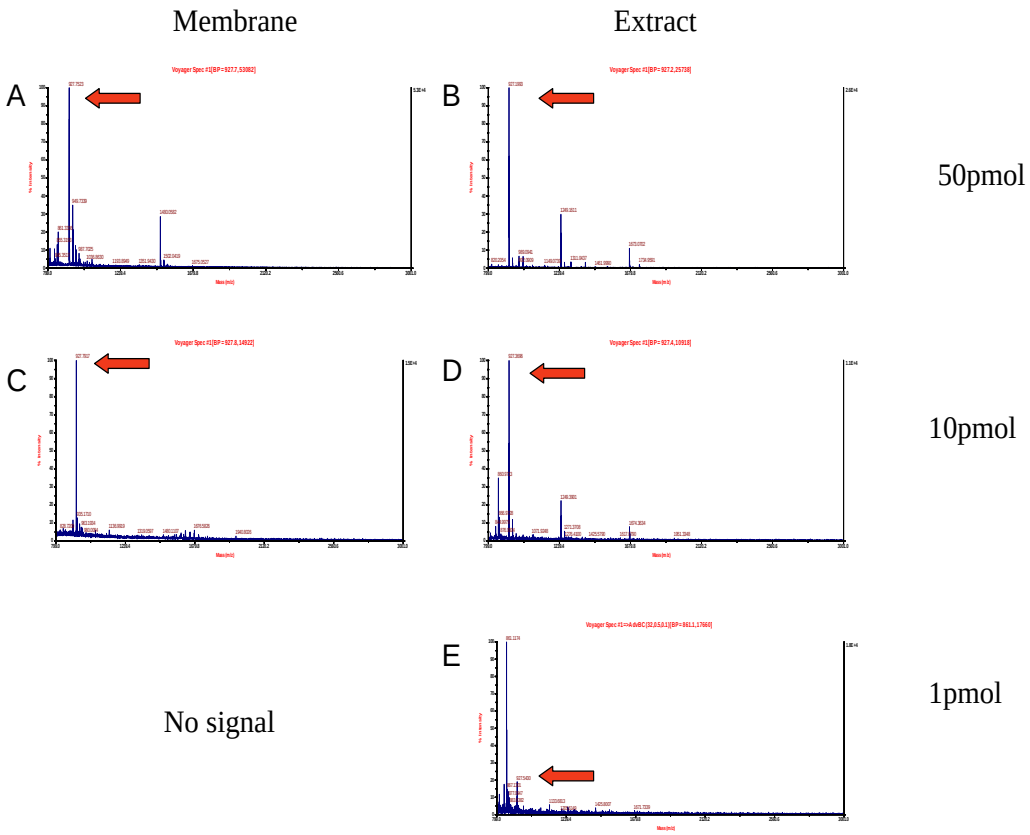


Figure 6. (A) and (C) represent MS spectra of BSA peptide mass 927 from the PVDF membrane. (B), (D), and (E) represent MS spectra of BSA peptide mass 927 from gel extraction

4.1.2 Discussion

The scope of this experiment was to compare the sensitivity of a technique that uses direct peptide extraction from the IPG strip (SIEF) versus a technique in which peptides are transferred onto a PVDF membrane by capillarity (ISIEF).

The transfer step is crucial in the imaging shotgun approach. It is supposed to preserve the spatial distribution of the peptides. The PVDF membrane has the advantage of being easy to manipulate and does not have to be frozen for storage purposes. In fact, even if preserved at room temperature and at normal atmospheric pressure for up to a week, no significant peptide diffusion is noticed. Nevertheless, during the transfer some peptides were not able to migrate toward the membrane and remained in the gel. It is therefore plausible that direct gel extraction has a slightly better yield. In a similar experiment aimed at comparing peptide recovery from selected fractions from the membrane as well as from the gel, we showed that after normalization of the pI window, in a $1cm^2$ fraction only 122 peptides were recovered from the membrane, compared to the 306 recovered from the gel. Peptide extraction from the membrane remains clearly a difficulty.

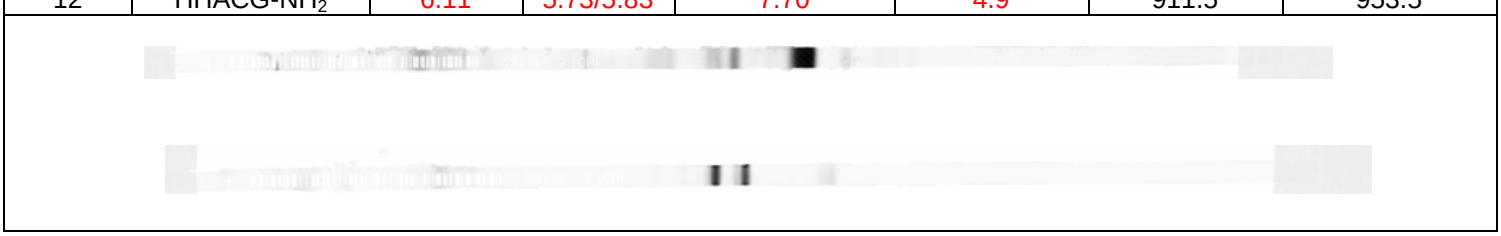
4.2. Fluorescent peptidic markers development

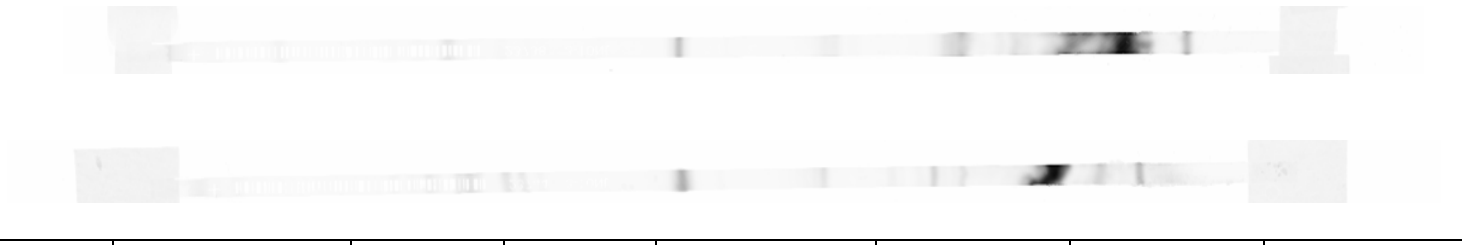
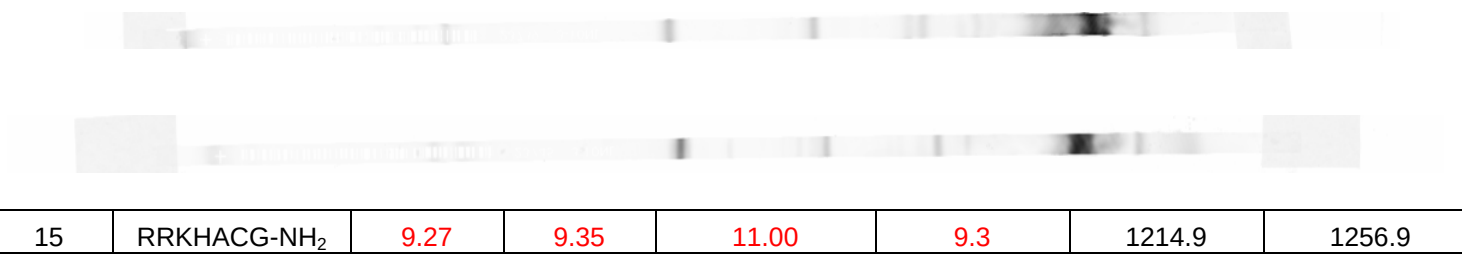
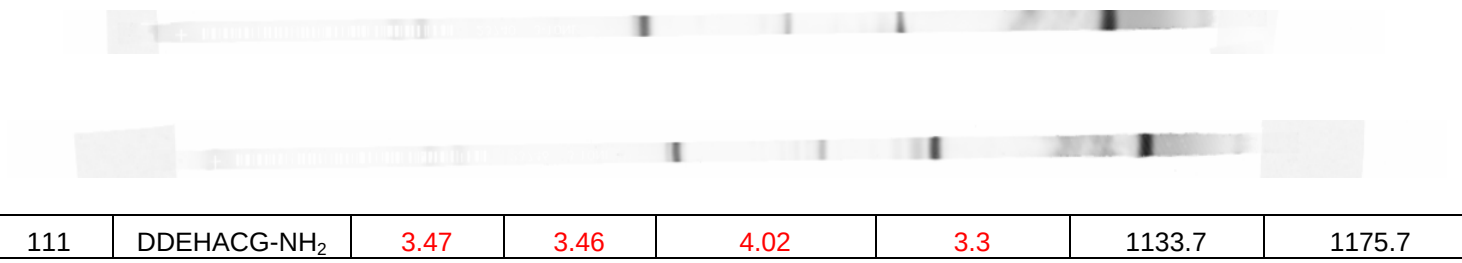
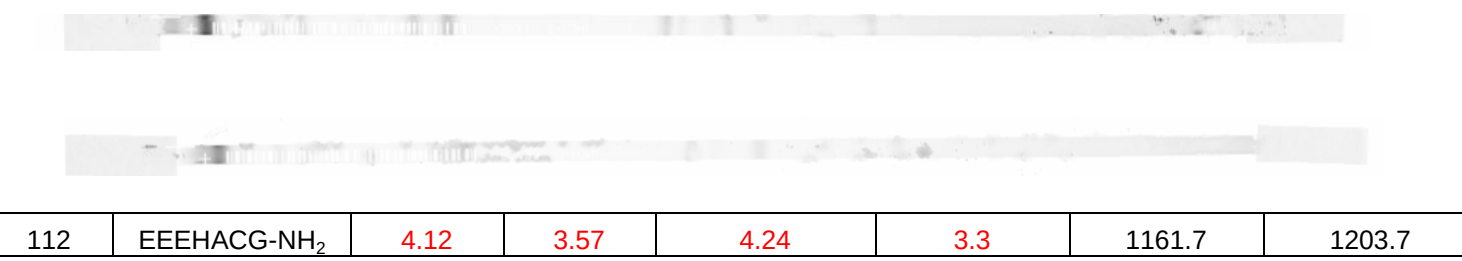
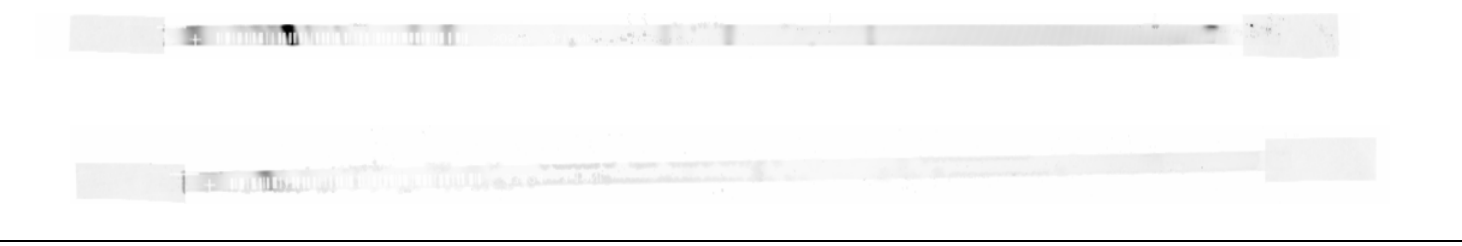
4.2.1 Results

The development of fluorescein labeled peptide markers required a certain degree of expertise in chemical protein synthesis. For this matter, a collaboration with Oscar Vadas from the Professor Keith Rose's group and with Professor Jacques Deshusses from the Department of Structural Biology and Bioinformatics at the University Medical Center in Geneva was envisaged.

A set of fluorescein labeled peptide markers expressing different isoelectric points was developed. The peptide synthesis was performed using Solid Phase Peptide Synthesis using a Fmoc/tBu technique [43].The fluorescein molecule was attached to cysteine residues. The oligopeptides shared three common residues at the C-terminus. Another series of peptides were acetylated at the N-terminus. Acetylation is a common post-translational modification in living cells [44]. The acetylation renders the peptide more hydrophobic and less basic, and adds 42 Da to the molecular weight. We therefore obtained a series of non-acetylated and a series of acetylated peptides, with slight differences in pIs. The markers had to be purified from contaminants by HPLC and Thin Layer Chromatography (TLC).

The list of fluorescein peptidic markers and their focusing on IPG strips is shown in figure 7):

marker	AA sequence	pI	Ac pI	Theoretical pI	Desired pI Including fluorophore	Mass (Da) Including fluorophore	Mass Ac (Da) Including fluorophore
11	EEHACG-NH_2	4.18	3.82	4.88	3.3	1032.6	1074.2
							
12	HHACG-NH_2	6.11	5.73/5.83	7.70	4.9	911.5	953.5
							

13	KKHACG-NH_2	8.83	8.32	9.43	7.2	1030.7	1072.7
							
14	RKHACG-NH2	8.83	8.51	9.42	7.2	1058.7	1100.7
							
15	RRKHACG-NH_2	9.27	9.35	11.00	9.3	1214.9	1256.9
							
111	DDEHACG-NH_2	3.47	3.46	4.02	3.3	1133.7	1175.7
							
112	EEEHACG-NH_2	4.12	3.57	4.24	3.3	1161.7	1203.7
							

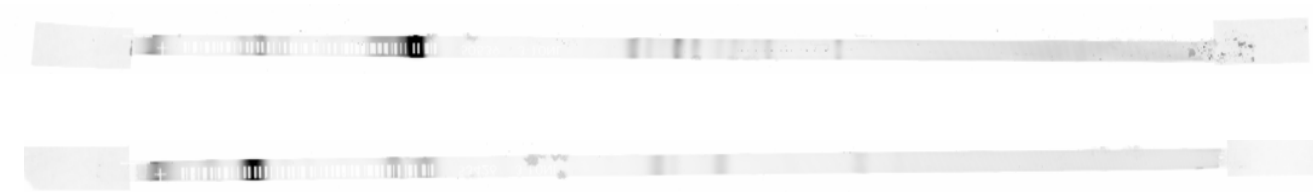
121	KEEHACG-NH_2	5.25	4.51	5.40	4.9	1160.8	1202.8
							
122	EHACG-NH_2	5.02	4.01	5.24	4.9	903.5	945.5
							
123	DHHACG-NH_2	5.73/5.82	4.92	5.97	4.9	1026.6	1068.6
							
131	DKHACG-NH_2	6.02/6.09	5.03	6.73	7.2	1017.6	1059.6
							
132	YDKKACG-NH_2	6.72	5.18	8.18	7.2	1171.9	1213.9
							

Figure 7. pI (of the peptides only) and molecular mass (including the fluorophore) of peptidic markers, acetylated and non-acetylated. The strip images with the non-acetylated marker are shown on top, the ones with the acetylated marker on the bottom. The theoretical pI was calculated from the peptide sequence only. The expected pI takes into account the variation due to the linking to the fluorophore. The peptide mass takes into account the mass of the fluorophore (388 Da).

Ideally, we were looking to obtain markers in different areas of the IPG strip. Four peptides were found having the desired isoelectric point: peptide markers 14, 15, 111, and 123 acetylated. Marker 15 had to be abandoned. In the focusing it showed a tendency to leave the 3-10 IPG strip at its basic end and enter the paper wick. Its pI fell outside the 3-10 pH range. Marker 14 was the second best choice for the basic region with a pI of 8.5, not excessively far from the desired pI of 9. Marker 111 and 123 acetylated were very close to the desired isoelectric point for the acidic region with pIs of 3.5 and 4.9. With these three markers, a good coverage of the whole pH gradient was achieved. From this point on, markers 111 (DDEHACG-NH_2), 123 acetylated (DHHACG-NH_2), and marker 14 (RKHACG-NH2) will be referred as to: marker 1, marker 2, and marker 3.

To test the reproducibility of our fluorescent peptidic markers two different samples were loaded onto four 13 cm 3-10 strips using marker 1, 2, and 3. The results are shown below:

Figure 8. Fluorescein pI markers 1, 2, and 3 focused on 3-10 L 13 cm strips in the presence of *S. aureus* S_{30} peptides H(-) first two from the top, and in the presence of *S. aureus* S_{30} peptides H(+) first two from the bottom.

This result showed that the markers had a constant isoelectric point.

We were interested in understanding if a change in the background medium would have an effect on the focusing of the markers. A test was designed to analyze the effect of BSA peptides on the quality of the marker focusing using 3-10 NL 18 cm IPG strips:

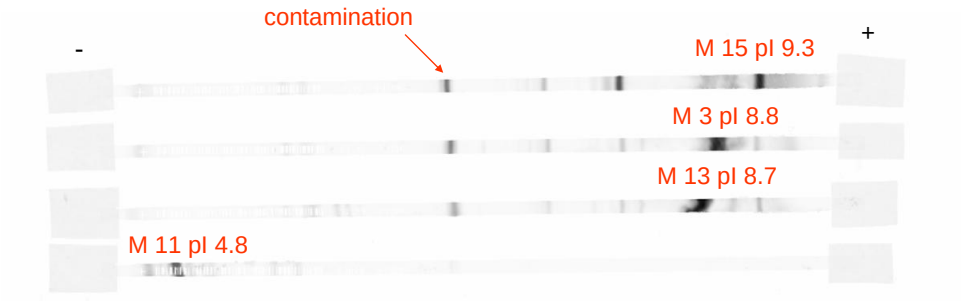


Figure 9. Fluorescein pI markers 11, 13, 3, 15 focused on 18cm 3-10 NL strips in the presence of BSA peptides. The other bands in the image represent contaminations.

In the presence of BSA peptides markers have very similar pIs, when compared with previous experiments using different media. These results are confirmed in an additional test using *E.Coli* and *S. aureus* S_{30} peptides:

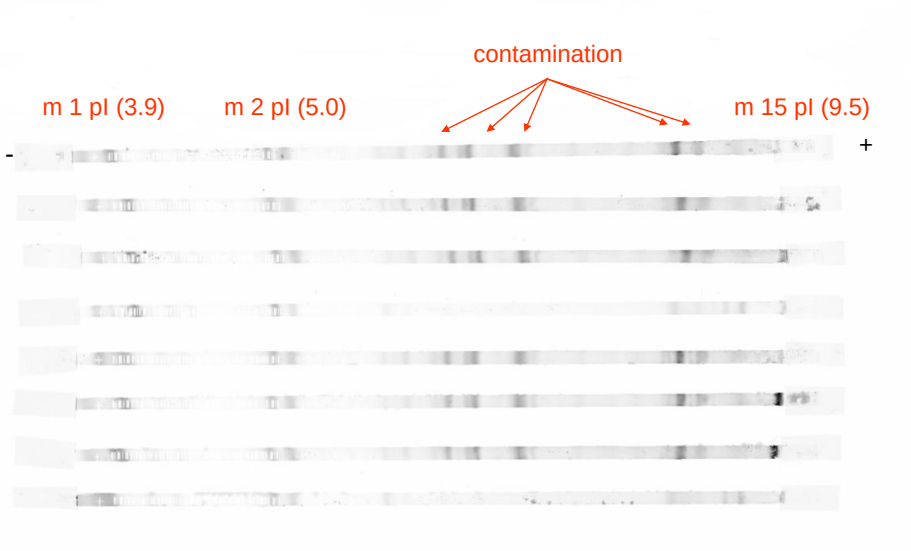


Figure 10. Fluorescein pI markers 1, 2, and 15 focused on 18cm 3-10 L strips in the presence of *S. aureus* S_{30} peptides H(-) (first four from the top), and *E.Coli* peptides (first four from the bottom). The other bands in the image represent contaminations.

Another test was set up to compare the effect of using *E.Coli* peptides on the quality of the marker focusing using once again 3-10 NL 18 cm IPG strips. Additionally, we were interested in analyzing the effect of background peptide quantity on the focusing of the markers. Different amounts of *E.Coli* peptides were used: 10, 100, and 200 µg respectively.

Figure 11. Fluorescein pI marker 15A focused on 3-10 NL 18 cm strips in the presence of 10, 100, 200 µg of *E.Coli* peptides (from the bottom top). The other bands in the image represent contaminations.

The concentration of the background peptides had an impact on the quality of the focusing. With 10 µg the marker band appeared blurred, and the marker did not focus at its expected isoelectric point. The sharpest focusing was obtained using 200 µg of *E.Coli* peptides. Not only the band appeared sharply focused, but the pI was consistent with previous results. Furthermore, the presence of satellite bands on IPG strips after focusing was observed.

In a similar experiment aimed at assessing the purity of the markers, SIEF was performed on a few *E.Coli* fractions containing contaminations. In parallel an imaging approach was performed to obtain MS spectra of the contaminants and the marker. The MS spectra confirmed the nature of marker 12 acetylated. The spectra of the two contaminants appeared to be a mixture of *E.Coli* peptides and unknown masses. No peaks belonging to marker 12 acetylated could be detected. However, due to the fact that we were not able to identify all of the most intense peaks, it remained unclear if the presence of satellite bands was due to the degradation of the markers or simply to other impurities.

Interestingly, during the marker focusing phase a few markers appeared in the IPG strip in the form of double bands. We decide to test the nature of this phenomenon to understand whether the two distinct bands substantiated the existence of two different states of the marker, or if they were the result of degradation of the marker itself during focusing. Marker 12 acetylated showed this peculiar characteristic. The gel fraction containing the double band was excised from the strip, and peptides of each band were extracted in two separate tubes. In a second run, the extracted peptides were refocused separately using the same 3-10 L 18 cm IPG strips.

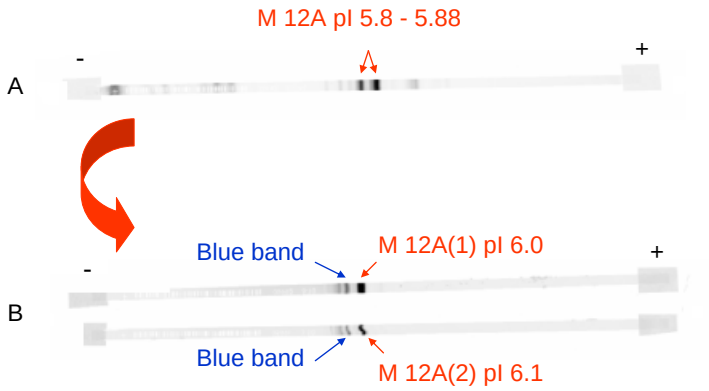


Figure 12. (A) Fluorescein pI marker 12A focused on 3-10 NL 18 cm IPG strips in the presence of 100 µg of *E.Coli* peptides. (B) extraction and refocusing of the two distinct bands. A double band was again observed on each strip. One of the bands had changed color from yellow to blue.

An interesting observation was the presence in each IPG strip of a single sharp band at a pI of 6.0-6.1 next to a blurred band in a slightly more acidic region. Surprisingly one of these two bands appeared to be blue. In order to further test the nature of this phenomenon, for both strips the intense band belonging to the marker and the presumed blue contamination were excised and passed on a MALDI-TOF instrument using the Imaging approach.

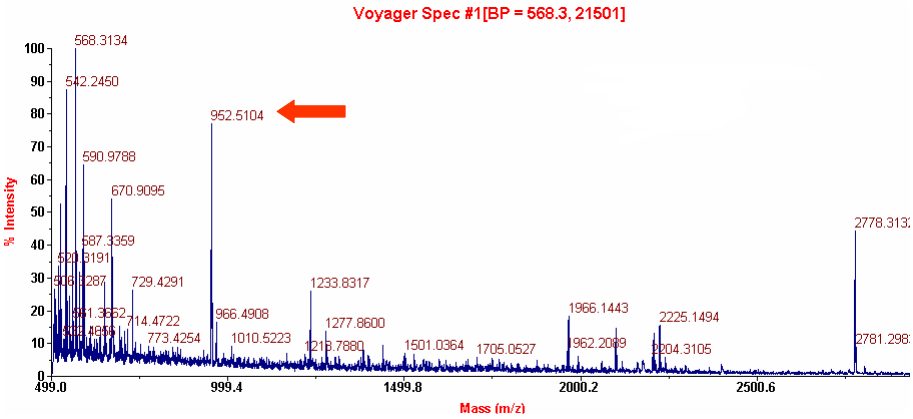


Figure 13. MS spectrum of the peptides present in the blue band. The arrow points to the peak belonging to fluorescent marker 12A with a mass of 952.5 Da.

The obtained spectra confirmed the identity of marker 12A in both strips. The blue band surprisingly contained marker 12A.

In order to find out the optimal amount of marker to be load for Imaging Shotgun IPG-IEF, different amounts of marker 15 were tested in the presence of *S. aureus* S_{30} H(+) peptides using the ISIEF pipeline. The MSight images are shown in figure 14:

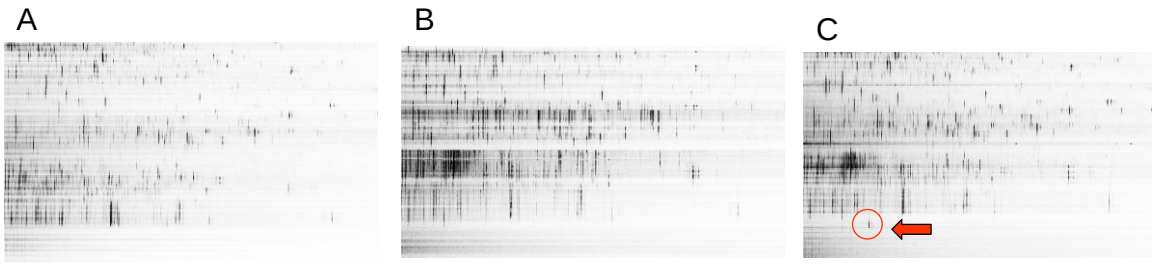


Figure 14. MSight images using different amounts of marker 15 in the presence of S.aureus S_{30} H(+) peptides on 7cm 3-10 L IPG strips: (A) absence of marker, (B) 4 µg and (C) 16 µg.

The presence of the marker was revealed only by using 16µg of marker 15; the sharper the focusing, the better the quality of the spot on the image. Unfortunately, when low marker amounts were used, the signal became indistinguishable from the background noise (figure 14.B).

4.2.2 Discussion

The distribution of the peptides along the pH gradient is not homogenous. Ideally, the fluorescent marker should be in these peptide poor regions of the strip, to be distinguishable from other peptide signals. Markers can be used to normalize irregularities in the strip gradient or differences in focusing. Moreover, they can be used as matching points for imaging superimposition in ISIEF (see section 2.1 and 2.2). One

of the challenges in designing such markers was the difficulty to obtain the desired isoelectric point. The theoretical pI can be calculated from the side chains of the composing amino acids using the pI prediction algorithm designed by Stephenson *et al.*. Nevertheless, the coupling of fluorescein (pK 6.4) to the peptides can considerably shift their pIs. The extent of this shift was unknown; a slight decrease in pK due to the mildly acidic character of the fluorophore was expected. Our results showed that the change in pI was in most cases inconsistent, leading to ambiguity in marker pI prediction from its amino acid sequence only.

The optimal marker quantity had to be found. If a small amount of marker is used its detection becomes problematic. On the contrary, if an excessive amount of marker is used it might suppress other signals. Furthermore, due to the fact that the markers were developed for matching purposes, it was extremely important that their pI was stable and that the band showed as sharp and as intense as possible. Inconsistency in markers focusing represents a serious problem in IPG-IEF. Not only the pI has to stay constant from one experiment to another, but it has to stay invariant when background peptides are substituted.

In order to assess the reproducibility of IPG-IEF experiments using our pI markers, it was important to test our markers in different media. For this purpose, peptides from BSA, *E.Coli* , and *S. aureus* S_{30} were chosen; small shifts in pI were observed when the background medium was altered, especially when BSA peptides were used. This is probably due to the ionic composition of the peptides themselves. The low number of peptides derived from BSA resulted in an incomplete coverage of the pH gradient and thus low focusing quality. Difference in strip length and gradient might also result in irreproducibility, which confirms the need for pI markers. Up to 1 mm in strip length differences were observed.

Another concern was related to the purity of the markers. Even after two consecutive rounds of purifications, contamination was still observed. The "double band" phenomenon is probably due to acid/base equilibrium around the neutral zone of the IPG strip. The fact that we observed the presence of the marker in the yellow band as well as in the blue band shows the instable nature of the coupling fluorophore-peptide. The marker may have changed its molecular configuration, as to appear at two distinct positions in the strip. Any improvement would have to be related with the chemistry of the molecules, by changes in amino acid sequences or by using a different fluorophore which has no charge. The synthesis would take time and considerable effort.

The quantity of marker used is a very important variable in an ISIEF experiment. The signal created by the marker has to be as sharp as possible. Ideally, the markers should appear as dots. In practice their signal is represented by a small vertical line, due to diffusion; the smaller the line, the more reliable the marker. Different markers showed dissimilar patterns in focusing. This became a problem when more than one marker was loaded onto the same IPG strip. We noticed that the focusing time differed from marker to marker, obliging us to be very attentive to the focusing process. Poor marker focusing inevitably translated into poor quality ISIEF images, rendering any subsequent comparison analysis problematic.

4.3. MS Imaging

4.3.1 Results

The development of fluorescent peptidic markers was intended for Imaging IPG-IEF. MSight images are the combination of two separate images. Due to the size of the MALDI target (4cm x 4cm), the 7 cm PVDF membranes had two be cut in two. This produced a gap in the image, because the matrix deposition at the extremity of the membrane is quite hard to achieve using the spotting robot. Therefore the peptides present in the central part of the image were lost, and the reconstruction of the whole image was hard to accomplish. For this matter, we decided cut the membrane diagonally. The two images would then contain a common pH region, rendering the matching easier, thus avoiding the loss of the peptides neighboring the excised area (figure 15.A).

In order to optimize the extraction of peptides, half of the membrane was covered with matrix and fixed onto the MALDI target for ISIEF, while the other half remained detached so that fractions can excised for extraction using SIEF (figure 15.B). Consequently, it was possible to perform SIEF and ISIEF in parallel. MSight images could be used for cutting out fractions directly from that same membrane, knowing that the pI gradient is exactly the same. This method would therefore increase the precision in the excising process enabling a more efficient peptide recovery.

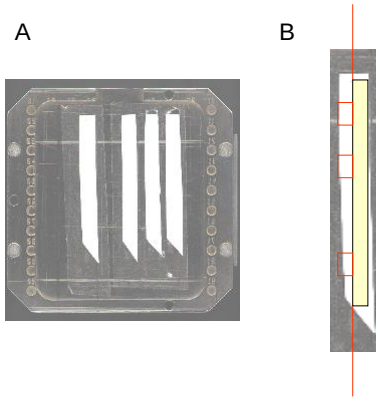


Figure 15. (A) Four PVDF membranes on a MALDI target using the diagonal cut approach. (B) Half of the membrane is coated in matrix and glued to the target while other half is not attached and ready to be fractionated.

Once all the variable were optimized, Imaging IPG-IEF was performed for the comparison of protein profiles in the presence of *S.aureus* S_{30} H(-) and *S.aureus* S_{30} H(+) peptides using markers 1, 2, and 3. This experiment was performed in two replicates to reduce the effect of technical irreproducibility on the differential comparison of the samples. The reconstructed MSight images are presented in figure 16:

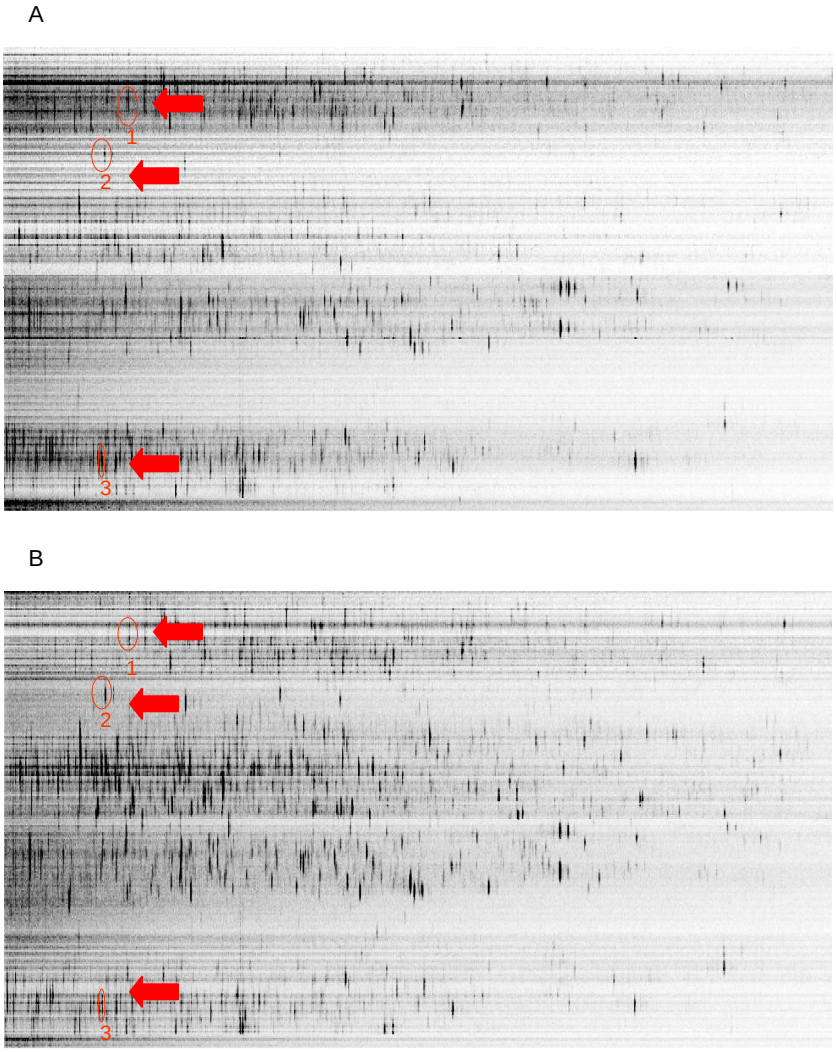


Figure 16. MSight reconstructed images of markers 1,2, and 3 in the presence of: (A) *S.aureus* S_{30} H(-) peptides and (B) *S.aureus* S_{30} H(+) peptides on 3-10 7cm L IPG strips, after two rounds of background noise removal and TIC normalization.

A comparison of the two images was carried out directly on MSight by zooming onto selected areas, and by using the "Differential Display" feature. This tool allows the direct superimposition of two images using chosen landmarks as reference points. The use of multiple landmarks resulted in an improvement in the quality of the alignment. For this reason, in addition to fluorescent markers, four more peptides were chosen as landmarks. A "Differential Display" image is shown in figure 14. Each image was reproduced in a different color to facilitate the recognition of differentially express peptides.

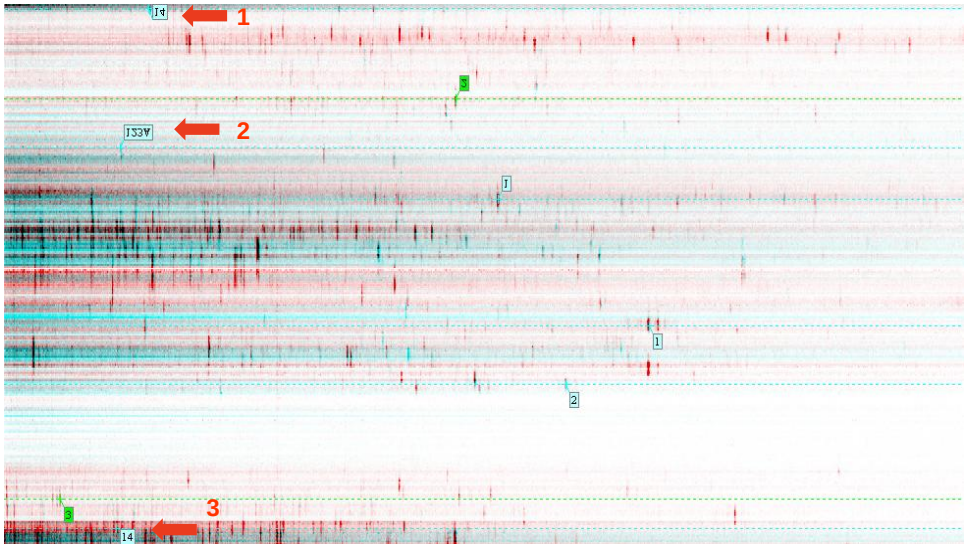


Figure 17. Differential Display image of *S.aureus* S_{30} H(-) peptides in blue and (B) *S.aureus* S_{30} H(+) peptides in red, using 3-10 L 7cm IPG strips with markers 1,2,3, and four other peptides as landmarks.

The "Differential Display" image revealed a few differences between the two samples. These differences were confirmed by comparing the two images on the screen and focusing the area of interest. Furthermore, the two replicates were internally compared to check the reproducibility of the results; the images appeared to be consistent with each other. Some peptides were indeed strongly visible as clean spots on one image and barely recognizable on the other image. These differences were detected also by cross-comparison with the replicate. Three peptides that satisfied both criteria were chosen for further analysis. The zoomed images and the spectra of peptide 2028 m/z are shown in figure 18:

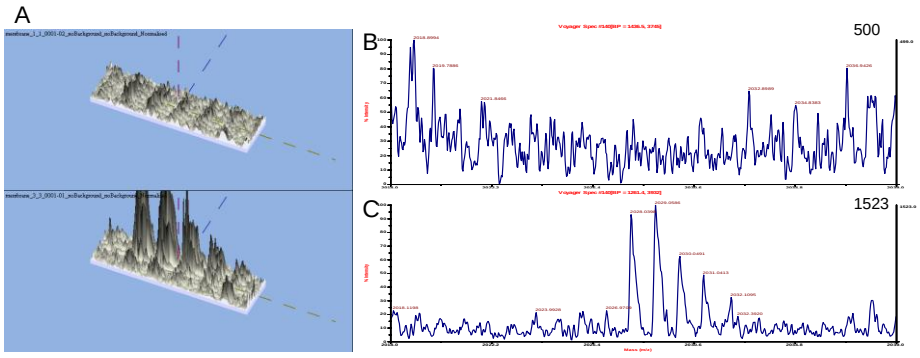


Figure 18. (A) MSight 3D zoomed image of peptide 2028 m/z in *S.aureus* S_{30} H(-) (on top), and in *S.aureus* S_{30} H(+) (on the bottom). (B) ISIEF MS spectrum of peptide 2028 m/z in *S.aureus* S_{30} H(-). (C) ISIEF MS spectrum of peptide 2028 m/z in *S.aureus* S_{30} H(+).

Peptide 2028 m/z showed a very intense signal on the 3D image of the MS concatenated spectra of *S.aureus* S_{30} H(+), yet not detectable in H(-) (figure 11.A). These results were confirmed by their respective MS spectra (figure 11.B and C).

A SIEF experiment was performed to confirm the results obtained with Imaging Shotgun IPG IEF. The position of peptides of interest 1636, 1677, and 2028 m/z was calculated using pI markers directly on the MSight image. The distance from the closest marker was calculated by multiplying the number of spectra by the distance between subsequent laser acquisitions (250 or 150 µm). This distance was reported on an IPG strip that was focused after data analysis on the membrane. The three fractions including the peptides of interest were excised directly from the strip, and peptides extracted in order to be analyzed by MALDI-MSMS.

The identification of peptides present in the fractions using Phenyx, confirmed the imaging results for peptide 1677 m/z (ALNHDFAEVFTGDIK), which was present in the *S. aureus* S_{30} H(+) fraction and absent in the H(-) fraction (figure 19). The peptide belonged to a Metal-dependent phosphohydrolase HD sub domain (A6TZN5). Unfortunately, we were not able to identify the other two peptides.

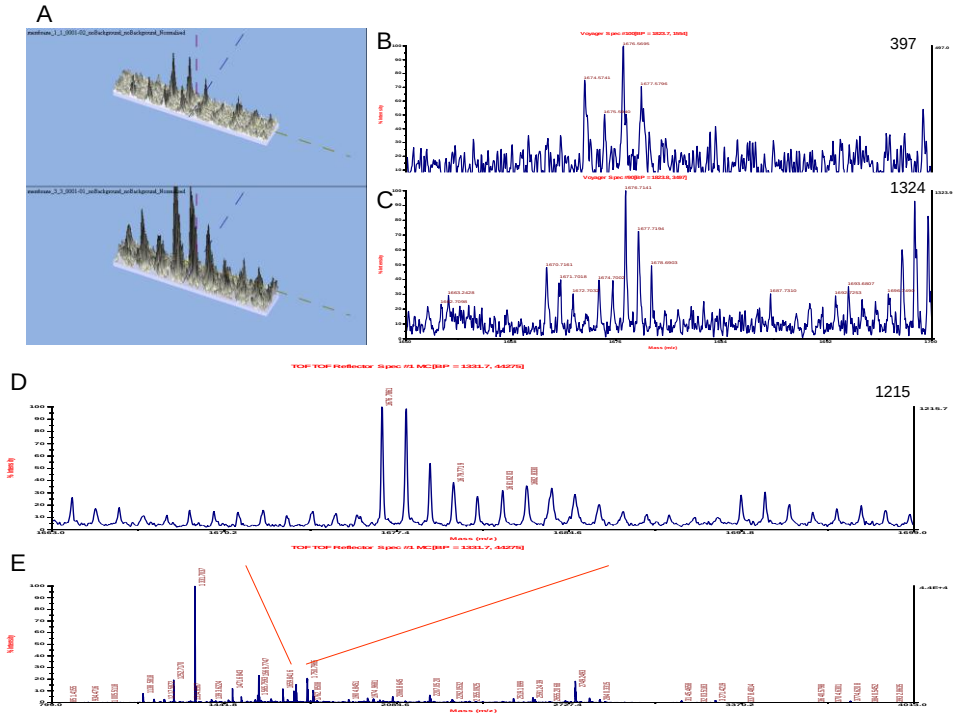


Figure 19. (A) MSight 3D zoomed image peptide 1677 m/z in *S.aureus* S_{30} H(-) (on top), and in *S.aureus* S_{30} H(+) (on the bottom). (B) ISIEF MS spectrum of peptide 1677 m/z in *S.aureus* S_{30} H(-). (C) ISIEF MS spectrum of peptide 1677 m/z in *S.aureus* S_{30} H(+). (D) SIEF MS spectrum zoomed on peptide 1677 m/z in *S.aureus* S_{30} H(+). (E) SIEF MS spectrum on the fraction including peptide 1677 m/z in *S.aureus* S_{30} H(+).

The results obtained with MSight for peptide 1677 m/z were confirmed using the Biomap software. Biomap provides a virtual representation of the distribution of peptides on the surface of the PVDF membrane (figure 20).

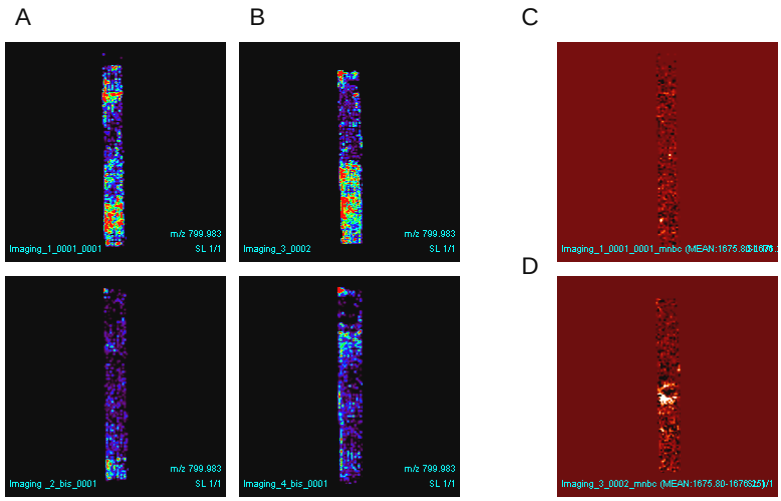


Figure 20. (A) Biomap image of the PVDF membrane using *S.aureus* S_{30} H(-) peptides (acidic side on the top image). Areas of high concentration of peptide are depicted in red. (B) Biomap image of the PVDF membrane using *S.aureus* S_{30} H(+) peptides (acidic side on the top image). (C) Biomap image of the distribution of peptide 1677 in *S.aureus* S_{30} H(-). (D) Biomap image of the distribution of peptide 1677 in *S.aureus* S_{30} H(+), white spots represents areas of high concentration of peptides.

4.3.2 Discussion

The diagonal cut approach was tested in order to avoid peptide signal loss due to the cutting of the membrane; up to 1.20 mm of peptide signal can be lost, and it becomes an issue when the membrane is cut in an area of high abundance of peptides. The new approach eliminated this problem, and allowed for a reliable reconstruction of the complete image. Unfortunately MSight is not yet adapted for the superimposition of only a portion of the image. The images had to be reconstructed by measuring distances between selected peptide and superimposing them in order to respect the relative distances. A great improvement would be the development of a tool on MSight that allows the partial superimposition of two images using once again landmarks as reference points. For image comparison the "Differential Display" future on MSight was used. The superimposition is not computed automatically; therefore the choice of good quality landmarks as reference point remains crucial. The images are stretched as to adjust to the reference points; the result is that some parts of the image are better aligned than others. An improvement would be the automatization of the alignment. Another issue is the choice of colors for the superimposition of the images. At the moment only two colors, light blue, and pink are available resulting in purple when superimposed. The colors are too similar, therefore differences are hard to spot. It would be beneficial to have a wider range of choice of colors, possibly by using two colors at the opposite side of the color wheel (complementary) in order to easily recognize matches at the first glance.

Recovering chosen peptides by means of extraction using an MSight image as a blueprint remains a major difficulty. Because peptides are not visible on the IPG strip (only the markers are), it is not possible to be exactly sure that the excised fractions will contain selected peptides in SIEF. In a new approach, the MSight image is used as a guide for the

positioning of peptides. The problem by running a SIEF and ISIEF pipeline in parallel is that the position of the peptides will not be the same, because different IPG strips are used. The matrix coating of one half of the membrane permitted to run the two pipelines in one single experiment. The position of the peptide for the extraction can be directly calculated from the image. Since half of the membrane is not fixed to the target, the excision is a simple process. The major drawback of this technique remains the lower rate of recovery of peptides from the PVDF membrane. Even though the methodology appeared to promising, we decided to abandon the project.

Unfortunately in our last SIEF experiment we were not able to recover two peptides out of three (1637 and 2028 m/z); only peptide 1677 m/z was rescued. A plausible reason might be that the two peptides were not included in the excised fraction. A human error regarding the positioning of the excision might have occurred. Therefore, we calculated the pIs of the two peptides and the average pI of peptides present in each fraction as well. The pIs of the two peptides were not too far away to the average pI of each fraction, confirming the fact that they should have indeed appeared in the fractions. Their absence might also be because of a loss during extraction and purification steps or random sampling in the LC-MS/MS step. Therefore, the best scenario for identifying these peptides would be direct tandem mass analysis on the membrane to identify the peptides. However, a MALDI instrument with an ion trap or a quadrupole mass analyzer is needed to be able to avoid the charging effects inherent to membrane acquisitions.

The quality of the MSight image is crucial for comparison experiments. Excessive background noise and insufficient amounts of background peptide may lower the quality of the image. The deposition of the matrix on the PVDF membrane remains a problem. A poor matrix deposition results in excessive background noise. Our results show that marker 1, 2, and 3 are clearly visible on the image; marker 2 is the only markers to be positioned in an area of low peptide content.

An important aspect in the development of Imaging Shotgun IPG-IEF was the choice of the biological sample. The purpose of our work was to show how this technology could bring up difference at the protein level between two samples even when the protein profiles show no difference in one or two-dimensional gels. We performed SDS-PAGE and 2D-PAGE for *S.aureus* S_{30} samples, and no difference was observed. Our intention was to find a few differences in peptides that we could have traced back to proteins. According to the transcriptomic data the samples in question were very similar. It seemed plausible that in our analysis with *S.aureus* S_{30} H(-) and *S.aureus* S_{30} H(+) peptides, the major difference in the images would be caused by helicase peptides, though other differences might have been caused by the absence of the helicase protein in biochemical pathways leading to up-regulation or down-regulation of downstream proteins. Surprisingly, the differences observed were not due to helicase, yet to other proteins. Furthermore, it was difficult to observe a lot of differences in the images. Even though we knew the differences at the genomics it was not straight forward to find these differences translated at the proteomic level and at a concentration detectable with this technology.

5. Conclusions and outlook

In this report, we described various developments on Shotgun IPG-IEF pipeline. The development of the fluorescent peptidic markers was challenging. Even after multiple rounds of purification, and refocusing, impurity patterns were detected. It is hard a priori to make a judgment on the nature of these contaminants. In the result section we showed the presence of fluorescent markers in satellite bands. Probably this was due to the fact that the markers degraded or interacted forming covalent structures with other molecules, shifting their original isoelectric point. Although contamination does not represent a major challenge in the focusing itself, it would be beneficial to design pI markers that are less subject to chemical alterations.

The three pI markers chosen for further analysis focused at a pH that was close to the desired one. Unfortunately, the rest of the markers fell far off the desired pI values. The coupling of the fluorophore to the peptide changes its isoelectric point in an inconsistent manner. Different pI shifts were observed for different markers. In order to obtain pI markers in low-populated peptides regions, we were obliged to develop a whole series with the hope that after the pI shift due to the fluorophore, the markers would be present in the region of interest Out of the three chosen peptides, only one fell exactly within the desired pI region, marked by the absence of peptides. Furthermore, it was very hard to obtain pI markers on the extremity of the IPG strip, especially on the basic side; we were obliged to choose marker 3, even though it focused in a region of high peptide density. A possible improvement would be the designing of at least one or two more pI markers that show the exact desired pI values, for a more efficient alignment of MSight images.

One major problem remains the inability to obtain a complete image of the membrane due to the size of the MALDI target. A different approach was therefore tested. The membrane was cut diagonally to avoid peptide loss around the excised area. Although this technique clearly represents an improvement, a problem persists. The matching of the two common areas in the two images has to be done by calculating distances between chosen peptides and make sure that the distances are respected when superimposing manually. MSight does not allow for an automated superimposition of portions of images. A great improvement would be the development of a partial superimposition feature. The only other way possible to avoid peptide loss would be the use of a bigger MALDI target, for the attachment of the whole membrane. Unfortunately, the Voyager-STR MALDI-TOF in our laboratory does not accept bigger targets. The acquisition must therefore be accomplished using a different instrument.

Another difficulty was the deposition of matrix on the membrane. To avoid diffusion an apposite spotting robot has been developed “in house”. Despite the fact that the spotting itself is automated, the configuration remains manual. The user has to define the membranes and the distance between the membrane and the spotter. If the spotter remains too far away from the membrane no significant deposition will be accomplished. On the other hand, if the spotter touches the membrane and scratches it, the surface becomes uneven rendering the subsequent acquisition problematic. Furthermore, scratches and pressure marks will be perceptible on the image. This problem is exacerbated by the fact that the spotting robot has to be configured each time it finishes the spotting of the half of

the membrane. Therefore, matrix deposition discrepancies exist even at the internal level, from the acidic to the basic half of the membrane. It would be extremely beneficial to improve the configuration capabilities of the spotting robot as to have to possibility to effectuate the configuration only once per target. The resulting images would therefore show the same matrix deposition pattern. Use of "none-contact" matrix deposition devices should further increase the reproducibility.

The images were created using MSight, which offers a wide range of analytical tools for the imaging analysis. Most of the features on MSight are user friendly and do not require a specific training.

To match the images, reference points have to be chosen for a proper alignment. Usually it is considered good practice to choose at least four reference points. In the resulting "differential display" image the combined images are represented in one set of fixed colors. It would be beneficial to have the possibility to choose from a wider selection of colors so that peptides differentially expressed can be easily spotted by the human eye. Color preferences may vary among individuals. Moreover, differences seem to be easier to spot when complementary colors are used. An automated image matching, similar to 2D gels image analysis, would allow higher throughput and more efficient comparisons to be performed.

Our main goal was to find difference in images using two similar biological samples. The selection of the samples was not trivial. Samples had to be similar in protein profile, yet show a few relevant differences.

We expected to find out differences in helicase peptides using the imaging approach. Surprisingly, we were not able to trace any difference back to the helicase protein. We believe that the absence of the helicase gene in the mutant, fostered or inhibited a series of biological cascades. The differences in the profile of peptides were probably due to this phenomenon. Further tests should be performed to find the best method to trace the differences in the images and identify their peptide sequences. Direct MS/MS on the membrane is a shortcut to such an identification approach. However, no stable method has been proposed yet, to overcome to charging effects on the PVDF membrane. We strongly believe that a hybrid MALDI instrument with an ion trap or a quadrupole is essential to allow direct tandem analysis of the membranes.

Despite some of the limitations described above, Imaging Shotgun IPG-IEF is a high-throughput proteomics pipeline, which allows rapid analysis of any proteome in a single day. Further investigation is needed to improve some of the deficiencies encountered during this work. However, taking into account the gain in time and energy one can obtain using this approach; any further enhancement on the pipeline would definitely worth the investement.

6. References

1. Wilkins K.L., Appel R.D, Hochstrasser D.F., *Proteome research: New frontiers in functional genomics.* 1997, Berlin:Springer

2. Mullis K., Faloona F., Sharf S., Saiki R., Horn G., Erlich H., *Specific enzymatic amplification of DNA in vitro: the polymerase chain reaction.* Biotechnology. 1992;24:17-27.

3. Eriksson J.,Fenyo D., *Improving the success rate of proteome analysis by modeling protein abundance distributions and experimental designs.* Nature Biotechnology 25, 651 - 655 (2007)

4. Spahr C.S., *Towards defining the urinary proteome using liquid chromatography-tandem mass spectrometry. I. Profiling an unfractionnated tryptic digest.* Proteomics, 2001; 1(1): 93-107.

5. Link A.J., *Direct analysis of protein complexes using mass spectrometry.* Nat. Biotechnology, 1999; 17(7): 676-682.

6. Aizawa P., Winge S., Karlsson G., *Large-scale preparation of human plasma.* Thromb. Res., mar 2008, 7.

7. Lacey J.M.,*Rapid determination of transferring isoforms by immunoaffinity liquid chromatography and electrospray mass spectrometry.* Clin. Chem, 2001; 47 (3): 513-518.

8. Chaussee M.A., McDowell E.J., Chaussee M.S.., *Proteomic Analysis of Proteins Secreted by Streptococcus pyogenes.* Methods Mol. Biol., 2007; 431:15-24.

9. Fuchs R., Ellinger I., *Free-flow electrophoretic analysis of endosome subpopulations of rat hepatocytes.* Curr. Protoc. Cell Biol., May 2002, Ch. 3:Unit 3.11.

10. Hernández-Zamora E., de la Luz Arenas-Sordo M., Maldonado-Rodríguez R., *Capillary electrophoresis for the detection of PMP22 gene duplication: Study in Mexican patients.* Electrophoresis, Mar 2008, 3.

11. Patel N., Solanki E., Picciani R., Cavett V., Caldwell-Busby J.A., Bhattacharya SK., *Strategies to recover proteins from ocular tissues for proteomics.* Proteomics, Mar 2008, 7; 8(5):1055-1070.

12. Smith L., Welham K.J., Watson M.B., Drew P.J., Lind M.J., Cawkwell L., *The proteomic analysis of cisplatin resistance in breast cancer cells.* Oncol Res. 2007; 16(11): 497-506

13. Sui J., Zhang J., Tan T.L., Ching C.B., Chen W.N., *Two-dimensional electrophoresis of proteins: Comparative proteomic analysis of vascular smooth muscle cells incubated with S- and R-enantiomers of atenolol using iTRAQ-coupled 2D LC-MS/MS.* Mol. Cell. Proteomics, Feb 2008, 11.

14. Chamrad D.C., Körting G., Stühler K., Meyer H.E., Klose J., Blüggel M., *Evaluation of algorithms for protein identification from sequence databases using mass spectrometry data.* Proteomics, Mar 2004, 4(3):619-28.

15. Klose J., Kobalz U., *Two-dimensional electrophoresis of proteins: An updated protocol and implications for a functional analysis of a genome.* Electrophoresis, 1995; 16:1034-1059.

16. Gygi S.P., Corthals G.L., Zhang Y., Rochon Y., Aebersold R., *Evaluation of two-dimensional gel electrophoresis-based proteome analysis technology*. Proc Natl Acad Sci, U.S.A., 2000 Aug 15; 97(17):9390-5.

17. Giddins, J.C.J. High Resolut. Chromatogr. Commun. 1987, 10, 319.

18. Wolters D.A., Washburn M.P., Yates J.R. 3rd. *An automated multidimensional protein identification technology for shotgun proteomics*. Anal. Chem. 73, 5683-90.

19. Searle B.C., Turner M., Nesvizhskii A.I., *Improving sensitivity by probabilistically combining results from multiple MS/MS search methodologies*. J. Proteome Res. 2008 Jan; 7(1):245-53.

20. Chen H.S., Rejtar T., Andreev V., Moskovets E., Karger B.L., *Enhanced Characterization of Complex Proteomic Samples Using LC-MALDI MS/MS: Exclusion of Redundant Peptides from MS/MS Analysis in Replicate Runs*. Anal. Chem., 2005, 77 (23), 7816 -7825.

21. Link A. J., Eng J., Schieltz D. M., Carmack E., Mize G. J., Morris D.R., Garvik B. M. Yates J. R. 3rd, *Direct analysis of protein complexes using mass spectrometry.* Nat. Biotechnol., 1999, 17, 676.

22. Washburn M. P.; Wolters D.; Yates J. R., *Large-scale analysis of the yeast proteome by multidimensional protein identification technology*. 3rd. Nat. Biotechnol., 2001, 19, 242.

23. Wolters, D. A.; Washburn, M. P.; Yates, J. R. 3rd., *An automated multidimensional protein identification technology for shotgun proteomics.* Anal. Chem., 2001, 73, 5683.

24. Washburn, M. P.; Ulaszek, R.; Deciu, C.; Schieltz, D. M.; Yates, J.R., 3rd., *Analysis of quantitative proteomic data generated via multidimensional protein identification technology.* Anal. Chem. 2002, 74, 1650.

25. Washburn, M. P.; Koller, A.; Oshiro, G.; Ulaszek, R. R.; Plouffe, D.; Deciu, C.; Winzeler, E.; Yates, J. R., 3rd., *Protein pathway and complex clustering of correlated mRNA and protein expression analyses in* Saccharomyces cerevisiae. Proc. Natl. Acad. Sci., U.S.A. 2003, 100, 3107.

26. Hattan S. J., Marchese J., Khainovski N., Martin S., Juhasz P., *Comparative study of [Three] LC-MALDI workflows for the analysis of complex proteomic samples*. J. Proteome Res., 2005 Nov-Dec,4(6):1931-41.

27. Yamada Y., *Isoelectric focusing with reduced cathodic drift and migration into the anode chamber*. J. Biochem. Biophys. Methods., Nov 1983;8(3):175-81.

28. Bjellqvist B., Ek K., Righetti P.G., Gianazza E., Görg A., Westermeier R., Postel W., *Isoelectric focusing in immobilized pH gradients: Principle, methodology, and some applications.* J. Biochem. Biophys. Methods, 1982, 6: 317–339.

29. Aebersold R., Mann M., *Mass spectrometry-based proteomics*. Nature. Mar 2003, 13;422(6928):198-207. Review.

30. Cargile B., Sevinsky J.R., Essader A.S., Stephenson S.L. Jr., Bundy J.L., *Immobilized pH Gradient Isoelectric Focusing as a First-Dimension Separation in Shotgun Proteomics.* Journal of Biomolecular Techniques, 2005, 16:181-189.

31. Cargile B., Talley D.L., Stephenson J.L. Jr., *Immobilized pH gradients as a first dimension in shotgun proteomics and analysis of the accuracy of pI predictability of peptides.* Electrophoresis, 2004, *25*, 936–945.

32. Essader A., Cargile B., Bundy J.L., Stephenson J.L. Jr., *Immobilized pH Gradient Isoelectric Focusing as a First-Dimension Separation in Shotgun Proteomics.* Proteomics 2005,5,24-34.

33. Caprioli R.M., Farmer T.B., Gile J., *Molecular imaging of biological samples: localization of peptides and proteins using MALDI-TOF MS, Anal. Biochem,* 69, 1997, 4751–4760.

34. Rohner T.C., Staab D., Stoeckli M., *MALDI mass spectrometric imaging of biological tissue sections*. Mechanisms of Ageing and Development 126 (2005) 177–185.

35. Bhattacharya S.H., Gal A.A., Murray K.K. *Laser capture microdissection MALDI for direct analysis of archival tissue*. J. Proteome Res. 2003 Jan-Feb; 2(1):95-8 .

36. Hochstrasser D.F., Appel R.D., Vargas R., Perrier R., Vurlod J.F., Ravier F., Pasquali C., Funk M., Pellegrini C., Muller A.F., *A clinical molecular scanner: the Melanie project*. MD Comput., 1991, Mar-Apr; 8(2):85-91.

37. Binz P.A., Müller M., Hoogland C., Zimmermann C., Pasquarello C., Corthals G., Sanchez J.C., Hochstrasser D.F., Appel R.D. *The molecular scanner: concept and developments*. Current Opinion in Biotechnology, 2004 Feb; 15(1):17-23

38. Palagi P.M., Walther D., Quadroni M., Catherinet S., Burgess J., Zimmermann-Ivol C.G., Sanchez J.C., Binz P.A., Hochstrasser D.F., Appel R.D., *MSight: An image analysis software for liquid chromatography-mass spectrometry.* Proteomics 2005, 5, 2381–2384.

39. Spengler B., Hubert M., *Scanning microprobe matrix-assisted laser desorption ionization (SMALDI) mass spectrometry: instrumentation for sub-micrometer resolved LDI and MALDI surface analysis*. J. Am. Soc. Mass Spectrom., Jun 2002, 13(6):735-48.

40. Hankin J.A., Barkley R.M., Murphy R.C., *Sublimation as a method of matrix application for mass spectrometric imaging*. J Am Soc Mass Spectrom., Sep 2007, 18(9):1646-52.

41. Flatmark T., Vesterberg O., *On the heterogeneity of beef heart cytochrome c. IV. Isoelectric fractionation by electrolysis in a natural pH-gradient.* Acta. Chem. Scand. 1966, 20(6), 1497-1503.

42. Righetti P.G., Gianazza E., *Method for detecting charged oligonucleotides in biological fluids*. J. Chromatogr. 1977, 137, 171-181.

43. Nakhleh E.T., Samra S.A., Awdeh Z.L., *Isoelectric focusing of phenanthroline iron complexes and their possible use as pH markers.* Anal.Biochem., 1972, 49,218-224.

44. Stastna M., Travnicek M., Slais K., *New azo dyes as colored isoelectric point markers for isoelectric focusing in acidic pH region*. Electrophoresis, 2005, 26,53-59.

45. Shimura K., Wang Z., Matsumoto H., Kasai K., *Synthetic oligopeptides as isoelectric point markers for capillary isoelectric focusing with ultraviolet absorption detection.* 2000 Feb, 21(3):603-10.

46. Shimura K., Kamiya K, Matsumoto H., Kasai K., *Fluorescence-labeled peptide pI markers for capillary isoelectric focusing*. Anal Chem. 2002 Mar 1; 74(5):1046-53.

47. Bridger P.S., Haupt S., Klisch K., Leiser R., Tinneberg H.R., Pfarrer C., *Validation of primary epitheloid cell cultures isolated from bovine placental caruncles and cotyledons.* Theriogenology. 2007 Sep 1; 68(4):592-603.

48. Zhou L., Li C.J., Wang Y., Xia W., Yao B., Jin J.Y., Gui J.F., *Identification and characterization of a MBP isoform specific to hypothalamus in orange-spotted grouper (Epinephelus coioides).* J Chem. Neuroanat. 2007 Sep; 34(1-2):47-59.

49. Smith S.A. and Pretorius W.A., The *conservative behaviour of fluorescein*, ISSN 0378-4738 = Water SA, October 2002, Vol. 28 No. 4.

50. Scherl A., *Non redundant mass spectrometry: a strategy to integrate mass spectrometry acquisition and analysis*. Proteomics, 2004, 4(4): 917-927.

51. Mc Devitt, D., *Molecular characterization of the clumping factor (fibrinogen receptor) of Staphylococcus aureus.* Mol. Microbiol., 1994, 11(2): 237-248.

52. Allard L., *Apo-C-I and Apo C-III as potential plasmatic markers to distinguish between ischemic and hemorrhagic stroke.* Proteomics, 2004, 4(8): 2242-2251.

53. Kuroda M., *Whole genome sequencing of meticillin-resistant Staphylococcus aureus*. Lancet, 2001, 357(9264): 1225-1240.