PROTEOMIC ANALYSES OF BIOLOGICAL SAMPLES BY USING DIFFERENT MASS SPECTROMETRIC STRATEGIES

A Thesis Submitted to
the Graduate School of Engineering and Science of
İzmir Institute of Technology
in Partial Fulfillment of the Requirements for the Degree of

DOCTOR OF PHILOSOPHY

in Chemistry

by
Melike DİNÇ

September 2018
İZMİR

We approve the thesis of **Melike DİNÇ**

Examining Committee Members:

Prof. Dr. Ahmet KOÇ
Department of Medical Biology and Genetics, İnönü University

Prof. Dr. Ahmet ÖZBİLGİN
Department of Parasitology, Celal Bayar University

Prof. Dr. Talat YALÇIN
Department of Chemistry, İzmir Institute of Technology

Dr. Nur Başak SÜRMELİ
Department of Bioengineering, İzmir Institute of Technology

Dr. Ümit Hakan YILDIZ
Department of Chemistry, İzmir Institute of Technology

3 September 2018

Prof. Dr. Talat YALÇIN
Supervisor, Department of Chemistry
İzmir Institute of Technology

Prof. Dr. Ahmet EROĞLU
Head of the Department of Chemistry

Prof. Dr. Aysun SOFUOĞLU
Dean of the Graduate School of
Engineering and Sciences

ACKNOWLEDGEMENT

My Ph.D. journey has been a period of intense learning, not only in the scientific arena, but also on a personal level. I would like to take this opportunity to express my sincere appreciation and convey my thanks to the people who have always been there for me and contributed in different ways. First and foremost, I would like to thank my supervisor Prof. Dr.Talat YALÇIN for his kindly support, excellent guidance, continuous encouragement and endless patience. I could never find enough words to express the influence he has had to my life. I have always felt blessed and honored to be part of such a wonderful team. I would like to offer my sincere gratitude to Prof. Dr. Ahmet ÖZBİLGİN for his cooperation, support and understanding. I am also very grateful to İbrahim ÇAVUŞ for his kind assistance and patience throughout our collaboration. I wish to thank to committee members Prof. Dr.Ahmet KOÇ, Assist.Prof. Dr.Nur Başak SÜRMELİ and Assist.Prof. Dr.Ümit Hakan YILDIZ for their valuable contributions. I would like to thank Prof. Dr.Günter ALLMAIER and Assoc.Prof. Dr.Martina MARCHETTI-DESCHMANN for inviting me to their lab and giving me the chance to experience the lovely Vienna. Also, I would like to thank to Dr.Güneş BARKA for his helps and kindness. My special thanks to my friends and former members of our laboratory Dr.Melda Zeynep GÜRAY, Dr.Ahmet Emin ATİK, Dr.Çağdaş TAŞOĞLU and Dr.Filiz YEŞİLIRMAK. I am so glad I met you. We may be apart in distance, but not in heart I believe. I would like to acknowledgement the grant from The Scientific and Technological Research Council of Turkey (TUBITAK) to the project 215Z257 and express my gratitude for the financial support. Lastly, I would like to add my deepest appreciation to my parents and sister for their unconditional love, faith and unlimited support; you are the best.

ABSTRACT

PROTEOMIC ANALYSES OF BIOLOGICAL SAMPLES BY USING DIFFERENT MASS SPECTROMETRIC STRATEGIES

The advent of soft ionization techniques MALDI and ESI have opened the doors of mass spectrometry to biological samples. Particularly protein research exploited this opportunity to the utmost that qualitative, quantitative, structural and distributional analyses of proteins have become achievable. Comparison of all expressed proteins belonging to two or more states is defined as differential proteomics and commonly performed either separating the proteins on a gel or separating the peptides in the mixture within a column. In this study, both of these methods were applied to a parasite that causes to the infectious disease, Leishmaniasis. Consequences of the disease become severe when parasite is multiplied in viscera rather than skin. Occurrence of different clinical manifestations is attributed to the variety in species however some species can cause to both such as *L.infantum* and *L.tropica*. The purpose of this study was to investigate this duality in *L.infantum*. Despite originating from two different species, cutaneous leishmaniasis isolates obtained from ten different patients exhibited considerably similar protein profiling on the gels however isolates belonging to five visceral leishmaniasis patients were discriminated from them with regards to protein abundances. Additionally, several differential proteins between cutaneous leishmaniasis and visceral leishmaniasis samples were determined by gel-free approach.

Apart from that, another mass spectrometric strategy for the abovementioned distributional analysis namely mass spectrometry imaging was included in the last section. Herein, a method incorporating the application of enzyme with matrix was developed for easy sample preparation to on-tissue digestion and promising results were obtained for a start.

ÖZET

BİYOLOJİK ÖRNEKLERİN FARKLI KÜTLE SPEKTROMETRİK STRATEJİLER KULLANARAK PROTEOMİK ANALİZLERİ

Yumuşak iyonlaşma teknikleri olan MALDI ve elektrospreyin gelişi, kütle spektrometri yönteminin kapılarını biyolojik örneklere açmış oldu. Özellikle protein araştırmaları bu fırsattan alabildiğine yararlandı o kadar ki proteinlerin kalitatif, kantitatif, yapısal ve dağılımsal analizlerleri yapılabilir hale geldi. İki ya da daha fazla duruma ait ifadelenen tüm proteinlerin karşılaştırılması ayrımsal proteomik olarak tanımlanır ve yaygın olarak ya proteinleri jel üzerinde ayrıştırarak ya da kompleks karışımdaki peptitleri kolonda ayrıştırarak yapılır. Bu çalışmada, bu metotların her ikisi de bulaşıcı hastalık leyişmanyöze sebep olan bir parasite uygulanmıştır. Parazit, deri yerine iç organlarda çoğaldığında hastalığın neticesi daha da şiddetli hale gelir. Farklı klinik belirtilerin oluşması türlerdeki çeşitliliğe bağlanır ancak *Leishmania infantum* ve *Leishmania tropica* gibi bazı türler her ikisine de sebep olabilir. Bu çalışmanın amacı, *L.infantum* türünde bu ikililiği protein düzeyinde araştırmaktı. İki farklı türden elde edilmelerine rağmen on farklı hastadan alınan deri leyişmanyöz numuneleri jel üzerinde oldukça benzer protein profilleri ortaya koydu ancak beş içorgansal leyişmanyöz hastalarına ait olan numuneler protein miktarları bakımından onlardan ayrıldılar. Bununla birlikte, jelsiz yaklaşımla deri leyişmanyözü ve iç organsal leyişmanyöz örnekleri arasında birçok ayrımsal protein belirlenmiştir.

Bundan başka, yukarıda bahsi geçen dağılımsal analiz için kütle spektrometri görüntüleme olarak adlandırılan bir diğer kütle spektrometrik stratejiye son bölümde yer verilmiştir. Burada, dokuda parçalama yöntemine kolayca örnek hazırlamak için enzim uygulamasını matriks ile birleştiren bir metot geliştirilmiş ve başlangıç için umut verici sonuçlar elde edilmiştir.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

LIST OF ABBREVIATIONS

MS	Mass spectrometry
2D-GE	Two-dimensional polyacrylamide gel electrophoresis
ABC	Ammonium bicarbonate
APCI	Atmospheric pressure chemical ionization
BSA	Bovine serum albumin
CBB-G-250	Coomassie brilliant blue-G-250
CE	Capillary electrophoresis
CHAPS	3-[(3-cholamidopropyl)dimethylammonio]-1-propanesulfonate detergent
CHCA	Cyano-4-hydroxycinnamic acid
CI	Chemical ionization
CID	Collision induced dissociation
CL	Cutaneous leishmaniasis
DDA	Data-dependent acquisition
DHB	2,5-Dihydroxybenzoic acid
DNA	Deoxyribonucleic acid
DTT	Dithiothreitol
EI	Electron ionization
ELISA	Enzyme-linked immunosorbent assay
emPAI	Exponentially modified protein abundance index
ESI	Electrospray ionization
FD	Field desorption
FDR	False discovery rate
FI	Field ionization
FTICR	Fourier transform ion cyclotron resonance
GS-MS	Gas chromatography-mass spectrometry
HCCA	Cyano-4-hydroxycinnamic acid
HPLC	High performance liquid chromatography
IAA	2-Iodoacetamide
ICP	Inductively coupled plasma
ICR	Ion cyclotron resonance
ICT	Immunochromatographic test

IEF	Isoelectric focusing
IPG	Immobilized pH gradient
ITO	Indium tin oxide
LC	Liguid chromatography
LC-MS	Liguid chromatography-mass spectrometry
LTOF	Linear time-of-flight
m/z	Mass-to-charge
MALDI	Matrix-assisted laser desorption/ionization
MCL	Muco-cutaneous leishmaniasis
MS/MS	Fragment spectrum
MSI	Mass spectrometry imaging
MudPIT	Multidimentional protein identification technology
MWCO	Molecular weight cut-off
Nd:YAG	Neodymium-doped yttrium aluminium garnet
PAGE	Polyacrylamide gel electrophoresis
PCA	Principal Component Analysis
PD	Plasma desorption
PDA	Photodiode Array
pI	Isoelectric point
PMF	Peptide mass fingerprinting
RF	Radio frequency
ROI	Region of interest
RP	Reverse phase
SCX	Strong cation exchange
SDS	Sodium dodecyl sulfate
SNP	Single nucleotide polymorphism
TEMED	Tetramethylethylenediamine
TFA	Trifluoroacetic acid
TGS	Tris-glycine-sds buffer
TOF	Time-of-flight
VL	Visceral leishmaniasis

CHAPTER 1

PROTEOMICS BY MASS SPECTROMETRY

1.1 General Aspects of Mass Spectrometry

Mass spectrometry (MS) is the name of analytical method which uses mass spectrometers to study gas-phase ions. Although the term refers to “mass”, it is actually mass-to-charge ratio (m/z), a dimensionless quantity that is measured by the instrument. For the output, a defined mass range is plotted on x axis while ion abundances are on y in a mass spectrum. Since it is not the quantity of absorption or emission of electromagnetic radiation, the term “spectroscopy” should be avoided. Mass spectrometry has gained a place in various fields with different applications due to its outstanding features gathered in a single exclusive instrument. Sensitivity, resolution, speed, detection range and analyte diversity were improved in time and are still being perfected.

In the evolution history of mass spectormetry, some milestones had a great impact on the advancement that their cumulative contributions shaped today’s sophisticated instrument (Table 1.1). Among them, combining with gas chromatography (GC) in-line has turned mass spectrometers to a special type of detector thus a more versatile method, GC-MS has been generated which can separate, concentrate, purify, measure and provide m/z spectra of a mixture at once. Later, advent of soft ionization techniques caused to expand the limited analyte type from small, volatile molecules to large, non-volatile, labile polymers. Thus liquid chromatography (LC) could enter to the game too as LC-MS. Actually this progress was a blessing for the analysis of samples from biological sources since these samples can be very complicated and minute. In deed, these two soft ionization techniques, matrix-assited laser desorption ionization (MALDI) and electrospray ionization (ESI) had emerged just in time that upcoming post-genomic era was boosted in a quite short time under the name of a new concept as proteomics. Although biological sample analysis started with proteins, today several types of biopolymers and small molecules including lipids, nucleic acids, oligosaccarites, metabolites, and even tissue sections and viruses can be analysed by

MS. In time, such kind of tasks are given the name by adding suffix, -omics to the target analyte. However, each branch requires expertise for sample preparation and data accumulation due to the natural characteristics of the analyte. Eventually, system biology aims to unravel the signal transduction and metabolic pathways by combining accumulated data and drawing networks[1].

Table 1.1. Chronological list of some important developments in mass spectrometry

Year	Developed/Introduced	by (Person/Company)
1897	Electron and its mass-to-charge ratio (Nobel Prize in 1906) [2]	J. J. THOMSON
1909	Elementary unit of charge[3]	R.A. MILLIKAN; H. FLETCHER
1912	Parabola spectrograph as the first mass spectrometer.	J.J. THOMSON
1918	Electron ionization source and the first spectrometer with a sector-shaped magnet [4]	A.J. DEMPSTER
1919	First mass spectrometer with velocity focusing[5]	F.W. ASTON
1948	Design of LTOF mass spectrometer[6]	A.E. CAMERON; D.F. EGGERS
1949	First application of ion cyclotron resonance in MS[7]	H. SOMMER; H.A. THOMAS; J.A. HIPPLE
1952	Development of double-focusing instruments[8]	E.G. JOHNSON; A.O. NIER
1953	Quadrupole analyse, the ion trap of quistor in a patent[9]	W. PAUL; H.S. STEINWEDEL
1956	Mass spectrometers coupled with a gas chromatograph (GS-MS)[10, 11]	F.W. McLAFFERTY; R.S. GOHLKE
1966	Chemical ionization (CI)[12]	M.S.B. MUNSON; F.H. FIELD
1967	Collision induced dissociation (CID) procedure[13, 14]	F.W. McLAFFERTY; K.R. JENNINGS
1968	First commercial quadrupole mass spectrometer	Finnigan
1969	Field desorption (FD) mass spectrometry of organic molecules[15]	H.D. BECKEY
1972	Reflectron [16]	V. KARATEV; B. MAMYRIM; D.V. SMIKK
1974	Atmospheric pressure chemical ionization (APCI)[17]	E.C. HORNING; D.I. CARROLL; I. DZIDIC; K.D. HAEGELE; R.N. STILLWELL
1974	First spectrometers coupled with a high-performance liquid chromatograph[18]	P.J. ARPINO, M.A. BALDWIN; F.W. MCLAFFERTY
1974	Fourier transformed ICR (FTICR) mass spectrometry[19]	M.D. COMISAROV; A.G. MARSHALL
1976	Plasma desorption (PD) source[20]	R.D. MACFARLANE; D.F. TORGESSON
1978	First triple quadrupole mass spectrometer[21]	R.A. YOST; C.G. ENKE
1980	Potential of inductively coupled plasma (ICP) MS[22]	R.S. HOUK; V.A. FASSEL; G.D. FLESCH; A.L. GRAY; E. TAYLOR
1982	Triple quadrupole mass spectrometers	Finnigan and Sciex
1987	Matrix-assisted laser desorption/ionization (MALDI)[23, 24]	T. TANAKA; M. KARAS; D. BACHMANN; U. BAHR; F. HILLENKAMP
1987	Coupling of capillary electrophoresis (CE) with MS[25]	R.D. SMITH
1988	Electrospray[26]	J. FENN
1994	Proteomics[27]	M. WILKINS
1994	Nanospray[28]	M. WILM; M. MANN
1999	Orbitrap[29]	A. A. MAKAROV

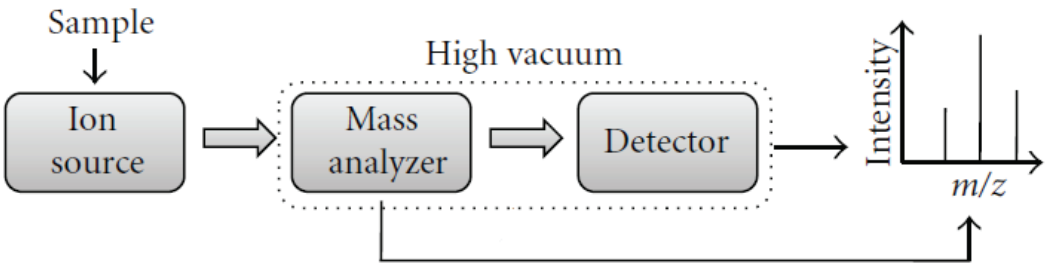


Figure 1.1. Basic components in a typical mass spectrometer

Electron is accepted as the earliest particle whose relative m/z was measured. Later, analyte type is expanded however for a long time tradional mass spectrometry was confined to small, volatile and thermally stable molecules. Since early 1990s, mass spectrometry has gained ground with regards to sample variety as well as resolution, sensitivity, speed, automation and high-throughput capability. Various types of mass spectrometers have been released to the market, some of which are equipped with specialized hardwares and softwares for particular researches however containing three main components (Figure 1.1) namely ion source, mass analyzer and detector is the shared base feature for all mass spectrometers. Among them, ion source is the part where analytes are ionized thus positively or negatively charged particles could be manipulated in the following section called analyzer. There are several different types of ion sources in mass spectrometry. In electron ionization (EI), chemical ionization (CI) and field ionization (FI), molecules must be introduced in gas phase therefore sample is either vaporized by heat or transferred from GC. Accordingly these ionization techniques are not suitable for thermally labile and nonvolatile analyte group which encompass biological macromolecules and polymers. Another concern is the integrity of the molecule to be measured. If the internal energy deposited during the ionization is too high, molecular ion of the analyte can not be observed. In electron ionization, electrons that are exposed to 70eV are released from a heated flament and collide with gaseous analytes. As a concequence, extensive fragmentation is observed on the spectrum therefore structures of small molecules, mostly organic compounds lower than 600 Da, can be investigated by this hard ionization technique. Although configuration of CI is very similar to that of EI, it is somewhat softer technique since less fragmentation occurs allowing the detection of molecular ion. Instead of energetic electrons, ions of the reagent gas that was generated by EI, react with the analyte and series of reaction involving transfer of electron, proton or other charged species take

place. Methane, isobutane or ammonia can be used as the reagent gas however fragmentation pattern of the same analyte changes. Field ionization (FI) and field desorption (FD) rely on the formation of high electric field due to potential difference of 8-12 kV between two closely placed electrodes. Special desing of the metal surface of the anode allows to remove an electron from the analyte by quantum mechanical tunneling. These two techniques are softer than EI and CI that they exhibit almost no fragmentation. In addition, field desorption technique is the combination of desorption and ionization thus sample does not have to be in the gas phase. Despite the analysis capability including termally labile and relatively large molecules, FD attracted attention for a short time due to the advent of more effective ionizations, electrospray (ESI) and MALDI. MALDI is an abbreviation for matrix-assisted laser desorption/ionization. These two techniques meet the need for mass spectrometric analysis of large molecules with minimal fragmentation and both have great contributions to biological sample analyses. In ESI, analytes are introduced to the system in liquid phase while samples are co-crystalized with a matrix compound thus in solid phase in MALDI. Process of desorption by laser ablation can only perform if matrix absorbs the light wavelength of the laser thereby rapid heat on the surface cause matrix to evaporate together with the analyte molecules. MALDI with nitrogen laser works at 337 nm (at 3,68 eV) and the absorption wavelength for neodymium-doped yttrium aluminium garnet (Nd:YAG) laser is 355 nm (at 3,49 eV). Ionization mechanism is not well understood although matrix is thought to be the proton source due to the acidic nature of many matrices. In addition, solvent of the analyte should dissolve the matrix, also matrix to sample ratio is usually kept high to protect the analyte from excess energy. Although MALDI tolerates salts and other contaminants more than ESI, sample preparation is critical for the quality of the spectrum. Shot-to-shot reproducibility is considerably affected from heterogeneous crystal formation which can be altered by matrix deposition methods. Another favorablable feature of MALDI is preferential acquisition of singly charged ions which renders the data interpretation simpler[30]. On the other hand, ESI produces dominantly multiply charged ions however this feature is exploited efficiently in the analyses of intact biomolecular structure associated with other tailored methods for this purpose. As an example, protonation of a folded protein is low than that of its unfolded state, from this point of view studies regarding the protein folding kinetics, protein conformational dynamics and noncovalent ligand interactions are conducted by the interpretation of charge

alterations and charge state distribution in the gas phase[31]. The reason of carrying so many charge is gradual formation of ions from a fine mist of droplet whose surface is surrounded with electrical charge. As the solvent evaporates, a critical limit namely Rayleigh limit is exceeded thus smaller droplets occur by electrostatic repulsion soon after outer molecules overcome the surface tension. This process is repeated several times and highly charged sample ions are conveyed to the analyzer in gas phase through the ion transfer capillary[32]. High voltage is applied to the ESI needle to during the formation of aerosol however protonation availability is related with the characteristics of the analyte (Figure 1.2.a). For instance, tryptic peptides are mostly observed as doubly charged because of the first and last amino acids in the peptide backbone. First amino acid holds the amine functional group and last amino acid holds amine group at the residue of arginine or lysine. If trypsin fails to cleave a lysine or arginine amino acid, triply charged peptide is expected to appear on the spectrum[33]. Low flow rate version of the electrospray called nanospray is predominantly preferred for today's proteomic analysis due to minute consumption of the sample[34]. Lastly, continuous flow of ESI has the capability of the quantification analysis comparing to MALDI which collects data intermittently through the laser pulses[35].

Analyzer is the essential component of a mass spectrometer in which ions are separated according to mass to charge ratio before reaching to the detector. Different type of mass analyzers can be listed as magnetic sector, electric sector, quadrupole, ion trap, time-of-flight, ion cyclotron resonance (ICR) and orbitrap. Among them, last five are taken place in proteomic studies. Double focusing instruments are the combined version of sector mass spectrometers. Modern sector mass spectrometers evolved to involve up to five sectors with different geometries however they are mostly used to study reactions in organic chemistry. ICR outperforms all others in terms of high resolution however it could not be widely used due to issues regarding maintenance and price. The newest technology, orbitrap offers affordable price and considerable resolution which caused it to draw attention in biological applications. Ion trap and quadrupole analyzers take advantage of being the earliest and cheapest instruments which results with accumulated information in the literature and improvements as well. TOF analyzers also provides quite high resolution especially after the incorporation of reflectron. Depending on the separation principle, accelerated charged ions are controlled by electrostatic field, magnetic field or electromagnetic (RF) field. On the other hand, ions are trapped in ion trap (Figure 1.2.b), FTIC and orbitrap based on the

resonance frequency of corresponding m/z. When ions are not manipulated with neither of these fields, separation relies on the velocity of the ions in the flight tube which is the case in TOF analyzers. After the accelaration to a constant kinetic energy, high mass ions are expected to travel slowly than the lighter ones because square root of an ion's m/z value is inverse function of its velocity. (Figure 1.2.c). In this context, length of field-free drift region becomes significant for the resolution therefore ions are manipulated with a device called reflectron. A series of electrical lenses in the reflectron first repel the ions to slow down to a point that they come to a rest than reaccelerates them into second field-free region to reach to the other detector. Finally quadrupole analyzers, in other words, mass filters allow only selected m/z to pass through the ion path by the force of oscillating electric fields applied to the quadrupole rods. Ions should have a stable trajectory to reach to detector otherwise they hit to the walls of the quadrupole's rods or escape through the rods[36].

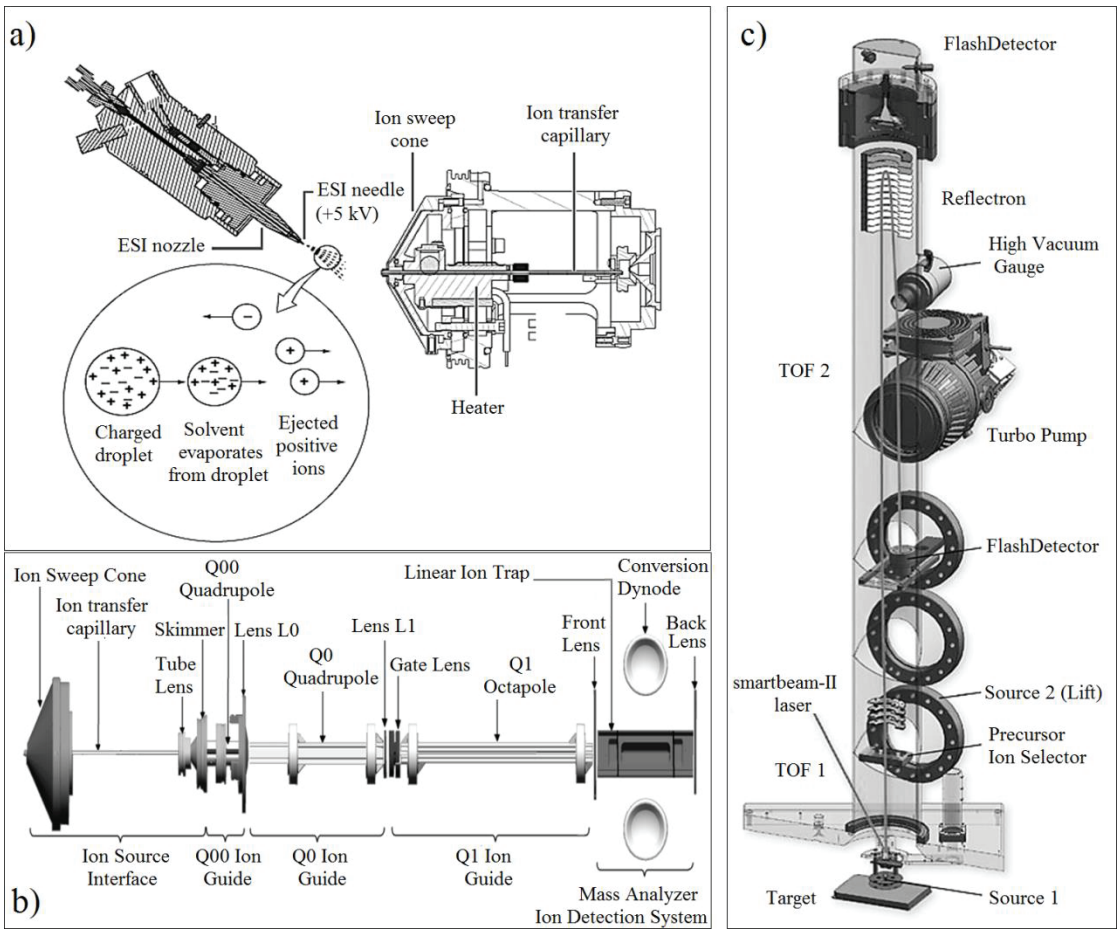


Figure 1.2. Schematic diagram of mass spectrometers used in this study. a) Electrospray b) Internal components of MS detector with the ion path units of LTQ XL^{TM} ESI-Linear Ion Trap MS lauched by ThermoFisher Scientific, c) Ion path units of Autoflex MALDI-TOF/TOF MS lauched by Bruker Daltonics

Fragmentation capability is the most effective feature what make mass spectrometry so versatile and powerful instrument since it enables scientists to elucidate the molecular structure and reaction monitoring. Tandem mass spectrometry refers to the combination of two mass analysers although it is also way of addressing the fragmentation process, MS/MS. Usage of trap type mass analysers (ion trap, FT-ICR) for the MS/MS analysis is defined with the term tandem-in-time. Unlike the mass spectrometers operate with tandem-in-space, these instruments provide higher-order fragmentation which is denoted with MS^n accordingly MS/MS corresponds to $MS^{'2'}$. On the other hand, tandem-in-space type mass spectrometers are simply manufactured by reuniting two analyzers with a collision cell[37].

Mass spectrometry determines isotopes as a reflection of their abundancy which especially helps to distinguish halogens, Brom and and Chlor. When it comes to carbon atom, the primary component of biological macromolecules, among naturally occurring stable isotopes of carbon 12 and 13, approximately 1% of matters is comprised of carbon 13. In conclusion, number of mass signals on the spectrum increases and isotope distrubution pattern changes as the molecule to be measured gets larger. When the calculation is based on the most abundant isotope, either monosiotopic mass or nominal mass terms are used. While monoisotopic mass is calculated according to the exact masses (i.e., $_1$H:1.007825, $_{12}$C:12.000000), non-integer numbers of decimal numbers (i.e., H:1, C:12) are dropped in the calculation of nominal mass. In the average mass calculation, all natural isotopes of the element are involved to the calculation of the elemental mass (i.e., H: 1.00794, C:12.011) which causes considerable shift towards higher masses for large molecules[38].

1.2. Separation Techniques Prior to Mass Spectrometric Analysis

Among two widespread methods, the one named as “gel-based” relies on two-dimensional polyacrylamide gel electrophoresis (2D-GE) which has been described by O’Farrell in 1975. This method is the combination of separations based on protein isoelectric point and protein molecular weight[39]. First separation of the 2D-GE is always isoelectric focusing. This technique benefits from proteins’ zwitterion character that is to say they carry functional groups holding at least one positive and one negative electrical charges. When positive and negative charges on a protein are equal, protein

exhibits zero net charge to the environment thus this pH value is identified as isoelectric point (pI). In a medium with pH gradient, protein moves towards the opposite charge under the influence of electric field. As the protein approaches to a certain point it slows down and eventually stops migrating. Finally, proteins are accumulated and condensed at their characteristic pI values. 2D-GE has gained much more reproducibility after the advent of immobilized pH gradient (IPG) strips that are replaced with carrier ampholytes. Today's routine IEF gel strips has fairly facilitate the association of IEF method to SDS-GE in comparison to cumbersome and unsuccessful tube gel mode.

Electrophoresis takes play in the generation of the second dimension. It is one of the most widely used analytical tool in which the charged molecules are migrated in an electrical field. Electrophoretic methods are usually carried out in the aim of separation rather than purification since proteins' structure and function are mostly affected adversely from the technique. Polyacrylamide gels are generated based on the free radical polymerization of acrylamide and cross-linking agent N,N'- methylene-bis-acrylamide. Thus, under the control of the initiator-catalyst couple, ammonium persulphate-N,N,N',N'-tetramethylethylenediamine (TEMED), a physically stable porous matrix is obtained to resolve proteins. Since dodecyl sulfate covers all proteins in proportion to their mass, negative charge to mass ratio is balanced for all proteins thus proteins are distinguished according to the size while they are dragged through the pores. The percentage of acrylamide, often referred to %T (total percentage of acrylamide plus crosslinker) determines the pore size of the gel and the relation is simply given as higher the %T, smaller the pore size[40]. 2D-GE technique provides a readout to visualize hundreds to thousand of proteins. Disease state differences, toxic influences and stress impacts have been revealed by comparing readouts of two samples in different states[41].

The method underlying gel-free proteomics is high performance liquid chromatography technique (HPLC), which incorporates many features including high recovery, reproducibility, speed and particularly superior resolution. In addition, HPLC provides various separation modes such as hydrophobic interaction, reverse phase (RP), hydrophilic interaction, ion-exchange, gel filtration, immobilized metal ion affinity and immunoaffinity. Combining HPLC to electrospray mass spectrometers generates an excellent on-line procedure, in which the ion-suppression effect is reduced, low abundance peptides are enriched, and salts are removed. Mostly RP mode is coupled to MS owing to its appropriate mobile phase content. Despite its efficiency and

practicality, one dimensional liquid chromatography seems insufficient considering the complexity of the proteomes and a vast amount of resultant peptides from the digestion. Therefore, multi dimensional or orthogonal chromatographic techniques are developed for better resolution. In these techniques, RP-LC is always placed prior to mass analyzer, and the preceding separation is primarily carried out by ion-exchange chromatography. On the other hand, affinity chromatography has been used efficiently for targeted proteins like phosphorylated or glycosylated ones; however, size exclusion chromatography is occasionally utilized because of its low resolution and limited loading capacity. Apart from these, LC is an inevitable segment for quantitative proteomics. Mostly gel-free method is referred to shotgun proteomics. In this approach, extracted protein mixtures are digested in solution and the resulting hundred of thousands peptides are fractionated, mostly by LC and then identified by data dependent acquisition mode of automated LC-MS/MS[42].

1.3. Proteomics

Image of a jigsaw puzzle can be recognized even it is incomplete however that image should be seen before and kept in the memory to be recognized at that moment. Similarly, to be able to identify a protein, we do not need to attain the whole sequece of the protein, selectivity of partial sequences enables us to relate it to the protein if it was somehow identified earlier and recorded in the databases. For sure, reconstructing the whole sequence is favoured than collecting the partial sequences, actually it is a must more than a preference to determine the complete characterization and function. However if we wish to perform *a global scale analysis of proteins* in a complex protein mixture, only identification of the proteins present at that predefined conditions and time will do the job to uncover protein interactions and eventually the metabolic pathways. Herein, original scope of proteome as *the entire protein complement expressed by a genome or by a cell or tissue type* seems somewhat utopic notion for a single proteome analysis since proteins are not constant like DNA, they have complex structure and dynamic nature. They are synthesized by demand in the cell and their copies can not be produced *in vitro* as can be done in genomics[43]. Therefore this scope is generally specialized and narrowed by defining the method, material, purpose, working area etc. such as bottom-up proteomics, plant proteomics, structural

proteomics, functional proteomics, targetted proteomics, phosphoproteomics and so forth. In conclusion, classical protein chemistry has turned into proteomic studies that are dependent to mass spectrometric methods and all accumulated protein contributed to validate the connection with the genes as well as to create organism's proteome[44]. Owing to the technological developments and intense user demand, simple protein identification methods enriched with more useful protein quantification workflows.

Total protein mixture obtained from a living contains several types of protein whose physicochemical property and amount dramatically differ from eachother. Since selectivity of the partial sequences from a distinct protein serve the purpose of identification of that protein, smaller pieces of the whole sequence namely peptides are exploited thus proteins could be investigated in simpler structure. For this purpose, proteins are treated with digestive enzymes which cleave from a specific position preferentially. The principal enzyme used in this procedure is trypsin which hydrolyzes the peptide bonds of lysine and argine amino acids from their C terminus. Trysin is the most suitable enzyme due to convenient size of acquired peptides because these two basic amino acids are common and well distrubuted in the sequence of almost all proteins. Alternative enzymes namely chymotrypsin, LysC, LysN, AspN, GluC and ArgC are used for further analysis which increase the identified sequence coverage by completing the missing parts[45]. Experimental setup is determined according to the conditions which maximize the catalytic activity of the enzyme therefore unfolding proteins, reduction of disulfide bonds and incubation at the optimum pH and temperature should be taken part in a typical digestion protocol[46]. Folded structure of a protein can be disrupted by using chaotropic agents or surfactants however this compound should not be interfere with enzyme digestion or mass spectrometric analysis. That's why molarity of widely used chaotropic agent, urea is lowered from 6-8 to below 1 by dilution or excess urea is removed by using centrifugal filters before the addition of trypsin to protect the enzyme from unfolding. If surfactant is used for this purpose instead of urea, mass spectrometry incompatible surfactant such as SDS should be avoided. Ammonium bicarbonate salt and Tris-HCl buffer are mostly used to keep the solution pH above 8. For completion of the digestion process, samples are kept at 37°C from four to sixteen hours[47]. Protein identification by the peptides recovered after protein digestion is called bottom-up proteomics and shotgun proteomics is a branch of bottom-up proteomics in which complex protein mixtures instead of isolated proteins are digested. Actually shotgun sequencing is a genome sequencing method in which

DNA is randomly broken and reassembled according to the overlapping regions[48]. Just like the analogy between proteomics and genomics, shotgun proteomics has taken its name from shotgun sequencing of genomics due to the technical similarity. The technique behind the scenes of shotgun proteomics is the data-dependent acquisition operation of mass spectrometers (Figure 1.3) in which both MS scan and MS/MS fragmentation are performed consecutively. In this automatic data collection mode, survey scan is followed by MS/MS analysis however only m/z values which were recorded according to predetermined rules are subjected to fragmentation[49, 50].

Protein pre-separation prior to peptide fragmentation by 2D-SDS-PAGE or LC fractionation reduce the complexity of total protein mixture or ideally isolate the protein. In such cases, connection between protein and peptide is maintained and the products obtained from enzymatic digestion are all attributed to that isolated protein. In adition to MS/MS fragmentations of the peptides, identification of protein from the masses of each peptides also called peptide mass fingerprinting becomes possible[51]. On the other hand, in shotgun proteomics this connectivity is lost therefore only fragmentational spectra can be utilized for the assignment of the peptide from a highly complicated peptide pool to the correct protein. High throughput data acquisition in short time without the laborious protein separation is the main strength of shotgun proteomics method however this feature comes with a price so called protein inference problem which is arisen from the peptides in common[52]. At least one distinct protein region which is unique to that protein meets the requirement to differentiate it from other proteins even if the rest of the peptides are shared between two or more proteins. More unique peptides will reinforce the reliability of protein identification. In the lack of unique peptide match, peptides from protein isoforms, protein of similar family and sometimes short sequences match to more than one protein while complicating the bottom-up route from peptide to protein due to indistinguishability. Identification of a protein by a single peptide is unfavourable that they are either considered as doubtful protein or completely disregarded[52].

In shotgun proteomics, we rapidly attain from hundreds of protein to thousands of peptides only by one step, tryptic digestion. Although no protein separation in the upstream processes is an advantageous with regards to less experimental burden, this time we end up with a great complex mixture of peptides which should be reduced to a measurable degree before LC-MS/MS in order to give a chance to more peptide for the fragmentation. It can be achieved by direct analysis in an on-line or by collecting

nonidentical fractions in an off-line manner[53]. In off-line LC, peptide elution is fractioned and those collected fractions are evaporated and then resuspended before the LC-MS/MS analysis. Another off-line method, in which LC is not involved, is isoelectric focusing of peptide mixture following IPG strip cutting. In this earlier fractionation method, recovery of the peptides from strip pieces is carried out by immersing pieces in a solution. For on-line 2D-HPLC, either two separate HPLC systems are connected or output of one sophisticated HPLC system with advanced pumping unit is used. This unit has two different pumps with separate flow lines thus flow management part can divert the flow from one column to the other. Apart from that, by using long packed columns in one dimensional LC can provide high peak capacity in peptide separation. Although self-packed extremely long columns up to 2-4 meters exhibited much better results, mostly 50 cm and 75 cm long columns are mostly available in the market[54,55]. Last but not least, column packed materials of cation exchange HPLC and reversed phase HPLC are combined by loading in series to generate a diphasic or triphasic HPLC column in the method named multidimentional protein identification technology (MudPIT). This connection become available due the compatibility between the strong cation exchange (SCX) chromatography mobile phase and second reverse phase separation thus peptides are retained on SCX while salts are sent away from reverse phase column as well as resolved better due to increasing peak capacity during the elution towards MS.

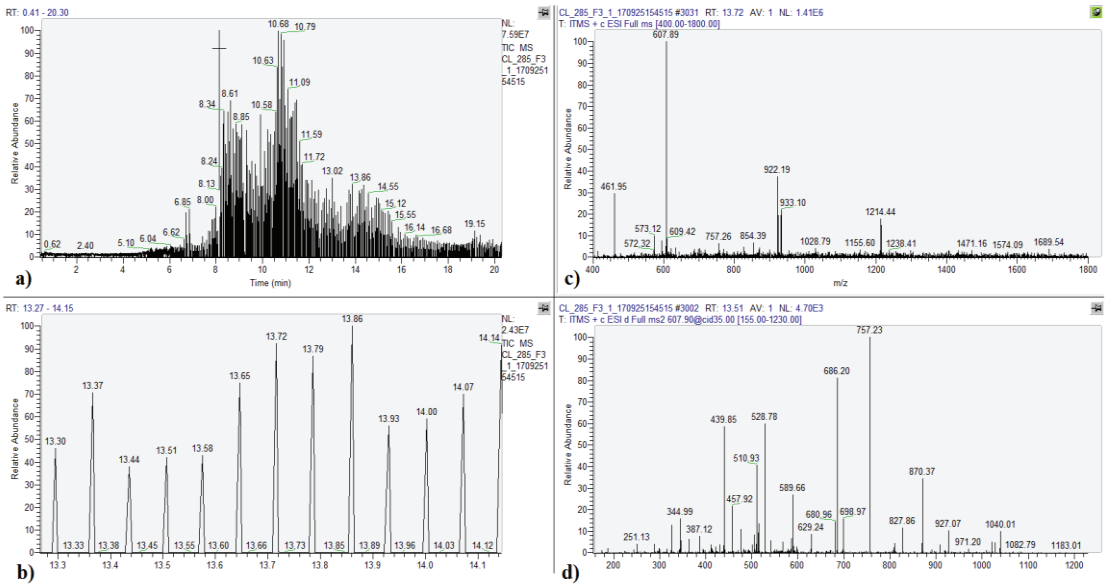


Figure 1.3. Outputs in automatic LC-MS/MS analysis with data-dependent acquisition. a and b) Mass chromatograms, c) MS Spectrum, d) MS/MS Spectrum

1.4. Peptide MS/MS Spectrum and Fragment Ions

Amino acid sequence of a peptide can be attained from the product ion spectrum that is acquired in a tandem mass spectrometry experiment through the collisionally induced dissociation (CID) mostly at positive mode. Fragmentation characteristics of peptides which is related with the position of the proton should be known to interpret those product ion spectra. In a doubly charged tryptic peptide, there are two distinct protonation sites, one is at the N- terminal amino group and other is at the amino group of basic residues, arginine and lysine. In the gas phase, charge on the less basic amino group at the N- terminal migrates on the peptide backbone and localized on the nitrogen atom of any one of the amide bond. As postulated in the mobile proton hypothesis, cleavage can be occurred from any of the peptide bonds. Since either fragments of N- or C- terminus can carry the charge, various masses of product ions are observed on the spectrum. Some type of product ions are observed more often with higher intensities on the MS/MS spectrum of a peptide. For the recognition of particular ions, product ion series are designated considering the N or C terminus retained in the ion. According to this nomenclature (Figure 1.4), ions contain the N- terminus are represented with a-, b- and c- letters while corresponding C- terminus ions are represented with x-, y- and z-. Apart from that, ions occur from the side chain cleavage due to higher energies at keV degree are labelled with w-, v- and d-. They are also called satellite ions and not observed in low CID. Since two functional groups, amine and carboxylic acid are attached to the α-carbon; β, γ, δ and ε carbons in order refer to the C atoms on the side chain or in other words R group. As can be seen on Figure 1.4, structures of w- and z-ions, similarly those of d- and a- ions are almost identical. Actually w- ions are formed from z+1 ions after the cleavage of β-γ bond and d ions from a+1 ions likewise. In the same vein, v-ions are formed from y-ions via cleavage of α-β bond corresponding to removal of the entire side chain. The w-and d- ions are important to distinguish Leu from Ile. For instance, it is known that w-ions of Ile is always 14 Da higher than that of Leu. In CID, loss of ammonia corresponds to 17 Da mass shift when R, K, N and Q amino acids are involed to fragmentation and they are represented with a*, b* and y*. Likewise, loss of water is recognised with 18 Da mass difference when S,T,E and D amino acids are involded and those ions are denoted a°, b° and y°. Finally, diagnostic ions observed at the low-mass end of an MS/MS spectrum are called immonium ions

and denoted by one-letter code for the corresponding amino acid. Single amino acid immonium ions have the structure $H_2N{=}CHR^+$ and formed by an internal fragment with just a single side chain[56].

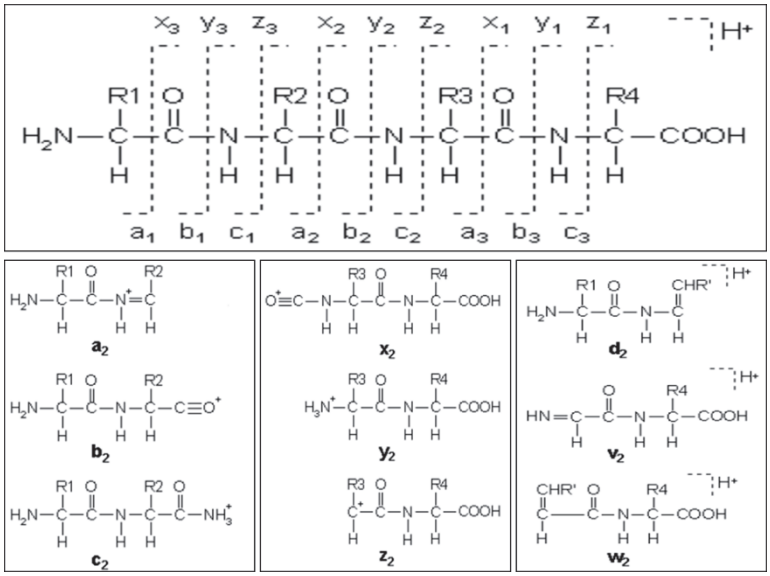


Figure 1.4. Nomenclature for peptide fragment ions and their structural formulas [56, 57]

1.5. Peptide Sequencing and Database Search

In biochemistry, traditional peptide sequencing method namely Edman sequencing is based on repetitive degradation of each amino acid starting from the one on the N-terminus. In the Edman reaction, phenylisothiocyanate molecule is attached to the N-terminus of the peptide thus a phelythiocarbamyl derivatized protein which has a cyclic intermediate encapsulating the N-terminal amino acid is generated for specific hydrolysis at the nearby amide bond. Released amino acid derivative is determined by HPLC and reaction cycle continues for the adjacent amino acids by using same reagents[58]. Although this technique provides accurate results, its capacity is not suitable for proteomics studies which requires high throughput analysis because it is only applied to pure protein or peptide sample and not convenient for complex protein mixtures. On the other hand, MS-based proteomics with database search owes it reputation largely to rapid sequencing of peptide mixtures as well as characterization of post-translational modifications[59]. As a part of MS-based proteomics, purified isolated protein is necessary for peptide mass fingerprinting (PMF) too, in fact, PMF is a protein identification technique which does not rely on peptide sequencing. In this technique,

mass spectrum containing mass values of individual peptides is searched in database for the potential protein whose products of enzymatic digest overlap with those masses. Before the PMF database search, interfering peaks originated from trypsin and keratin peptides, matrix signals and other salt adducts should be removed from the m/z values list not to mislead the search engine if they are certainly determined earlier[60]. However attribution of m/z values to a peptide in the MS sepctrum does not provide an identification as selective as the fragment ions in the MS/MS spectrum therefore identification through the database search of MS/MS spectra has become the mainstream in protemics. Although peptide inference from the MS/MS spectrum without the need of database which is called *de novo* sequencing is possible through the interpretation of MS/MS spectrum, an extensive fragmentation involving all complementer ions is required for a correct detection. Unfortunately, most of the peptide MS/MS spectra do not exhibit qualified fragmention to that extent. PepNovo, PEAKS, Lutefisk and Novor are some examples to softwares that can be used for this purpose. Apart from that, searching tandem mass spectrometric data in the peptide spectral libraries by specified algorithms started in the mid 2000s however this technique is not widespread and mostly limited to targeted analysis. Nevertheless, high-quality mass spectra are being accumulated by collective contribution and potential of the method is increasing by the coverage of spectral libraries expands.

Searching uninterpreted MS/MS data in predefined database is the method of choice for the large scale proteomic studies thanks to sensitive, rapid and automatized mass spectrometry systems. However this technique requires a database as is evident from the name therefore success of proteomic analysis depends on attainable genomic sequence. Genetic code is translated to amino acid sequence in FASTA format which represents nucleotids and amino acids with single letter. In-house softwares or internet based resources like NCBI blastx and expasy/translate can be utilized to translate the genes when database of interest is not available and built manually. Since high abundant and frequent proteins are well examined in time, their corresponding genes, function and features are documented in detail. Uncharacterized proteins, in other words hypothetical proteins correspond to proteins whose expression or functions are not well defined yet. When genetic information is not available for the organism of interest, the closest option in the taxonomy is chosen to catch a homology between the proteins in the database or alternatively it is searched against the all entries. Database search engines work on similar principle that relying to corresponding score for each matching.

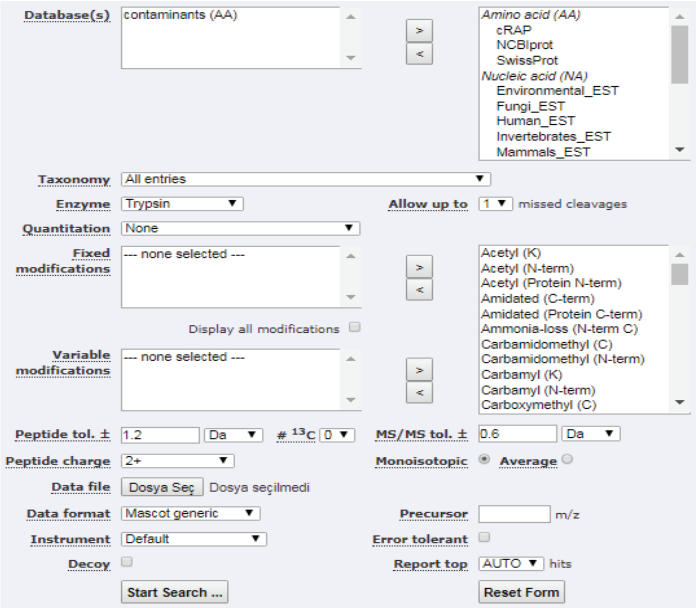


Figure 1.5. Setting interface of Mascot (Matrix Science)

Several search engines have been developed to assign automatically collected large number of spectra to the in-situ digested peptides from the proteins in the database. Before we execute the search, some particular parameters are required to be set (Figure 1.5). Minimum requirements including the selection of database, enzyme, precursor tolerance (peptide tol.), fragment tolerance (MS/MS tolerance), miscleavage, peptide charge must be enterred. Additionally, determination of the expected modifications improves the results. When search is completed, we obtain a collection of proteins with a corresponding scores. This scoring function enable us to evaluate the quality of the match. To control the error rate, false identifications should be known. We can not discriminate the false identifications however we can control the degree of false identifications by searching the data against a fake database. This database contain decoy sequences that should not be matched to any spectrum if it does we can be sure that decoy hit is a false identification. Decoy database should resemble the target database in size, amino acid distribution and protein length therefore revesed proteins are being commonly selected to create an incrorrect protein list[61]. 1% is widely accepted value for proteomics studies though lower values can be chosen. Additionally, search engines also provides a numerical value (e.g. emPAI[62]) which is calculated based on spectrum counting according to the rational, abundant proteins spend more time in the system during data acquisition thus spectral counting enables relative quantification however high resolution MS, specialized softwares and expertise are required.

CHAPTER 2

GEL-BASED COMPARATIVE PROTEOMIC ANALYSIS OF LEISHMANIA ISOLATES OBTAINED FROM VISCERAL AND CUTANEOUS LEISHMANIASIS PATIENTS

2.1. Introduction to Leishmaniasis

Infectious diseases whose pathogens are spread through vectors like mosquitoes, insects, ticks, fleas, flies etc. fall under the group of vector-borne diseases. Source of these infections can be viral, bacterial, eukaryotic unicellular or large multicellular (worm-like helminths)[63]. As a vector-borne disease, leishmaniasis is caused by a parasitic protozoans of leishmania genus and transmitted by female *Phlebotomine* sandflies. This disease is widespread on many countries in tropical and subtropical regions where the climate is appropriate for flies therefore it is not seen in Antarctica. Since flies thrive under poor living conditions and rural area where clean water, sanitation and nutrition are inadequate, communities in poverty especially living in underdeveloped countries predominantly confront with this problem[64]. Although disease targets large populations, generates serious healt problems and causes significant economic burden in many countries, it does not receive much attention and funding in comparison to some other infectious diseases such as AIDS, malaria and tuberculosis. Today twenty disease sharing the similar fate are gathered under the title of "neglected tropical diseases" to eradicate the infections and put a spotlight on public-health interventions. Buruli ulcer, chagas disease, dengue and chikungunya, dracunculiasis, echinococcosis, foodborne trematodiases, human african trypanosomiasis (sleeping sickness), leprosy, lymphatic filariasis, mycetoma, chromoblastomycosis and other deep mycoses, onchocerciasis, rabies, scabies and other ectoparasites, schistosomiasis, soil-transmitted helminthiases, snakebite envenoming, taeniasis or cysticercosis, trachoma and yaws (endemic treponematoses) are the other members of this categorization[65]

According to the World Health Organization (WHO) reports, over one billion people are at risk in endemic areas. Estimated numbers of new cases vary between 600.000 to 1.000.000 in cutaneous leishmaniasis and 50.000 to 90.000 in visceral leishmaniasis. Actually, a great majority of the parasite infection does not result with the disease which implies the significance of the host's immune system however this should not be confused with asymptomatic infections. In asymptomatic leishmaniasis, positive results are seen on the test although patients do not seem sick. In time, disease can develop in those patients, if not, their surveillance should be continued since they are still potential reservoirs.

Depending on the continent, leishmaniasis is divided into two groups, New World is referred to Americas and Old World to Africa, Asia and Europe. While sandflies of Lutzomyia genus are found in the New World, vectors from Phlebotomus genus transfer the parasite to host in the Old World. About 100 countries are affected from the disease. Currently, it does not pose danger in Australia as human pathogen however a leishmania species namely *Leishmania australiensis* has been encountered in kangaroos[66]. Unfortunately, some leishmaniasis species are transmitted from animal to human. Dogs are one of the major domestic reservoirs therefore control of zoonotic leishmaniasis is directly related with protection of pets from the bites of sandflies and surveillance of their health as well. Other animal reservoirs involve cattles, horses, cats and rodents. Causative agent of zoonotic visceral leishmaniasis is *Leishmania infantum* on the other hand visceral leishmanisis caused by *Leishmania donovani* is counted as anthroponotic type in which infection is spread between humans and from humans to animals.

For infectious diseases, public awareness is very important and leads to a strong prevention. For this reason, we should better recognize the transmitter and its habits. Sandflies are quite small, approximately one-third of a mosquito. Their light and brownish color renders them hard to notice however characteristic vertical shape of their wings at rest may help to discriminate from other flies. Only females can transmit the disease since they need to suck blood for the egg development. They can not go far away, fly maximum 300 meters further in the neighborhood. They usually do not feed daytime even so wearing clothes that cover the body is advised when entered to endemic regions. In addition, usage of mosquito nets, insecticide sprays and repellent collars for dogs are recommended. Despite all of the above, permanent solution can only be attained with a dedicated national control program[67].

2.1.1. Clinical Forms of Leishmaniasis and Treatment

Depending on the body location where immotile form of the parasite namely amastigote is proliferated, three main clinical forms (Figure 2.1) as cutaneous, muco-cutaneous and visceral leishmaniases are defined. While manifestation of cutaneous leishmaniasis is skin lesions on the exposed part of the body, enlarged lymph nodes, spleen or liver along with fever, fatigue and weight loss are the symptoms of visceral leishmaniasis. The most dangerous form is visceral leishmaniasis since it may end up with fatalities if patients are not treated. Although, curation in cutaneous leishmaniasis is easier, it leaves behind unpleasant permanent scars. Muco-cutaneous form remains somehow between these two since destruction on mucosal membrane may be resulted with mutilation in nose, mouth or throat. Various species of leishmania genus is the underlying reason of these alterations in the clinical outcomes associated with the immune response of the host. Although there have been more than 50 leishmania species defined, 20 of them are pathogenic for humans. Clinical types of the disease are related with the species and one species may be the causative agent of more than one type. While *Leishmania donovani* and *Leishmania infantum* are primarly causative agents of visceral leishmaniasis; *Leishmania major*, *Leishmania mexicana*, *Leishmania tropica*, *Leishmania braziliensis* and some others are responsible for cutaneous leishmaniasis. Finally *Leishmania braziliensis* is seen in the cases of muco-cutaneous leishmaniasis with or without cutaneous leishmaniasis. Another issue about leishmaniasis, is the co-infection with human immunodeficiency virus (HIV) which causes more treatment failure and reduces survival chance[68]. After the infection, incubation period varies from a few weeks to several months. Additionally, cutaneous leishmaniasis has resemblance to other skin conditions such as pyodermitis, psoriasis, wart, sarcoidosis, etc. therefore many times it is confused and ignored. Similarly, except the swollen organs fever, weakness and weight loss do not count differential symptoms

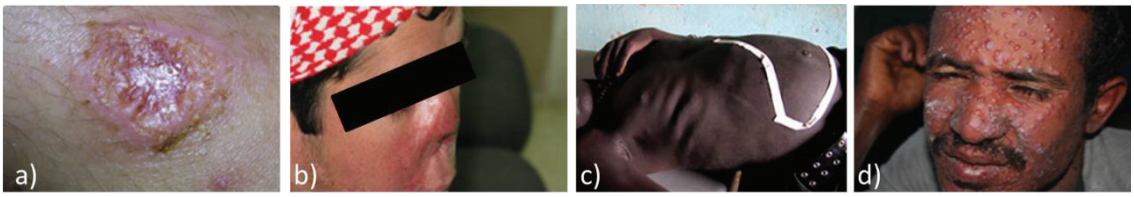


Figure 2.1. Clincal types of Leishmaniasis disease[69]: a) cutaneous, b) muco-cutaneous c) visceral, d) post kala-azar dermal leishmaniasis

Epidemiological information, characteristic symptoms and laboratory tests are incorporated into assessment criteria during diagnosis and treatment of leishmaniasis. Since skin ulceration is easier to notice and examine, diagnosis and treatment in tegumentary leishmaniasis (cutaneous and muco-cutaneous forms) is rather easier than visceral leishmaniasis. Amastigote cells are searched in the Giemsa stained biopsy specimen, scarping or impression smears under the light microscope or some specimen from lesion is cultivated in an appropriate medium. For the visualization of amastigotes in visceral leishmaniasis, surgical biopsy is applied to internal organs such as lymph nodes, bone marrow or spleen. PCR is another method for the detection of the parasite though main purpose with PCR is to identify the species. Recently, mass spectral libraries have been established for the rapid identification of leishmania spp. by MALDI-TOF mass spectrometers. In addition two serological tests are being used widely outside the laboratory for rapid examination. In direct agglutination test (DAT), agglutination of the stained and killed parasite due to specific antibodies in the blood if present, becomes observable after18 hours. The other test relies on a recombinant antigen (rK39) which is a 39-amino acid length repeat from kinesin related protein and exhibits great success with ELISA however easy access to favorable and user-friendly diagnostic equipment could be possible with the development of rK39-based immunochromatographic test (ICT) and dipstick products as well. Since drugs used in the treatment are quite toxic and symptoms are hardly shown up after the infection, accurate diagnosis with highly selective and sensitive techniques is of the utmost importance. In the first-line therapy, pentavalent antimonials in the formulations of either meglumine antimoniate and sodium stibogluconate are chosen however Amphotericin B, its liposomal form and miltefosin are started to take place of antimonials. Alternatively, some anti-fungals and antibiotic which show good anti-leishmanial activity and combination therapy are applied when necessary[70]. Although vaccine researches and trials against leishmaniasis are being carried on, a completely successful vaccine has not been developed yet.

2.1.2. Challenges in Leishmaniasis

Clinical polymorphism in cutaneous leishmaniasis (Figure 2.2) is so broad that various type of atypical lesion including erythematous, chancriform, annular, lupoid,

sporotrichoid, eczematiform, erysipeloid, verrucous, psoriasiform, pseudotumoral may emerge. Although their incidence rates are not high as many as those of typical forms; papulanodular, plaque and ulcer lesions, treatment delay and failure in those cases are more likely[71]. Infection of some particular parasite strains has self-healing feature however unusual and dreadful cases associated with weakened immune system underline the importance of host's immune state. As an example, in an atypical disseminated cutaneous leishmaniasis caused by *L.major,* one arm of the patient with history of opium addiction and obstructive lung disease, was covered with multiple papular, crusted, severely ulcerated creamy and wet-looking lesions[72]. Similarly, HIV-coinfection is very common in visceral leishmaniasis which results with more failures in treatment. Drug resistance and relapses after treatment are other concerns related with the disease. Moreover, a complication of visceral leishmaniasis may emerge as macular, maculopopular and nodular rash in recovered patients namely post kala-azar dermal leishmaniasis (Figure 2.1, d). Since it is caused by *L.donovani*, most of the cases are seen in Sudan and India. Those skin rashes acts as a reservoir for parasite therefore treatment should not be ignored. Last but not least, risk of parasite transmission to nonendemic regions even to cold Northern countries through the vector sand fly increases if warming effects of climate change is feeled worldwide.

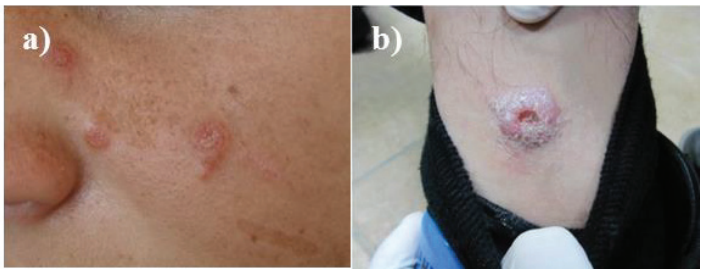


Figure 2.2. Cutaneous leishmaniasis. a) Volcano type lesion caused by *L. tropica*, b) papulonodular type caused by *L.infantum* (Dr. Özbilgin's collection)

Over 20 leishmania species can infect human and this number may increase by zoonotic species through the animal reservoirs in time. On the other side, cross-species genetic exchange between visceral and cutanous strains of *Leishmania* in the sand fly vector, as well as generation of *in vitro* hybrids are complicating pathogenicity of the disease[73,74] . Actually, genomes of *L.infantum*, *L.braziliensis* and *L.major* are quite similar that more than 99% of genes are conserved between these three species therefore differentiations were first attributed to pseudogene formation and gene loss. Later on,

single nucleotid polymorphism (SNPs) and gene copy number variations were suggested for the altered pathology of *L.donovani* in a research comparing genomes of VL and CL forms[75]. Recent studies have revealed that Leishmania parasites succeed to adapt new environments through aneuploidy, chromosomal abnormalities. More importantly, this genomic instability that occurs as changes in chromosome copy number is carried out without losing the genetic diversity. Maintenance of the diversity is explained by the ability of keeping the same combination of alleles in genetically diverse individuals during the selection of new alleles. Scientists of this collaborative research believe that these findings will bring a new perspective to the mechanisms involved in drug resistance, tissue tropism and parasite survival[76].

As the saying goes "geography is destiny", endemicity is the main determinant for the species that cause to infection. In addition, some regional variations are observed between the same incidences such as less accuracy in the rK39 ICT test in East Africa and occurrence of different clinical presentations (e.g. frequent enlarged lymph nodes in Sudanese)[70]. According to World Health Organization records, VL is mostly observed in Bangladesh, Brazil, Ethiopia, India, South Sudan and Sudan. Bolivia, Brazil and Peru are the countries that mucocutaneous leishmaniasis is mostly encountered. On the other hand, about 95% of CL cases occur in the Americas, the Mediterranean basin, the Middle East and Central Asia. In these regions, the most affected countries are Afghanistan, Algeria, Brazil, Colombia, Iran and Syria. As can be understood, our country is on the target of the disease by being a country that is bordered by Mediterranean sea and neighbore to Syria[64]. In deed, fifty thousands in round numbers cutaneous leishmaniasis cases have been reported in Turkey from 1990 to 2010 and it is anticipated that this value has been increased dramatically after migrations from Syria due to the civil war[77,78]. Apart from that, variety of species is also increased from two to four by the addition of new species namely *L.major* and *L.donovani* to our accustomed species *L.tropica* and *L.infantum* [79,80,81]. As a general problem, these different species may cause both visceral and cutanous clinical forms while it is thought that they used to cause either one. Taking into account the drug resistance and species alterations in tropism, as natural selection is working well and turning to a game changer against us, more efforts should be spent to battle against this major public health problem by raising the awareness. Besides more researches should be carried out to update our knowledge.

2.1.3. Aim of the Study

The purpose of this thesis is to study the difference between the protein expression profiles of leishmania isolates obtained from visceral and cutaneous leishmaniasis patients. To compare two sample groups to each other genetically, *L.infantum* was chosen since it causes both visceral and cutaneous leishmaniasis. However species *L.infantum* is predominantly responsible for visceral leishmaniasis and cutaneous leishmaniasis caused by *L.infantum* is not very common therefore we also included new sample group as the cutaneous leishmaniasis caused by *L.tropica* to correlate cutaneous leishmaniasis between two different species. Our goal from the comparison of two clinical forms to determine the proteins that can be used as vaccine candidate to target all parasites regardless of the species and proteins that can be useful for the differential diagnosis. The power of proteomics is arisen from collectivity thus more than one protein that are relevant to a specific function can be brought to light. In this respect, involvement of those associated proteins in a drug has the potential to increase the efficacy of the threatment. There are two well-accepted methodical approaches in proteomic studies. Gel-based approach relies on identification of protein which are isolated by gel electrophoresis and the other one does the same application to protein mixture without separating proteins. Both methods have strengths and drawbacks therefore we carried out both approaches to the same sample set to handle the problem from different angles. For statistical consideration and unbiased evaluation, we examined odd numbers of patient (5 in gel based approach and 3 in gel-free approach) for each sample group.

Leishmaniasis researches are mostly carried out by using defined model strains whose genetic material do not change however recent findings show us these parasites have the ability to exchange their genetic information selectively. This feature enables them to adapt to more severe conditions and survive in unusual habitats. As mentioned earlier, natural hybrid parasites and drug-resistant strains are explicit indicators for this notion. Therefore, investigations on newly-emerging strains instead of stationary strains in the laboratory should be increased to keep up with the upgrades in the genetical changes of leishmania parasites. Accordingly, usage of clinical isolates from different patients is one of the main strengths of this study. In order to take unbiased decision, equal treatment was applied to all samples and only common proteins were focused.

2.2. Experimental Methods

In the gel-based approach part, 15 promastigote isolates obtained from different patients were augmented to obtain enough protein extract. 9 samples among these 15 were split into two parts to consume in the gel-free approach. Total protein solutions of 15 samples used in gel-based approach were first separated by 2D-SDS polyacrylamide gel electrophoresis method and proteins of interest were later digested with trypsin enzyme. On the other side, total protein solutions of 9 samples used in gel-free approach which was discussed in chapter 3 were digested without separating proteins beforehand.

2.2.1. Growth of *Leishmania* isolates and Total Protein Extraction

Protein samples used in this study were categorized into three classes and coded as VL(I), CL(I) and CL(T) apart from labeling for each samples. CL(I) and CL(T) refer to isolates obtained from patients with cutaneous leishmaniasis caused by *Leishmania infantum* and *Leishmania tropica* respectively. Similarly VL(I) corresponds to visceral type leishmaniasis whose isolates were obtained from patients infected by *Leishmania infantum*. All parasite cultivation, isolation and total protein extraction steps were carried out by Prof. Dr. Ahmet Özbilgin's group in Celal Bayar University, Faculty of Medicine, Department of Parasitology. Frozen promastigote cells whose cryopreservation was maintained by liquid nitrogen were first cultivated in enriched NNN (Novy-MCNeal-Nicolle) medium at 25 $^{\circ}$C and later transferred to RPMI-1640 medium including 10% FCS, 200 U of penicillin/ml, and 0.2 mg of streptomycin/m. When promastigote population reached to logarithmic phase with approximately 10^8 cell/ml concentration, growth medium was centrifuged at 4400 RPM for 10 minutes at 4 $^{\circ}$C in a clean 15 ml tube. After centrifugation, pellets were incubated in RPMI 1640 medium without FCS overnight. Finally, promastigotes were harvested again by centrifugation at 4400 rpm. Washing with PBS (Phosphate-buffered saline) buffer was repeated three times to remove the medium thoroughly. Cell lysis was carried out by adding 1 ml of Mammalian Cell Lysis Reagent (Fermentas Life Sciences). After the incubation at room temperature on a shaker at 900 rpm, total protein extract was collected by centrifuging at 14500 rpm for 15 minutes at 4 $^{\circ}$C. The supernatant was stored at -20 $^{\circ}$C for the upcoming proteomic analysis.

Table 2.1. Patient records for some samples

Sample codes	Province	Age	Sex	Lesion location	Lesion duration	Lesion Type	Treatment	Genotyping Results
C259	Adana	17	M	Forehead	36	Dry	No	*L. infantum*
C615	Diyarbakır	18	W	Face	1	Recidivant	Yes	*L. tropica*
C628	Manisa	50	M	Liver	6	VL	No	*L. infantum*
C285	Hatay	28	M	Knee	24	Dry	No	*L. infantum*
C495	Manisa	30	M	Liver	3	VL	No	*L. infantum*
C121	Osmaniye	26	M	Arm	18	Dry	Yes	*L. infantum*
C239	Manisa	57	M	Liver	6	VL	No	*L. infantum*
C240	Manisa	9	M	Liver	6	VL	No	*L. infantum*
C250	Osmaniye	8	M	Lower lip	36	Dry	Yes	*L. infantum*
2FA	Urfa	3	M	Nose	5	Dry	No	*L.tropica*

2.2.2. Protein Desalting/Concentration and Quantitation

Desalting and buffer exchange processes provide both sample clean-up and concentration. Acetone precipitation of proteins and filtration with disposable centrifugal devices are two different examples of those upstream sample preparation step. In acetone precipitation, 4 volumes of -20 ˚C acetone is mixed with 1 volume of sample and incubated in -20 ˚C for 3-4 hours. After centrifuging at 14000 rpm at 4 ˚C supernatant is decanted without disturbing the protein pellet at the bottom. Protein is then dissolved with appropriate buffer solution following the air-drying of acetone. Drawbacks of this method are possible chemical artifacts (e.g. +98 u) and insoluable pellet formation due to over-drying. In protein filtration, liquid sample is filtered from molecular weight cut-off (MWCO) membrane by centrifugal force. To determine the protein concentration method among these two, we performed a comparative analysis relying on 2D-SDS-PAGE. As can be seen on figure 2.3, there is not significant difference between the gel images of same sample with same amount therefore 10K Da cut-off filter was preferred due to its easier application. In addition, cut-off filters were also used to remove rehydration buffer of 2D PAGE which contains high amount urea for in-solution trypsin digestion that is the protein cleavage step of shotgun proteomics in chapter 3.

400µL volume from each sample was centrifuged with 10K Da MWCO filter devices (Amicon Ultra-0,5 mL; Millipore) at 4 ˚C 14000 rpm for 15 minutes. Total sample volume was approaximately 800µL therefore this step was repeated twice for the rest of the sample. Concentrated 20-30µL protein slurry was dissolved with 350µL rehydration buffer and transferred into a clean 1.5 mL ultra high recovery microcentrifuge tube (STARLAB). Another 350 µL rehydration buffer was added to

collect the remaining sample and added to previous tube. Thus 700 µL total protein sample in rehydration buffer was obtained. Cut-off filters were washed with 500mL three times to reuse in tryptic in-solution digestion. Rehydration buffer used in this stdy involves 7 M urea, 2 M thiourea, 4% CHAPS, 65 mM DTT (Dithiothreitol) and 2% ampholyte solution for pH 3-10 range. DTT and ampholyte were added just before EIF.

For protein quantitation with Bradford assay, 3 fold dilution was necessary. 8µL of protein solution was mixed with 16µL rehydration buffer. Reaction was taken place in 96-well plate according to microplate procedure of Coomassie Plus Assay Kit; Thermo. Briefly, 290µL coomassie reagent, 10µL rehydration buffer and 10µL BSA solution of different concentrations were added to draw the graph of standard curve and 290µL coomassie reagent, 10µL H_2O and 10µL sample were added for the unknown samples.

2.2.3. Two Dimensions: Isoelectric Focusing and Gel Electrophoresis

All protein samples were adjusted to an amount of 400 µg in 330µL volume by adding the calculated volume of rehydration buffer. Before the deposition onto IPG strips, 20 µL rehydration buffer containing 3 mg DTT and 9µL ampholyte were added to each sample. Non-linear pH 3-10, 17 cm IPG strip (ReadyStrip™, Bio-Rad) was used for the isoelectric focusing step (Protean IEF cell, Bio-Rad) which constitutes the first dimension in 2D-SDS PAGE method. Active rehydration was started at 50 Volts for 16 hours. Next day, before the initiation of real separation program, wetted paper wicks were placed between electrode wire and strips on both sides. It is recommended to use paper wicks for the absorption of low-molecular-weight components such as salts and proteins that are out of the pH range. Finally the predetermined IEF program was set and initiated. After the completion of program, strips to be subjected to second dimension with gel electrophoresis were conditioned by the incubation in equilibration buffer-1 and equilibration buffer-2. Equilibration buffer-1 contains dithiothreitol (DTT), a reducing agent which cleaves the disulfide bonds by reducing sulfhydryl groups. Equilibration buffer-2 has the same content as equilibration buffer-1 except DTT instead an alkylating agent, iodoacetamide is added to buffer to prevent the reformation of cysteine disulfide bonds. After 15 minutes incubation on horizontal shaker at 200 rpm, equilibration buffer-1 is decanted and equilibration buffer-2 for another 15

minutes. Since iodoacetamide is sensitive to light, tray is covered to keep strips in dark. After the equilibration step, strip was washed by immersing into running buffer tris-glycine-sds (TGS). Later, strip was placed carefully onto the polyacrylamide gel that was casted before and covered with overlay agarose. Resolving gel for Tris-glycine SDS-Polyacrylamide gel electrophorsis are mostly prepared either %12 or %15 percents of acrylamide and we chosed 12% of acrylamide gel. In the Appendix A, solution components used in the preparation of SDS-polyacrylamide gel and EIF program are given in detail. Finally, proteins were separated according to their molecular weight under the electrical field which was run 16 mA for one hour and 8-9 hours at 180 Volt.

2.2.4. Gel Stainings: Coomassie Staining and Silver Staining

CBB stainin provides good quantitative linearity although it is not as sensetive as silver staining. Colloidal state offers higher sensetivity and reproducibility and clearer background since it does not react with the gel. Gel staining with Coomassie Brilliant Blue G-250 is pretty straightforward compared to silver staining. Briefly, as aluminum sulfate is dissolved in water, ethanol and CBB-G250 are added respectively after the homogenization each time. When the solution is completely dissolved phosphoric acid is mixed causing the coomassie molecules to aggregate to colloids. Finally, it is made up with water. (0.02%w/v CBB-G250, 5%w/v aluminum sulfate-(14-18)-hydrate, 10% v/v ethanol, 2%v/v orthophosphoric acid (85%)). Destaining of CBB is simple too that gels are only kept in ethanol (10%v/v) and orthophosphoric acid (2%v/v) for awhile. In some protocols, aluminum sulfate is substituted with ammonium sulfate and ethnaol is substitute with methanol. The detection limit of this method is 1–5 ng of protein per spot or band. Silver staining is based on the affinity of the proteins for a cation, here silver. Other substances, like SDS, which also have a high affinity for silver, must be removed prior to staining. This means that a fixation step is required. The sensitivity of the staining reaction can be enhanced by the use of sensitizers, here sodium thiosulphate, between fixation and silver impregnation. 100 ml of silver solution for a big gel (20 x 20 cm) is necessary. A suitable amount for the rest of the solutions is about 200 ml. Formaldehyde or glutaraldehyde should not be used in the fixation or silver solutions because this will fix the proteins permanently in the gel. Below is the

solutions used in silver staining protocol. It should bnot be forgotten that sensitization solution, silver solution and development solution should be prepared freshly.

- Fixation solution: 300ml ethanol and 100ml acetic acid is added to 600ml water.
- Sensitization solution: 300mg sodium thiosulphate pentahydrate in 1L water.
- Silver solution: 2g silver nitrate is dissolved in 1L water.
- Development solution: 30 g potassium carbonate and 15 mg sodium thiosulphate pentahydrate is dissolved in 1L water. 700 µl of 37% formaldehyde is added.
- Stop solution: 50 g Tris base is dissolved in 1L of 2.5% acetic acid in water.
- Silver destaining solution: 30 mM potassium ferricyanide is mixed 1:1 with 100mM sodium thiosulfate.

After the gel electrophoresis run, gel is placed in a glass or plastic tray in fixation solution for minimum 1 hour and maximum 18 hours on a platform shaker. After the fixation step, gel is rinsed with 20% ethanol for 15 minutes and with water for 15 minutes. Then gel is sensitized with sodium thiosulphate for exactly 1,5 minutes on the platform shaher. Next, gel is rinsed with plenty of ultra pure water for 20 seconds twice. Gel is stained with silver solution for 30 minutes on the platform shaker. After rinsing gel for 20 seconds with water, development solution is added while shaking on the platform shaker until the spots almost have desired intensity. This usually takes 2-5 minutes. Finally stop solution is added to stop development and waited for two minutes. Then gel is rinsed with water twice for 10 minutes. The gel can be stored in water for several days. Before starting the in gel-digestion procedure, gels are destained with destaining solution until brown color disappears and then they are washed with ultrapure water at least five times.

2.2.5. In-gel Digestion of Protein Spots

This protocol takes two days or more depending on the removal of the dye. First protein bands are cut from the gel as closely as possible with a sharp scalpel, and divided into smaller pieces that are approximately $1mm^3$ to 2 mm^3. Gel pieces are placed in a new plastic microcentrifuge tubes and 200 μL of the wash solution (50% (v/v) methanol and 5% (v/v) acetic acid) is added and they are rinsed overnight at room temperature. If desired, this washing step can be carried out over the weekend or, alternatively, for 4 h. Reagents of in-gel digestion protocol are as follows.

- Wash solution: 50% (v/v) methanol and 5% (v/v) acetic acid.
- 100 mM ammonium bicarbonate in water:
- 50 mM ammonium bicarbonate in water
- 10 mM DTT in 100 mM ammonium bicarbonate
- 100 mM iodoacetamide in 100 mM ammonium bicarbonate
- Extraction buffer: 50% (v/v) acetonitrile and 5% (v/v) formic acid.
- Trypsin solution is prepared by adding 1.0 mL of ice cold 50 mM ammonium bicarbonate to 20 μg of sequencing-grade modified trypsin. The final concentration is 20 ng/μL trypsin.

On the second day, the wash solution is removed from the sample with a plastic pipette and discarded. Then 200 μL of acetonitrile was added to dehydrate the gel pieces for ~5 min at room temperature. When dehydrated, the gel pieces will have an opaque white color and will be significantly smaller in size. Carefully the acetonitrile is removed from the sample with a plastic pipette. Completely the gel pieces are dried at ambient temperature in a vacuum centrifuge for 2 to 3 minutes. Then 30 μL of 10 mM DTT is added and proteins are reduced for 0.5 hour at room temperature. DTT solution is removed from the sample carefully and 30 μL of 100 mM iodoacetamide is added to alkylate the protein at room temperature for 0.5 hour. After 30 minutes the iodoacetamide solution is removed from the sample with a plastic pipette carefully. Next 200 μL of acetonitrile is added to dehydrate the gel pieces for ~5 min at room temperature. Acetonitrile is removed from the sample with a plastic pipette again. Once more the gel pieces are rehydrated in 200 μL of 100 mM ammonium bicarbonate for 10 minutes at room temperature. After that the gel pieces are dehydrated last time with acetonitrile and then they are dried at ambient temperature in a vacuum centrifuge for 2 to 3 minutes. Finally 30 μL of the trypsin solution is added to the sample and the gel pieces are allowed to rehydrate for 10 minutes with occasional vortex mixing. 5 μL of 50 mM ammonium bicarbonate is added to the sample to cover the gels when necessary. The sample is drove to the bottom of the tube by centrifuging for 30 sec, and the digestion is carried out overnight at 37 °C.

On the third day peptides produced by the digestion are extracted in three steps. First 30 μL of 50 mM ammonium bicarbonate is added to the digest and incubated fr 10 minutes with occasional gentle vortex mixing. The digest is drove to the bottom of the tube by centrifuging the sample for 30 second. Then 30 μL of the extraction buffer is

added to the tube containing the gel pieces and incubated in the sample for 10 minutes with occasional gentle vortex mixing. The extract is drove to the bottom of the tube by centrifuging the sample for 30 seconds. Supernate is collected carefully with a plastic pipette and combined in the 0.5-mL plastic microcentrifuge tube. A second 30-μL aliquot of the extraction buffer is added to the sample and last step is repeated. The volume of the extract is reduced to < 20 μL by evaporation in a vacuum centrifuge at ambient temperature. The volume of the digest is adjusted ~ 20 μL, as need, with 1% acetic acid. At this point, the sample is ready for analysis however a final step for the salt remove is needed and. C18 ZipTip™ from Millipore Corporation, a 10 µL pipette tip with a bed of chromatography media fixed at its end is mostly used for this purpose. To equilibrate the ZipTip pipette tip for sample binding, maximum volume setting of 10 µL 0,1% TFA in water is aspirated as wetting solution and dispensed to waste three times. After equilibrating the tip, peptides are bound by fully depressing the pipette plunger to a dead stop. The sample is aspirated and dispensed 7-10 cycles without dropping the sample to the waste. Then pipette is again washed with wash solution at least twice. 1 to 4 µL of elution solution which contains 50% acetonitrile in 0,1% TFA is dispensed into a clean vial.

2.2.6. MS/MS Analysis and Protein Identification

The matrix used for the peptides is alpha cyano-4-hydroxycinnamic acid (CHCA) with a two-layer MALDI deposition method. This involves the deposition on the MALDI target of a microcrystalline matrix layer via fast evaporation from 0,6-1 μL solution of CHCA (10 mg/ml) dissolved in 20% methanol in acetone. A 1 μL aliquot of the digested protein sample was mixed with either 2 or 4 of saturated CHCA solution in 40% methanol in 0,1% TFA-water. The peptide-matrix solution was vortexed and 1 μL portion was deposited on top of the first matrix layer. Spectra are processed and analyzed by Autoflex III Smartbeam (Bruker) which uses internal MASCOT software (Matrix Science, London,UK) for searching the MS/MS data. This type of MALDI TOF/TOF works with programs flexControl 3,0 and flexAnalysis 3,0. Processed data by flexControl is transferred to MASCOT software by another licensed program biotools 3.1. NCBIProt and Swiss-Prot protein sequence databases were used for searches under “other eukaryota” taxonomy . Other database search parameters were as follows:

carbamidomethylation (C) as fixed global modification allowance for up to one tryptic cleavage. The peptide mass tolerance was either 200 ppm or 1 Da and fragment ion mass tolerance was 0,5 Da. Charge state was 1+ and monoisotopic mass was chosen. Mass range of the analyses was set to 700-3500 Da.

2.3. Results and Discussion

In this study, both gel-based and gel-free proteomic approaches were utilized for the comparison of different clinical forms in leismaniasis disease. In the gel-based section, pI 3-10 IEF strips were chosen for the general comparison therefore all promastigote isolates obtained from 15 different patients were run on gel by 2D-SDS-PAGE. In addition, two gels from pI 4-7 strips and one gel from pI 5-8 strips were tested with remaining samples with lesser amount to estimate the degree of resolution in narrower pI ranges. Among these pI ranges, pI 3-10 strips exhibited the best coverage for the protein mixture with rather less separation resolution as expected. According to the gels of narrower range strips, pI 4-7 and pI 5-8 strips can be useful for deeper analyses of the targeted proteins whose pI were predetermined. Reproducibility of electrophoresis is quite notorious that certain conclusion can be attained after at least three replications for one sample but the fact remains that replication of several sample is laborious and uneconomic therefore we built the comparison on one repeat from fifteen different biological samples. No specialized tool or software were used for the spot picking thus visual comparison were carried out only by eye detection. Striking changes as newly appeared spots or conversion from pale to dark color and vice versa between visceral and cutaneous groups at least three times were mostly taken into consideration. In this context, focused areas from each gel are arranged in a row to depict the region of interest (ROI) for ease of seeing and the whole images of the gel and details about all identified proteins are given in Appendix B. Spots identified with in-gel digestion and MALDI-TOF/TOF MS with subsequent Mascot search are indicated with a number on the figures.

Before switching formerly used acetone protein precipitation method with 10K MWCO filtration, their performance were compared on the same sample with same amount. Judging by the close similarity between gel images, filtration method (figure 2.3, left column) was considered as good enough to be replaced with acetone

precipitation (figure 2.3, right column). In addition, our conventional gel staining with Coomassie Brilliant Blue G-250 (CBB) (figure 2.3, upper row) had to be changed with a more sensitive gel staining method, silver staining (figure 2.3, lower row) to gain more clarity.

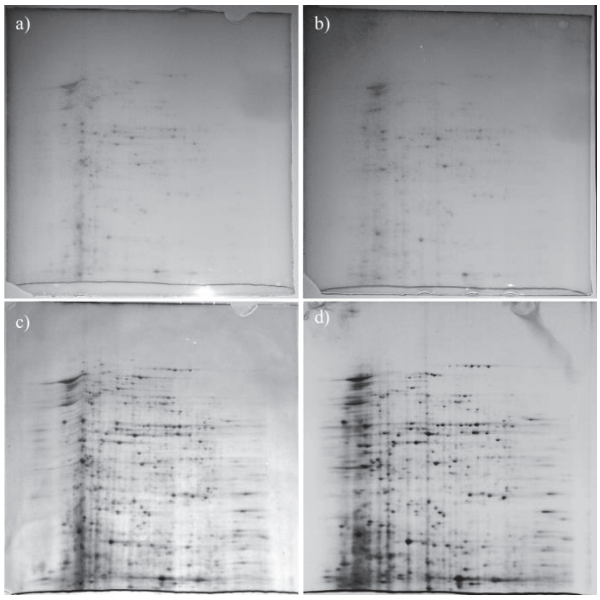


Figure 2.3. Gel images of protein concentration with filtration (a,c) and acetone precipitation (b,d) after CBB staining (a,b) and silver staining (c,d)

In the figures of ROIs (figure 2.4, figure 2.5, figure 2.6), samples of visceral leishmaniasis are laid on the top row while samples of cutaneous leishmaniasis from *Leishmania infantum* follow them in the middle and cutaneous leishmaniasis samples from *Leishmania tropica* are aligned down on the third row. Exceptionally, images from different pH ranges other than pH 3-10 are attached to figure 2.4, on the rightest column to incorporate the consistent results obtained from these gels.

The area of ROI(1) is very close to the bottom edge of the gel therefore proper images could not be obtained for some samples (CL121, VL239, CL-Crt02). Apart from that, gel images of the same sample, VL628 with pH 3-10 and pH 4-7 show that pH 3-10 range may not be sufficient to resolve proteins whose molecular weight and pI values are too close. Notwithstanding, this may not be always valid since this resolution were attained on some samples such as VL-Crt07, CL-Cri08, CL-259. Nevertheless, a new protein spot (CL-Crt03; spot #1) which was pointed by an arrow was only detected on the samples of cutaneous leishmaniasis either caused by *L. infantum* or *L.tropica* three times. Appearance of superoxide dismutase (SOD) protein as a noticeably new

spot solely on the gels of cutaneous type samples brought to mind that it might be an indication of a biomarker protein between visceral and cutaneous disease types however this protein was determined on a different location on the gels of visceral type samples (VL-Crt07; spot #4). Superoxide dismutase is a metalloenzyme which catalyzes the conversion of harmful oxygen species to molecular oxygen or hydrogen peroxide (H_2O_2) thus it takes places in detoxification process in the cell. As a reminder, reactive oxygen species are also produced against microorganism invasions in macrophages through the respiratory burst[82]. Although phagocyte NADPH oxidase and nitric oxide synthase enzymes take major role in this defense mechanism, SOD activity increases in parasite infected macrophages[83]. Among four antioxidant defence enzymes including catalase, glutathione peroxidase, superoxide dismutase and peroxidoxin, last two are expressed in Leishmania. Additionally, a recent research demonstrated that SOD deficiency cause to loss of virulence by rendering parasite more susceptible[84]. In the light of these, that SOD might be an implication for parasite transformation between the disease types and deserves further analyses to relate it to vicera tropism.

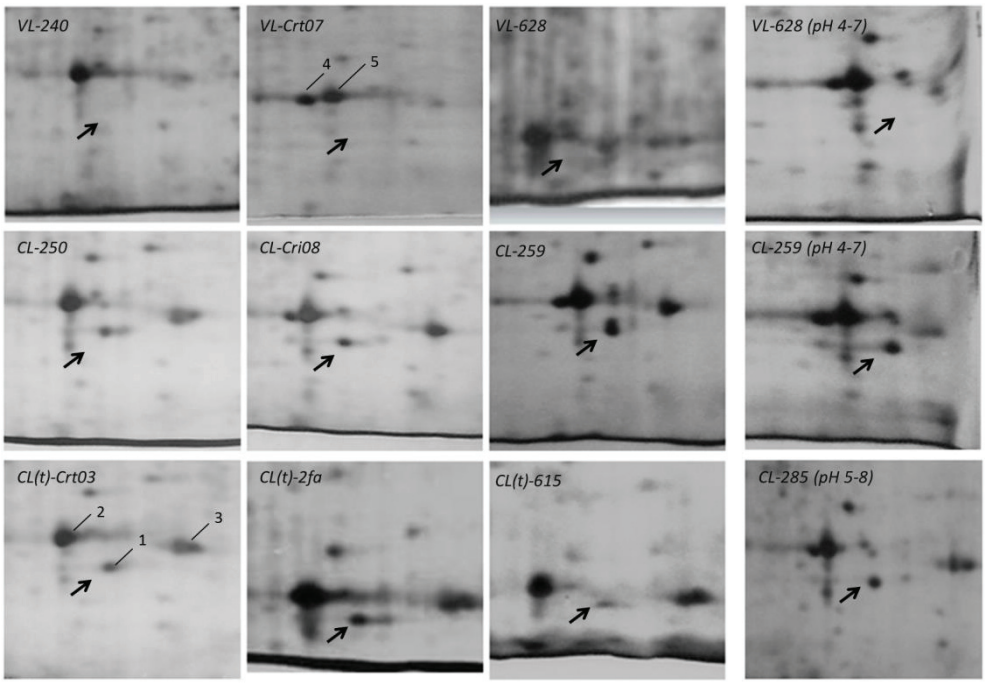


Figure 2.4. Focused images of region of interest, ROI (1)

In ROI (2), two collateral spots (Figure 2.5, #25 and # 26) on the direction of arrow can be detected clearly on all cutaneous samples while this location is either entirely empty or barely stained in visceral samples. Both of these spots were identified

as putative N-acyl-L-amino acid amidohydrolase however spot with #26 could only be determined by peptide mass fingerprinting. This enzyme exhibit aminoacylase activity by acting on carbon-nitrogen bonds other than peptide bonds. Three different accession codes, A4I6N8, A4I6N7, A4I5G9 correspond to N-acyl-L-amino acid amidohydrolase in Uniprot database. Genes of first two are on the same chromosome (ch_31) while gene of latter is located on another chromosome (ch_30). According to shotgun analysis, this protein is also expressed in the samples of visceral type however it is represented only with A4I6N7 which may be an explanation for down regulation. Although, changing expression levels of aminoacylase enzyme is mentioned in the literature, this protein is not in spotlight like SOD. Downregulation of aminoacylase along with several metabolic enzymes in metacyclic virulent form comparing to procyclic avirulent is concluded to conservation of energy from cell division[85]. On the other hand, it is demonstrated that gene expression of N-acyl-L-amino acid amidohydrolase increases in miltefosine resistant *L. donovani* according to genomic microarray analysis[86]. Additionally, upregulation of heat shock protein 83-1 (position of spot#33) and downregulation of enolase (EC 4.2.1.11) (position of spot#27) can be implied from color intensity. Lastly, protein namely metallo-peptidase family 32 was identified multiple times with spots having different locations (spots #22, #28, #30).

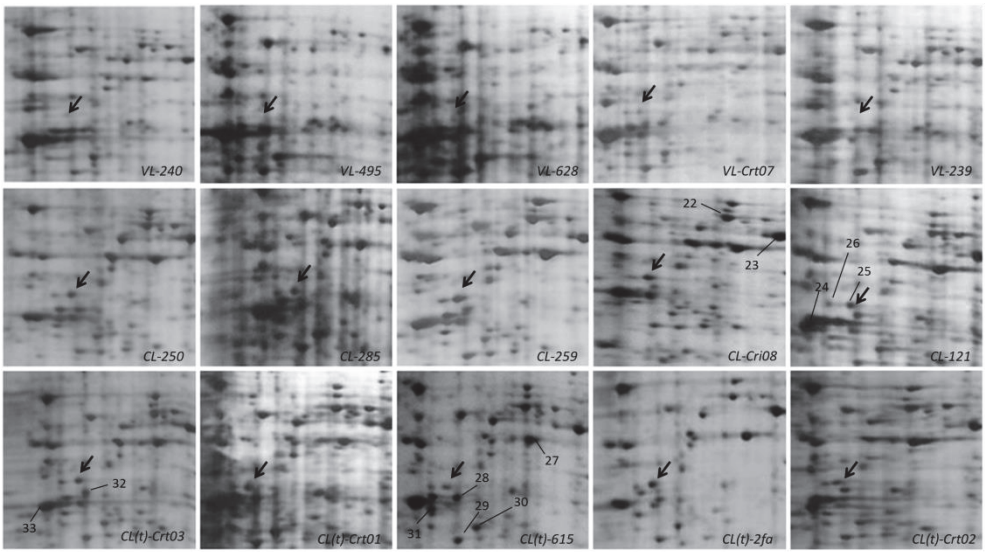


Figure 2.5. Focused images of region of interest, ROI (2)

Prominent advantage of gel-based method is enabling approximate information about those of isolated proteins' molecular weight and pI values however proteins

sometimes do not show this correlation. Metallo-peptidase clan MA(E) family 32 (CL-Cri08; #22, CL(t)-615; #28, #30) and heat shock protein 83-1 (CL(t)-Crt03; #33) are two proteins which can be counted as example to this inconsistency. Molecular weight of the latter is around 80 kilo Dalton and it is observed multiple times (CL(t)-613; #31 and Appendix B; #34, #35, #36) on the gels however none of those spots correspond to 80 kilo Dalton. Likewise spots of protein (CL(t)-615; spot#28, spot#30) namely metallo-peptidase do not match with the molecular weight except one (CL-Cri08; spot#22). This undesirable situation is actually common and may be arisen from modifications due to intracellular activities or experimental applications during protein extraction and electrophoresis. Nonetheless, if the spot is identified with a reliable match, expression of that protein in the cell is accepted.

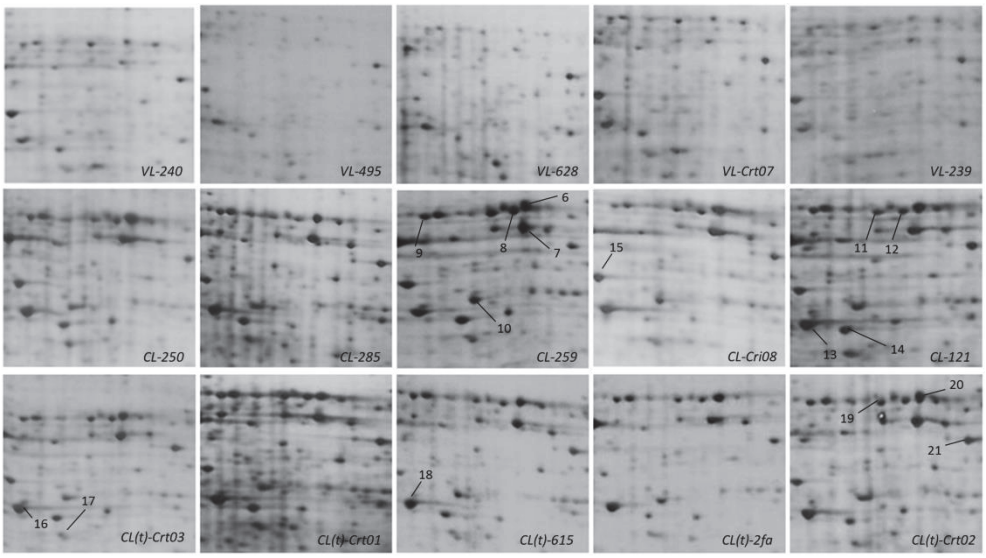


Figure 2.6. Focused images of region of interest, ROI (3)

When it comes to ROI(3), downregulation on particular spots can be observed explicitly (figure 2.6). Color intensity and size of spot#8, spot#6 and spot#7 (CL-259) which are identified as putative dihydrolipoamide dehydrogenase (EC 1.8.1.4), pyruvate kinase (EC 2.7.1.40) and putative succinyl-coA:3-ketoacid-coenzyme A transferase (EC 2.8.3.5) respectively decrease by far, especially in VL-495 and VL-628 samples. These proteins seem directly or indirectly related with the production of a crucial metabolite, acetyl-CoA (figure 2.7). In fact, pyruvate kinase and enolase take role in the steps of glycolysis, a series of reactions in which glucose is converted to pyruvate. Released free energy from this metabolic pathway is used in ATP gain. Much more ATP is attained

from citrate cycle and pyruvate enters this cycle through acetyl-CoA. In addition, acetyl-CoA molecule is released and participated in many other biochemical reactions of carbohydrate, protein and lipid metabolism. On the other hand, although dihydrolipoamide dehydrogenase enzyme performs quite a simple reaction of lipoamide conversion from dihydrolipoamide, it participates to enzyme complexes including pyruvate dehydrogenase, α-ketoglutarate dehydrogenase, branched-chain oxoacid dehydrogenase and glycine decarboxylase because lipoic acid is an essential cofactor for those enzymes. Subunits of α-ketoglutarate dehydrogenase complex involve citrate cycle enzymes which are required to produce metabolite succinly CoA from 2-oxoglutarate. In the light of information above, probably parasites of viseral leishmaniasis regulated their system according to lesser energy demand which implies the avoidance of rapid growth. Apart from that, aspartate aminotransferase protein designated with spot#10 (CL-259) is notably downregulated in visceral isolates. Other downregulations were observed on putative aldose 1-epimerase (CL-121; spot#13, spot#14) protein and trypanothione reductase protein (CL-259; spot#9).

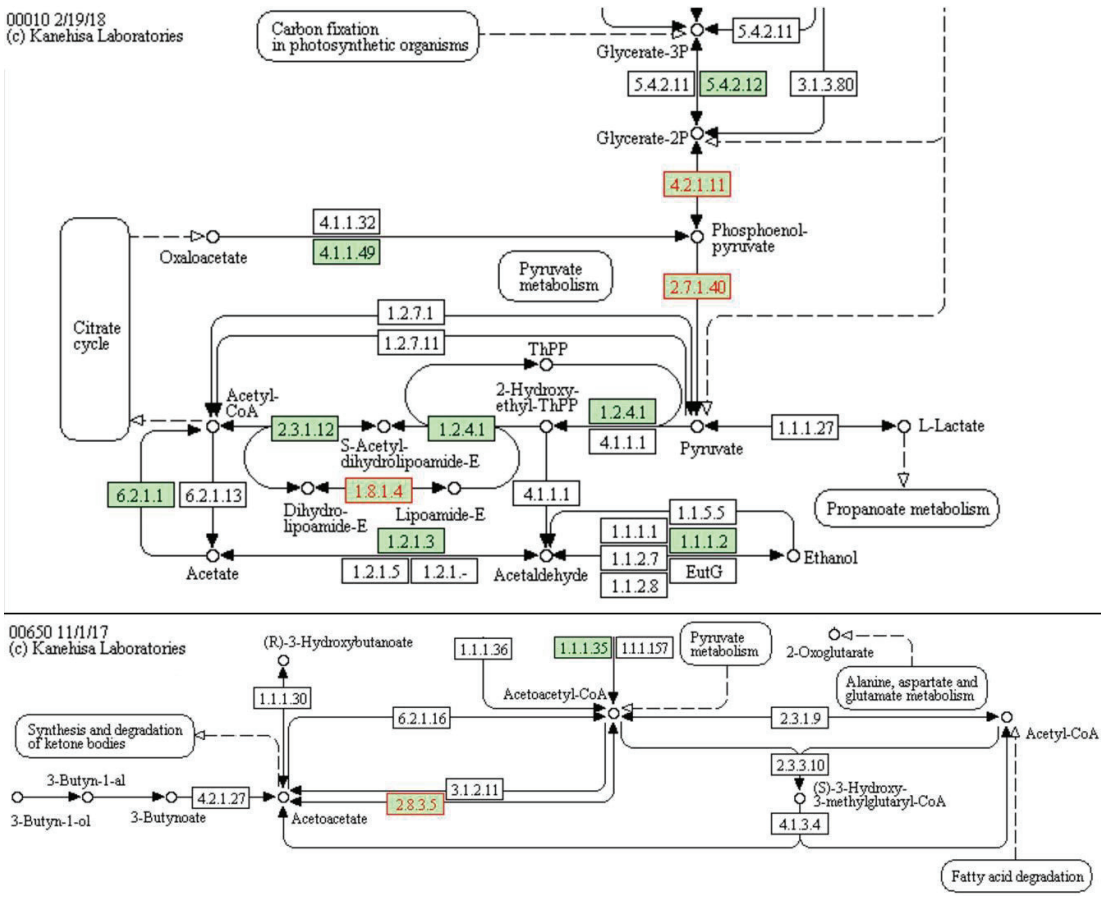


Figure 2.7. Detected proteins with their EC numbers are written in red on partial pathways of Glycolysis/Gluconeogenesis (up) and Butanoate metabolism

To the best of our knowledge, there is only one study which compares protein expressions of viseral and cutaneous samples originated from same species. In this study, promastigote isolates corresponding to cutaneous and visceral leishmaniasis were both obtained from a dog which was infected by *Leishmania tropica* in Iran. According to their results obtained from 2D-GE with pH 4-7 and LC-MS/MS analysis, more proteins (77) were downregulated in viseral samples compared to cutaneous samples (30)[87]. Although, our results do not correlate with regards to identified proteins in this study, we both reached to the point that parasites of visceral leishmaniasis refrain from heavy protein synthesis compared to parasite of cutaneous leishmaniasis. Apart from that, downregulation of dihydrolipoamide dehydrogenase along with pyruvate dehydrogenase (EC 1.2.4.1) and isocitrate dehydrogenase (a citrate cycle enzyme) were demonstrated in mitochondrial proteome of drug resistant *L.infantum* parasites (both antimony and miltefosine; *in vitro*)[88]. However succinly-CoA:3-ketoacid-coenzyme A transferase levels were higher than that of wild type in contrast with our results. Nevertheless, some differential proteins obtained from both gel-free and gel-based methods are interestingly encountered in Leishmania researches regarding drug resistancy therefore possibility of a connection between these two acquired mechanisms should be taken into consideration since both are related with adaptation of the parasite to harsh conditions.

2.4. Conclusion

Leishmaniasis is the name of an infectious disease that is caused by a unicellular parasite and spread through the bite of vector, sandfly. Main manifestation is the emerging of unpleasant scars on the skin however it may cause deaths when invasive parasites grow inside organs. These two clinical form are defined as cutaneous leishmaniasis when the outer parts of the host's body are affected and likewise visceral leishmaniasis when the inner parts are the target of the parasites. This classification is generally related with the species of the leishmania parasite however some species can cause both clinical forms. *Leishmania infantum* is among them, yet the majority of the cases are resulted visceral leishmaniasis. Similar trait is observed reversely in the species, *Leishmania tropica* which means it is responsible for cutaneous leishmaniasis to a high degree, yet rare cases of visceral leishmaniasis caused by *L.tropica* is

encountered. In the last decade, reports of such unusual cases have increased tremendously. Furthermore, same trend started to appear on other species too. In our country, these two species, *L.infantum* and *L.tropica* are mostly detected in the leishmaniasis cases however they influence west side and east side respectively. The reason why some species cause to cutaneous infections while others do fatal visceral disease has alway been wondered however this discrimination seems to be losing as mention earlier by the increase of unusual cases. Diminishing of the connection between species and disease forms results with complications in diagnosis and treatment. For this reason, mechanism underlying the habitat selection of parasite on the host's body and viscera tropism should be elucidated to take the control on this public healt problem as well as to unravel the mystery of progress in leishmania parasites. Comparative genetic analysis on different leishmania species do not provide sufficient information since the synteny is highly conserved on the other side proteomic and metabolomic studies can reveal differential profiles despite the similarity between genomes.

In this study, comparative proteomic analysis of leishmania isolates obtained from visceral and cutaneous leishmaniasis is carried out by using gel-based and gel-free proteomic approaches. Results of the gel-based approach are shared and discussed in this chapter and those of gel-free proteomic approach is given in the following chapter. One of the outstanding observation of this gel-based comparative proteomic analysis was the explicit existance of a spot found only in the samples of cutaneous leishmaniasis either caused by *L.infantum* or *L. tropica*. MS and MS/MS spectra belonging to the peptides of this spot which were measured by MALDI-TOF/TOF and searched with Mascot, matched to super oxide dismutase protein (SOD). Same protein was detected in a sample of visceral leishmaniasis however it was located to different position. Morever, results of gel-free approach which was discussed in chapter 3 also support that all three visceral leishmaniasis samples contain SOD protein. In the Leishmaniasis researches, importance of SOD for parasite survival have taken attention for a long time and its immune response have being investigated to consider as drug or vaccine candidate[89-91]. In this regard, we estimate that protein composition of SOD protein is affected during conversion from either one leishmaniasis form to the other. This protein holds the potential to be a biomarker in the protein blotting for the parasite of cutaneous leishmaniasis if those regions generating the differentiation in the protein can be determined and a custom antibody can be produced for that specific regions.

Another differentiation between visceral and cutaneous leishmaniasis were observed on protein expression levels of many proteins in terms of relative quantity which was reflected on the gels as spot size and intensity. Among them, expression of putative dihydrolipoamide dehydrogenase, pyruvate kinase, enolase, putative succinyl-coA:3-ketoacid-coenzyme A transferase, and N-acyl-L-amino acid amidohydrolase were distinguishably higher in the samples of cutaneous leishmaniasis than those of all five visceral leishmaniasis samples. These proteins are member of energy methabolism except the last one. Since production of energy is the indication of cell proliferation, parasites causing visceral leishmaniasis probably refrain from rapid propagation not to disseminate their signals of existence to the defense mechanism of the host. Apart from that, visceral leishmaniasis parasites do not seem to compromise heat shock proteins 83-1 even they are up-regulated in some of the samples to overcome the stress conditions.

During the comparison of the gels, we have concentrated on readily detectable protein spot variations that were consistent in majority of the samples in a group. Polyacrylamide gel electrophoresis (2D-SDS PAGE) analysis revealed that protein profilings of cutaneous leishmania isolates belonging to different species of leishmania (*L. infantum* and *L. tropica*) are indistinguishably similar wherease those of visceral leishmania isolates are remarkably different from cutaneous leishmania samples. Reduction in protein divesity and downregulations of many proteins which imply the limitation of protein biosynthesis in parasites causing visceral leishmaniasis is consistent with the results in literature regarding isolates from a dog infected by *L.tropica* although no correlation was found with regards to identified proteins[87].

CHAPTER 3

GEL-FREE COMPARATIVE PROTEOMIC ANALYSIS OF LEISHMANIA ISOLATES OBTAINED FROM VISCERAL AND CUTANEOUS LEISHMANIASIS PATIENTS

3.1. *Leishmania* Parasite

Leishmania is a single cell parasite belonging to the class kinetoplastida which encompasses flagellated protozoans of the phylum euglenozoa. These organisms are recognized by the presence of bizarre DNA namely kinetoplast in their single large mitochondrion. Kinetoplastida is divided into two subgroups as bonodina and trypanosomatida according to the number of flagella. Another distinguished feature is unique compartmentation of some glycolytic pathway enzymes in the organalle called glycosomes. Members of trypanosomatida group has single flagella and patogenic to vertebrates including humans. For instance, other species of trypanosomatida, *Trypanosoma brucei* and *Trypanosoma cruzi* cause abovementioned African sleeping sickness and chagas disease respectively. Lastly, these parasites undergo morphological alterations during their life cycles in order to adapt to the new environments in the host.

Life cycle of leishmania begins with a bite of infected female sandfly thus metacyclic promastigotes are released and enter the host from the bite site. Mostly parasites invade macrophages in the skin although they can attach to the surface of different cell types such as neutrophil, dendritic cell and fibroblast. During phagocytosis, a vacuole forms around the parasite as the pseudopod encloses the parasite and herein parasite transforms from promastigote form to amastigote. Later, lysosomes of host cell infuse with parasite vacuole however it does not damage the parasite. As the amastigotes accumulate by multiple divisions, vacuole starts to occupy larger area in the cytoplasm and approaches to the cell membrane. At last, enlargement of vacuole culminates with the burst of the cell thus amastigotes reach the extracellular space and blood. Distribution of released amastigotes in the bloodstream continues in the body and it is transmitted if another sandfly sucks this blood. In the system of

sandfly, amastigotes pass through the throat to abdominal midgut. In this new environment they transform into procyclic promastigote and adhere to midgut epithelium by their flagellum. Some portion of parasite is taken out with blood meal and digested. When the parasite detached from the epithelium, promastigotes migrate towards the stoma. Thereby parasites transform to a more infective paramastigote form, metacyclic paramastigote whose cell body is smaller in size and flagellum is longer and more vibrant (Figure 3.1).

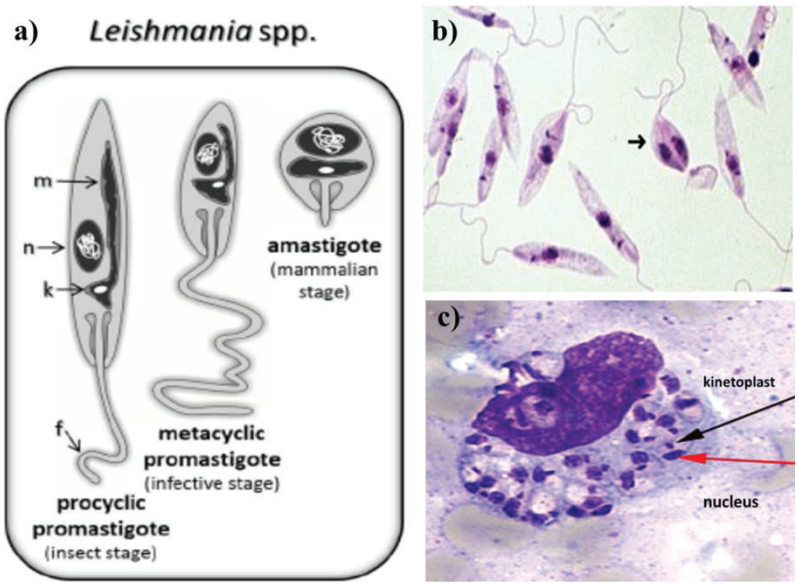


Figure 3.1. a) Stages of Leishmaniasis spp. in life cycle[92], Giemsa-stained parasites under the microscope b) Promastigote[93], c) Amastigote[94]

Early proteomic studies on leishmanial parasites are usually carried out to determine the proteins regulating the differentiation between the stages[95]. These proteins correspond to the response mechanisms in the parasite for the adaptation to the new acidic environment of phogolysosome which are necessary for intracellular survival of amastigote forms. To eliminate the factors originating from the host cell, transformation of the parasite to amastigote forms is employed axenically through the pH and temperature controls in vitro thus time-based observation become possible although it ends up with the lack of virulence ability. Since sugar and amino acids are rich in the midgut of sandfly, promastigotes are inclined to glycolysis whereas amastigotes tend toward gluconeogenesis for the glucose and prefer higher TCA and oxidative phosphorylation by exploiting the metabolites from amino acid and fatty acid oxidation. Therefore enzymes of cytosolic glycolysis including phosphoglycerate

mutase, enolase and pyruvate kinase down-regulate as gluconeogenesis enzymes up-regulate. In addition, mitogen-activated protein kinases (MAPK) are other regulatory actors through the phosphorylation activity. Apart from these, elevated expression of oxidative stress enzymes, superoxide dismutase and thypanothione reductase accompany to amastigote differentiation[96-99]. By the way, thypanothione associated enzymes draw special attention in drug development as an exclusive target since this antioxidant comprised of two glutathione molecules that are attached together by a spermidine bound, is unique to kinetoplastida and is not synthesized in human[100, 101]. Comparative proteomic analyses of strains belonging to drug-susceptible and -resistant parasite and similarly strains with passive and aggressive virulences are also taken place excessively in the literature[102-107]. Additionally, several specialized topics have been investigated basing on the profiling of phosphoproteome, exoproteome, membrane proteome, glycoproteome and immunostimulatory proteome, especially to find drug targets effect on all species[108-114]. Although comparative genetic analysis does not provide much information about the divergence between the distinguishing parasites belonging to same the species obtained from the same source, proteomic and metabolomics analyses reveal the implication of gene expressions[115]. Contrary to expections, it was demonstrated that synteny between complete genomes of *Leishmania infantum*, *Leishmania braziliensis* and *Leishmania major* are highly mainted across these three species[74]. Furthermore, among all *Leishmania spp*., only *Leishmania braziliensis* is known to hold the RNA interference mechanism[116]. Despite the proximity within the gene locations of three different species, leishmaniasis clinical forms of visceral, cutaneous and muco-cutaneous are mostly appreared by the infection of those species respectively. For this reason, alterations in pathogenesis is attributed to protein expressions regulated with the translation or post-translational modification rather than gene expressions[117].

Regarding the comparison of clinical manifestations of leishmaniasis, we have encountered to only one research whose parasite source is a dog exhibiting visceral and cutaneous leishmaniasis simultaneously. They have carried out 2D-SDS PAGE method and mass spectrometry-based identification to the isolates of *Leishmania tropica* and determined 135 differentially expressed proteins. Their results demonstrate that proteins involved in protein folding, antioxidant defense and proteolysis are up-regulated in viceral type, on the other hand, amount of protein expressions related with energy metabolism, cell signaling and virulence decrease are decreased[87].

3.2. Experimental Methods

Since aliquotes of the same sample were used, protein extraction procedure was explained earlier in the experimental section of the chapter 3 under the titles of "*2.2.1 Growth of Leishmania isolates and Total Protein Extraction*" and "*3.2.2. Protein Concentration and Quantitation*"

3.2.1 Protein In-Solution Digestion

As mentioned before, molecular weight cut-off (MWCO) filter was used to remove excess urea and detergent which interfere with trypsin during the digestion process. In this filter-aided sample preparation method, 330µl rehydration buffer containing 400µg total protein was pipetted into specialized 10kDa MWCO filter tube. 5µl of 0,2M reducing agent Dithiothreitol (DTT) was added and incubated for 1 hour. Later, 20µl of 0,2M alkylating agent 2-Iodoacetamide (IAA) was added and incubated for 1 hour in the dark. Next, 20µl of 0,2M reducing agent Dithiothreitol (DTT) was added again and incubated for 1 hour to neutralize the excessive IAA. In order to remove the bulk solution containing urea, thiourea and CHAPS, sample were centrifuged at 14000 rpm for 10 minutes. Before the addition of trypsin enzyme, samples were washed with 200µl 50 mM Tris-HCl at pH 8,5 and centrifuged at 14000rpm for 10 minutes. Finally, 4µg trypsin (enzyme to protein: 1:100) was prepared in a solution containing Tris-HCl (50mM, pH 8,5) and 10% acetonitrile was added and incubated at 37°C overnight. To decrease the volume to 100 µl which is the maximum injection volume of HPLC sample loop, samples were evaporated by speed-vac concentrator.

3.2.2 Off-line HPLC and LC-MS/MS

After the digestion of proteins, we obtain a very complex mixture of peptide. When this complexity is reduced prior to mass spectrometric LC-MS/MS analysis, number of protein is elevated because more peptides can get the chance to be measured and fragmented. Therefore we applied a pre-separation by off-line HPLC method whose details are given in Table 3.1. To decrease the number of fractions collected in a 96 well

plate, six wells were pooled into a new 1,5 mL tube by skipping each time 15 fractions. In this way, we aimed to combine unlike peptides as possible. Later, all 16 fractions were completely evaporated. Cell volume of PDA detector was too large for a sensitive measurement therefore LC chromatograms were quite similar to each other and were not useful.

Table 3.1. Off-line HPLC conditions and parameters

HPLC System/ Software	SHIMADZU Prominence UFLC LC Solution Version 1.25 SP3	***LC Program:***	
Column	Reverse Phase C18-Teknokroma Mediterraneasea18 (25 cm x 0,46 cm; 5µm)	*Time*	*Mobile Phase B*
Column Temperature	40 ˚C	*5.00*	*0%*
Mobile Phase A	10 mM Ammonium formate-Water	*5.10*	*5%*
Mobile Phase B	10 mM Ammonium formate-Acetonitrile 90%	*40.00*	*25%*
Mobile Phase pH	A and B both adjusted to 10	*47.00*	*30%*
Injection Volume	100 µl	*54.00*	*40%*
Flow	0,5 mL/minute	*61.00*	*80%*
Fraction Collector	Sunchrom Micro Fraction Collector	*70.00*	*80%*
Fraction Duration	~35 second/well (loaded to 96 well plate)	*85.00*	*0%*

On the left side of Table 3.2, features of the system and the program for gradient flow elution of the LC-MS/MS analysis are given. PDA detector was not suitable for this analysis thus it was excluded from the HPLC system and outlet tubing from LC coulmn was directly connected to the inlet of electrospray. Data-dependent setting were entered as they are given on the right side of Table 3.2, other parameters were left as default values.

Table 3.2. LC-MS/MS conditions and parameters

LC Conditions	**HPLC System**	DIONEX UltiMate 3000	**ESI**	**Spray Voltage (kV)**	5
	Software	Chromeleon Version 6.80		**Capillary Temperature (˚C)**	200
	LC principle	Reverse Phase		**Capillary Voltage (V)**	7
	Column brand	Sigma Supelco Ascentis ®	**Data-Dependent Settings for Automated MS/MS Data Acquisition**	**Acquire time (minutes)**	50
	Column size	15 cm x 500 µm; 2,7 µm		**Scan events (SE)**	10
	Column oven	25 ˚C		**MS Scan ranges (m/z)**	400-1800
	Mobile Phase A	0,1% Formic acid-Water		**Polarity**	Positive
	Mobile Phase B	0,1% Formic acid-Acetonitrile		**Data type**	Centroid
	Injection Volume	10 µL		**Activation type**	CID
	Flow	5 µL/minute		**Default charge state**	2
LC Gradient Program	**Time**	**Mobil phase B**		**Isolation width (m/z)**	2.0
	0.00	2%		**Activation time Q (ms)**	30.00
	4.00	2%		**Dynamic Exclusion**	Enabled
	5.00	5%		**Repeat count**	1
	6.00	8%		**Repeat duration (s)**	30
	36.00	22%		**Exclusion list**	500
	39.00	35%		**Exclusion duration (s)**	30
	42.00	60%		**Exclusion mass width**	± 1.5
	45.00	90%		**Reject charge states**	1, 4&up
	50.00	90%		**Signal threshold (counts)**	10000
	55.00	2%		**Normalized collision energy**	35

16 fractions of each leishmania total protein digests were analysed from fraction 1 to 16 consecutively. When the first runs were completed, a second run for all fractions were carried out as the technical repeat. All collected data obtained from 16 LC-MS/MS analyses of the first run and 16 LC-MS/MS analyses of the second run were merged to obtain single file. Data collected at thermo's raw format were converted to mascot generic file format (mgf) by using the tools of Proteome Discoverer 1.4.

3.2.3. Peptide Database Search

Database searches with search engines Comet[118], MSAmanda[119], Myrimatch[120], MS-GF+[121] and xTandem[122] were run by using Search GUI[123] program which provides a graphical user interface for the configuring and running single or multiple search engines. Same data were run with in-house Mascot[124] server (version 2.3) installed in our laboratory [125]. Same database and search parameters were entered for all searches as the following: fixed modifications is carbomidomethyl, variable modification is oxidation, peptide tolerance is 1.0Da, MS/MS tolerance is 0.5Da, peptide charge is both 2+ and 3+, instrument is ESI-TRAP and experimental mass values are monoisotopic. The only difference between Mascot searches and other searches was the method used in the generation of decoy databases. While Mascot uses randomised protein sequences in Mascot 2.3. for the automatic decoy search, reversed protein sequences are appended to the real database in Search GUI. Data were searched against two different databases. Leishmania infantum database which contains 8,331 entries in October 2017 was downloaded from UniProt[126] website by writing "Leishmania infantum" to the the taxonomy section under supporting data (www.uniprot.org)[126]. As recommended in the database generation tutorial of CompOmics group[127], a list of common contaminants which was obtained from the Global Proteome Machine (GPM)[128] and Trypsin protein from pig: "P00761" were concatenated to the former leishmania infantum database thus the final file in FASTA format obtained 8,907 entries and 5,458,420 residues. An alternative database which involves other species was downloaded from NCBI database[129] website (https://www.ncbi.nlm.nih.gov/). This large database of leishmania genus that contains 113761 proteins (October 2017) is actually the accumulated entries of PDB[130] (Protein DataBank, 748), RefSeq[131] (NCBI's reference sequence, 48487), UniProtKB[126] (*UniProt Knowledgebase, 228),* DDBJ[132] (DNA Data Bank of Japan,

174), EMBL[133] (The European Molecular Biology Laboratory, 50090), GenBank[134] (NIH genetic sequence database, 14004) and PIR[135] (The Protein Information Resource, 42). Search results obtained from Search GUI were visualized and analyzed by using Peptide Shaker[136]. Differential and common proteins were determined by sorting the elements in the textual output as a list of intersections and uniques. For this purpose, an on-line free website[137] which is designed by VIB from Gent University to draw venn diagram was used. Finally possible protein-protein interactions were checked at string-db.org, a platform that provides information about protein-protein interactions.

3.3. Results and Discussion

In this chapter, results and discussion is divided into two parts. In the first part, performaces of bioinformatic tools (six different search engines) and importace of search parameters (peptide mass tolerance and MS/MS tolerance) were evaluated..

In the second part, differential proteins obtained from shotgun proteomic analysis are given as lists. FDR value is kept under 1% for all results. Herein, a consistency was observed as a justification for the applicability of label-free quantification however statistical validation and at least three technical replication are requisites for quantitative analysis and specific softwares as well. Lastly, we would like to add a notification about the accession to the results. Since one search result contains hunders of proteins and more than 50 searches were run in this study, it was not possible to put the database search results into hardcopy. Possibly softcopy of all raw data and Mascot search results will be available soon via submitting to internet (possibly to PRIDE Archieve) however full results of searches can be sent in excel if requested.

3.3.1.Comparison of the Search Results with Different Settings

Before searching all samples, we first performed several Mascot searches to assess how relaxing the error window will effect on results and how proteins are distributed. Mascot default setting for peptide tolerance and MS/MS tolerance are 1.2Da and 0.6Da respectively. If data is collected from a high resolution mass spectrometer such as orbitrap or FT-ICR, peptide tolerance can be set to a low ppm value (e.g. 10 ppm). As it can be seen on Figure 3.2, too tight or too loose ms tolerance should be

avoided not to jeopardize the accessible yield of the database search. Nevertheless choosing the smaller number will be better when the gap becomes negligible therefore 1.0 Da was determined for ms tolerance during the searches.

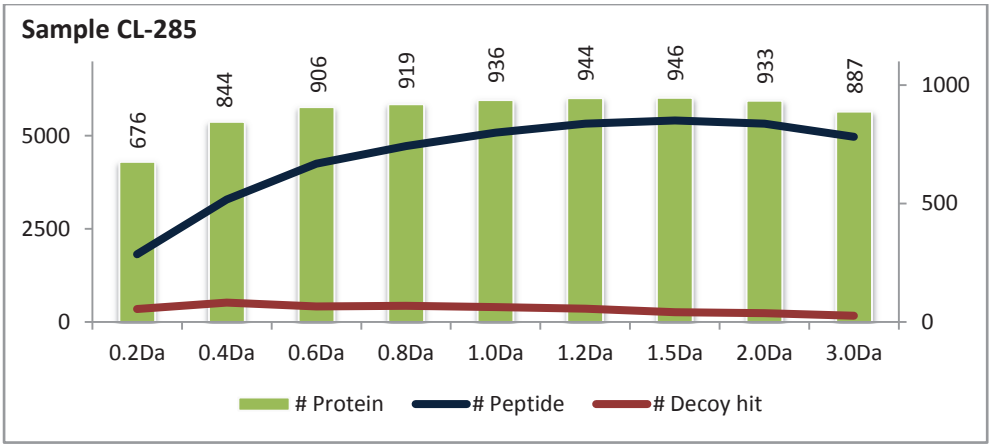


Figure 3.2. Effect of MS tolerance on quantitative results of Mascot search

From Figure 3.3, it can be conclude that change on ms/ms tolerance has more drastic effect on the search results than that of ms tolerance. Although searching with higher ms/ms tolerance gives considerably improved results, judging by the rise in decoy hit we observe that software's ability of matching the true sequence becomes questionable. According to this graph, 0.5 Da emerges as a good compromise considering the aimed "*highest the number of peptide and lowest the number of decoy*" restriction therefore it was assigned to ms/ms tolerance during searches of all samples. 0.6 Da was an accepable value but we did not want to risk degree of true matches.

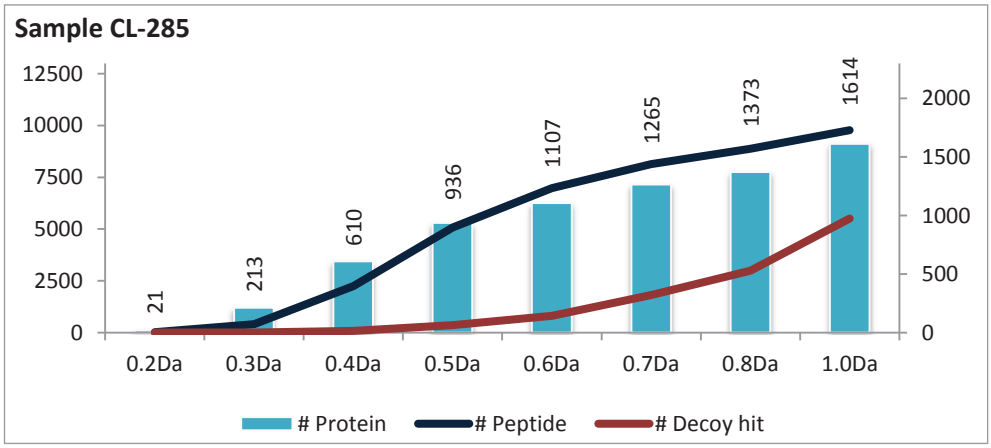


Figure 3.3. Effect of MS/MS tolerance on quantitative results of Mascot search

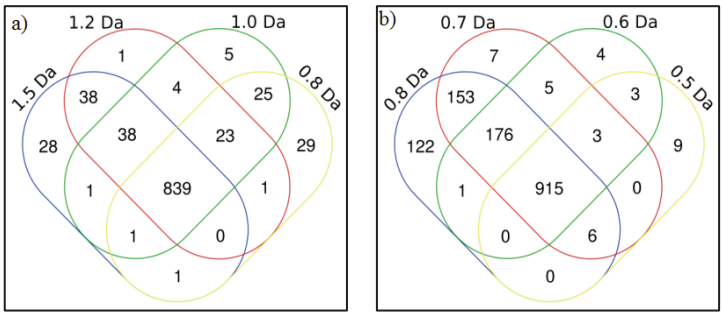


Figure 3.4. Distribution of common proteins in sample CL-285 obtained from Mascot searches with four MS tolerance (a) and four MS/MS tolerances settings (b)

Apart from the result based on protein numbers, particular four search results were compared with respect to protein identities. Although a great majority of proteins are shared between searches with different ms tolerance (Figure 3.4, /839) and ms/ms tolerances (Figure 3.4, /915), proteins in the other intersection sets cause inaccuracy for the search results. To elucidate this discrepancy, 29 proteins (Figure 3.4) which were identified only in the search with 0.8 Da ms tolerance were examined.

Table 3.3. CL-285 sample proteins unique to Mascot search with 0.8Da MS tolerance

Code		Identity	Mw	#	Code		Identity	Mw	#
1	A4I9P1	Elongation factor 1-beta	25.9	3	14	A4IB09	Oligosaccharyl transferase-like protein	87.6	1
2	A4IDB2	Putative translation elongation factor 1	23.2	3	*14	A4IB08	Oligosaccharyl transferase-like protein	96.9	1
3	A4I708	Putative 3,2-trans-enoyl-CoA isomerase, mitochondrial	38.7	1	15	A4HRZ6	Calcium-transporting ATPase	112.5	1
4	A4I760	Uncharacterized protein	87.1	1	16	A4I5W4	Uncharacterized protein	13.2	1
5	A4ICY2	Putative L-ribulokinase	62.3	1	17	A4I7A9	Uncharacterized protein	27.5	1
6	A4I619	Uncharacterized protein	38.2	1	18	A4I8S4	Uncharacterized protein	101.5	1
7	A4I1Q1	Uncharacterized protein	39.6	1	19	A4HWJ8	60S ribosomal protein L6	21.2	1
8	A4I7S4	Putative small nuclear ribonucleoprotein	8.7	1	20	A4I994	Isocitrate dehydrogenase [NADP]	46.5	1
*8	A4I304	Putative small nuclear ribonucleoprotein	11.8	1	21	A4HRP7	Uncharacterized protein	218.0	1
*8	A4HZX8	Putative small nuclear ribonucleoprotein	11.2	1	22	A4I2G5	Uncharacterized protein	421.7	1
*8	A4IAF5	Putative small nuclear ribonucleoprotein	12.2	1	23	A4I988	Uncharacterized protein	159.0	1
9	A2CIE9	Cytosolic NADP-malic enzyme (Fragment)	60.5	3	24	E9AHY3	Uncharacterized protein	108.3	1
*9	B7SP15	Cytosolic NADP-malic enzyme (Fragment)	60.5	3	25	A4I394	Histone H2B	15.7	1
10	A4I504	Uncharacterized protein	26.7	1	26	A4IDB9	Uncharacterized protein	100.6	1
11	A4HT41	Anamorsin homolog	13.2	1	27	A4I290	Uncharacterized protein	206.6	1
12	A4I0S2	Uncharacterized protein	94.8	1	28	A4HZE7	Uncharacterized protein	63.3	1
13	A4IBA7	Putative kinesin	77.5	1	29	A4HUV8	Putative pumillio protein 10	81.2	1

As shown in Table 3.3, almost all of these exceptional proteins were identified through a single peptide sequence (depicted with # on the graph) therefore their identification is in doubt. Since uncharacterized proteins are hypothetically assumed proteins which were deducted from organism's genome, they definetely require more peptide match to prove their expression. Another reason is the redundancy in the database due to protein isoforms. In the Table 3.3, proteins indicated with asterisks actually correspond to the same protein however their total protein sequences are not identical as can be understand from differentiations in molecular weight.

Mostly it is not recommended to apply filter to the raw data before the database search since unbiased elimination may cause to lose the informative spectra however a proper selection may help to reduce the size of the data and search duration. On the other hand, post-processing of the results by taking into account the assigned scores, reinforces the reliability of the study. In mascot search engine, results can be formatted three different ways. Default significance threshold, p can be decrease from 0,05 to 0,01 or a cut-off can be set to the ion score (IS) which will remove the individual peptide spectrum matches (PSM) below that threshold. On the search results page, mascot reports a secore value for individual ions which indicates identity or extensice homology for that search and it was given as 30 for all searches with 1,0 Da and 0,5 Da for ms and ms/ms tolerance respectively. According to general opinion, PSMs with ion scores lower than 20 should be avoided to expel the random matches.

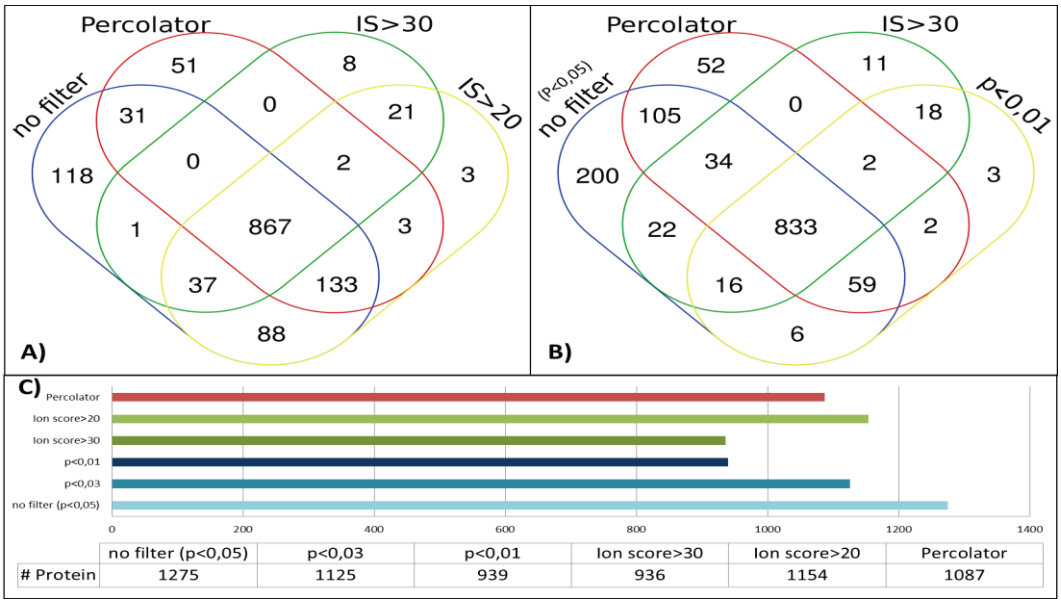


Figure 3.5. Distribution of common proteins obtained from Mascot searches. a, b) After post processings (CL-285), c) Total quantities on bar graph

In addition mascot provides percolator option which can be perfomed only by box-ticking. In Figure 3.5, comparison between these filtre options applied on sample CL-285 is given. Again, majority of the identified proteins over eight hundreds were shared in all searches. MS tolerance and ms/ms tolerance graphs given earlier above in Figure 3.2 and Figure 3.3 were plotted according to identity threshold (IS>30) however percolator results were more coherent to the results of other search engines in terms of FDR values therefore mascot percolator was chosen.

In this study, we also wanted to evaluate the performance of other database search engines and benefit from their result during final interpretation of differentially expressed proteins. For this purpose, xTandem, MS Amanda, Comet, MyriMatch and MS-GF+ search engines were selected. In addition, multiple search by combining Comet, xTandem, MS Amanda and MS-GF+ was carried out. Among the six aforementioned search engines, overlapping proteins whose uniprot codes are same displayed on a venn diagram in Figure 3.6, a for the sample CL-285. Since it is quite difficult to handle a venn diagram with six sets, new venn diagrams with five sets by substacting MS-GF+ or Mascot percolator were put in Figure 3.6, b. Lastly, venn diagrams drawed from the results of four different search engines and multiple search with those four were given both with full results and confident results. Judging by the total number of identified proteins, Myrimatch, MS-GF+ and MS Amanda are more selective than Mascot, xTandem and Comet. 517 proteins shared by all search engines correspond to 46% of xTandem total search, 47% of Comet total search, 67% of Myrimatch total search, 47% of Mascot total search, 76% of MS-GF+ total search and 64% of MSAmanda total search. Proteins common in searches of five programs (excluding MS-GF+), MS Amanda, MyriMatch, xTandem, Comet and Mascot were 91, 11, 38, 2, 3 and 33 respectively. Proteins shared in only two searches can give idea about the resemblance between search engines. In this context, xTandem-Comet pair gives the hightest number with 64, Mascot-xTandem pair follows them with 34 and Mascot-Comet pair comes after with 25. According to this assumption, closeness between xTandem and Comet seems much more than others. Lastly, we obtained the highest number of proteins from the multiple search, “4 combined” however it was not expected to see any protein in the set of only “4 combined” (60 unique proteins). In the result of protein list, PeptideShaker marks some of the validated matches as confident and some as doubtful. Figure 3.6, c section shows the venn diagrams which were drawed according to all proteins (left) and confident proteins (right) separately.

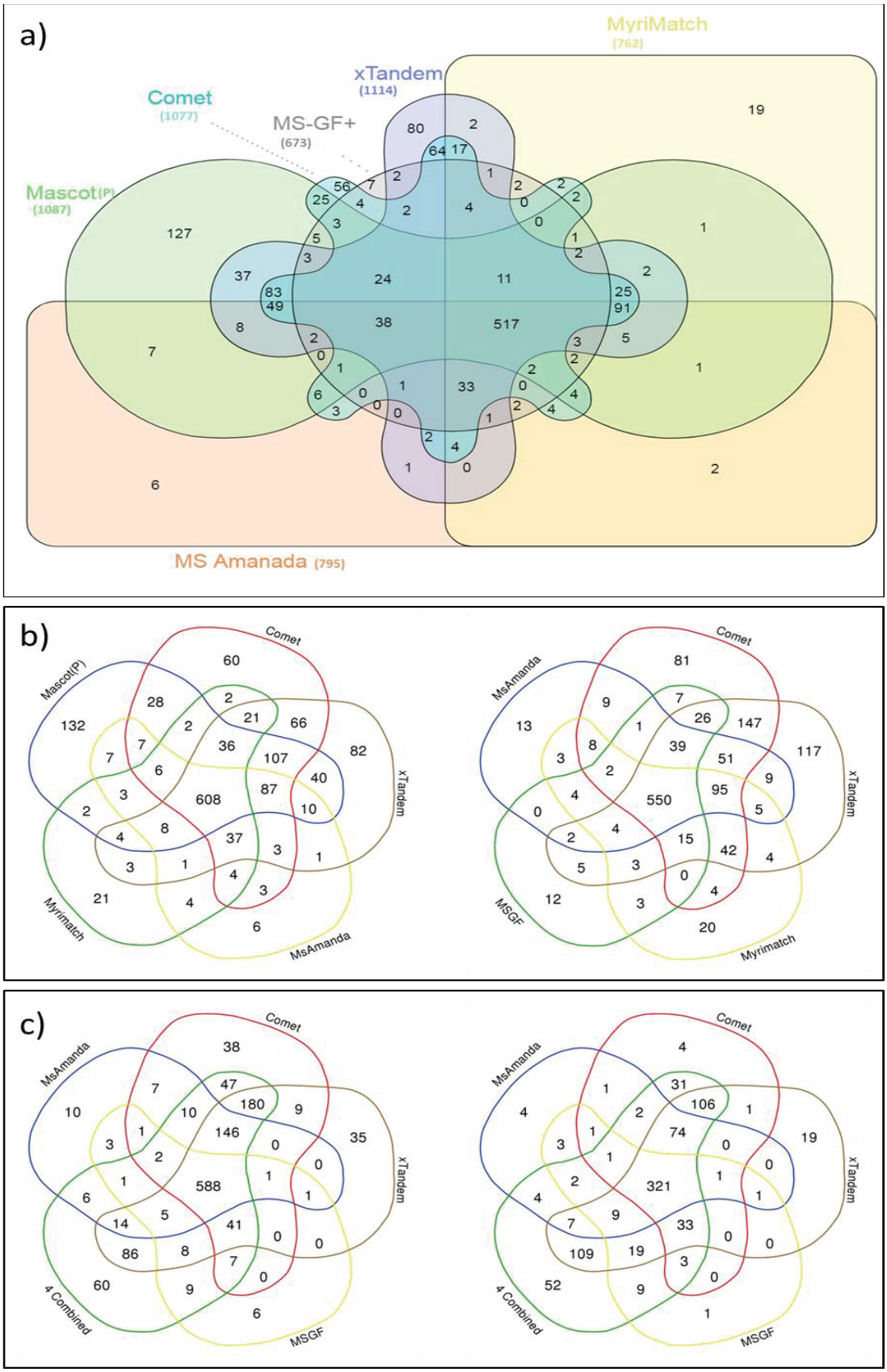


Figure 3.6. Venn diagrams for search results of a) all 6 search engines, b) 5 search engines without MS-GF (left) and Mascot-percolator, c) 4 search engines including their combined search (left) and same search engines by considering confident proteins

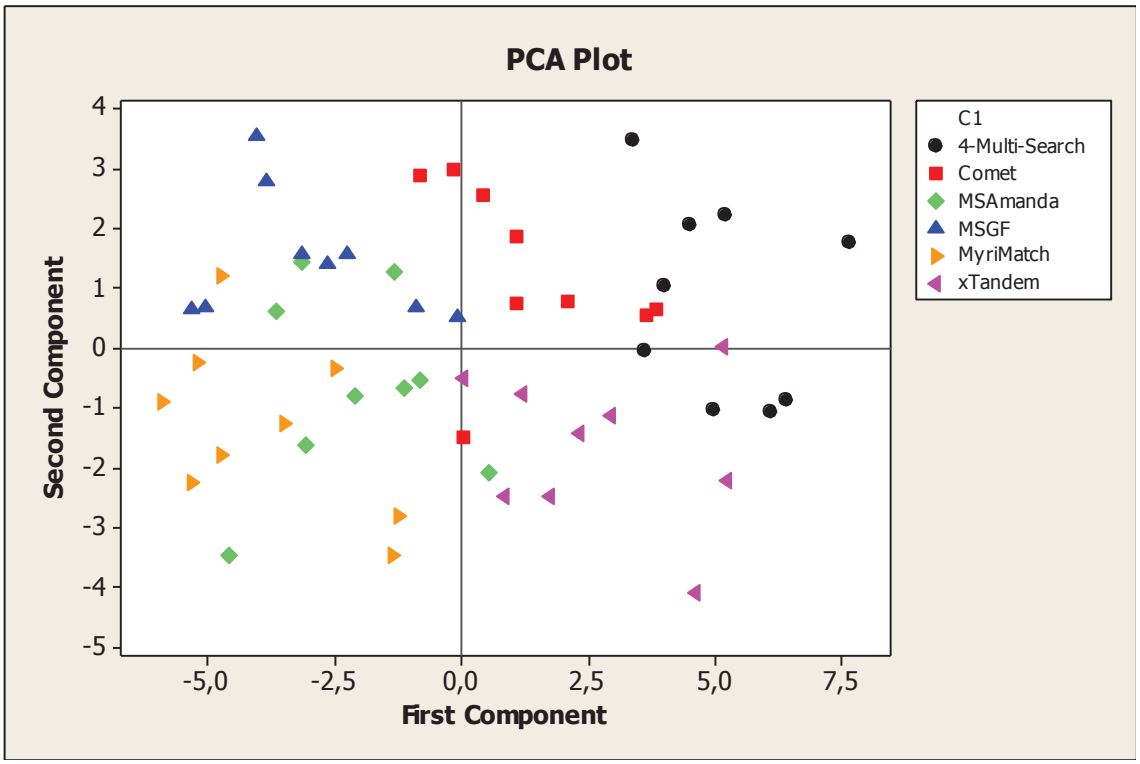


Figure 3.7. Principal component analysis for the results of five search engines

PeptideShaker program provides detailed information about the statistical validation that is comprised of several entries including validated number of proteins, peptides and peptide to spectrum matches in addition corresponding false discovery rate, false negative rate, condifence limit, total possible true positives. Information in the validation reports of all 9 samples are given in Appendix D and those values are used to draw a principle component analysis plot. Mascot searches were not included since they were not processed with PeptideShaker. As can be seen on the PCA plot (Figure 3.7), data points of xTandem and Comet as well as 4-Multi- Search (or 4combined) drops on the right hand side which overlaps with the earlier anticipation about the closeness between xTandem and Comet according to the number of shared proteins. Similarly, majority of datapoints belonging to MS Amanda and MyriMatch gather on the left hand side down and MS-GF+ explicitly dripts apart from them. Again, this results is coherent with the venn diagrams. Figure 3.6.b, shows that 5 specific proteins are shared between only MyriMatch and MS Amanda however this number is 0 for MS-GF+ and MyriMatch and likewise MS-GF+ and MS Amanda. As a reminder, venn diagrams were drawn considering only one sample (CL-285) however all samples contributed to the PCA plot, yet we obtained good correlation between PCA plot and venn diagram which can be attributed to the accuracy of PCA analysis.

Number of total protein identified from the corresponsing search conditions are given in Table 3.4 for all 9 samples. Elevated number of identified proteins in the searches against NCBI_L.genus database is arisen from the redundant proteins in this database since all species of Leishmania are involved here. On the other side, we specified UniProt database to *L.infantum* species thus redundancy is mostly confined to protein isoforms. As long as the necessary information is defined about the search, scientist does not have to defend his/her choices. If high reliability is preferred, either narrower mass tolerance before the search or strict filtration after the search should be applied. When it comes to bioinformatic tools, we can conclude about the performance of search engines judging by the search results (Figure 3.8) however we can not claim the superiority of a bioinformatic tool in terms of correctness.

Table 3.4. Number of proteins identified in different searches for all samples

Searches against NCBI_L. genus (MS tol:1.0 Da; MS/MS tol:0.5 Da unless indicated otherwise)									
	CL (L.infantum)			**VL (L. infantum)**			**CL (L.tropica)**		
	121	**259**	**285**	**240**	**495**	**628**	**615**	**Crt01**	**Crt02**
Mascot (MS tol: 04 Da)	900	1224	1051	897	849	931	830	931	1006
Mascot (MS tol: 0.5Da)	1133	1490	1432	1165	1142	1207	1183	1265	1250
Mascot (>20 Ion Score)	1039	1403	1332	1082	1048	1127	1100	1191	1186
Comet	1048	1506	1442	913	976	1070	1164	1312	2882
Mascot Searches against UniProt_L.infantum (MS tol:1.0 Da; MS/MS tol:0.5 Da unless indicated otherwise)									
	CL (L.infantum)			**VL (L. infantum)**			**CL (L.tropica)**		
	121	**259**	**285**	**240**	**495**	**628**	**615**	**Crt01**	**Crt02**
MS/MS tol: 0.4 Da	755	1025	907	785	728	839	703	813	866
MS/MS tol: 0.5 Da	1026	1289	1275	1014	980	1119	1068	1112	1102
Filtration >30 Ion Score	723	998	936	716	701	795	759	836	840
Filtration >20 Ion Score	893	1182	1154	913	888	989	961	1026	1016
Mascot Percolator	822	1081	1087	806	819	869	877	978	929
Searches against UniProt_L.infantum with Search GUI & Peptide Shaker (MS tol:1 Da; MS/MS tol: 0.5 Da unless indicated otherwise)									
	CL (L.infantum)			**VL (L. infantum)**			**CL (L.tropica)**		
	121	**259**	**285**	**240**	**495**	**628**	**615**	**Crt01**	**Crt02**
Comet	836	1171	1071	839	755	880	848	841	898
MS Amanda	699	809	795	634	622	682	670	724	772
MS-GF+	530	722	673	575	562	684	459	506	714
MyriMatch	604	744	762	581	587	580	591	634	688
xTandem	781	1036	1114	830	795	824	859	897	900
Multiple Search(- MyriMatch)	951	1197	1210	911	940	982	1004	1061	1048

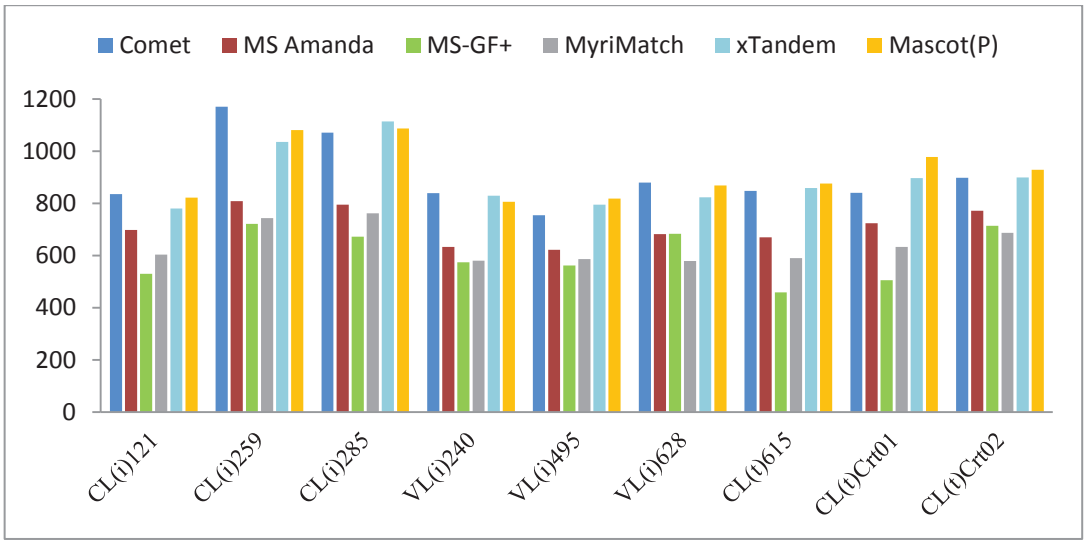


Figure 3.8. Number of identified protein by 6 different search engines for all samples

Certainly, simultaneous detection of a protein in the search results of different search engines can be postulated as more confident identification however doing several searches and compling the shared proteins as given in Figure 3.9 is not a practical way therefore a criteria which requires at least two or more peptide sequences (two-peptide rule) may be applied optionally to call a protein "identified" for a search with one search engine, not to mention, this may cause to lose some valuable information.

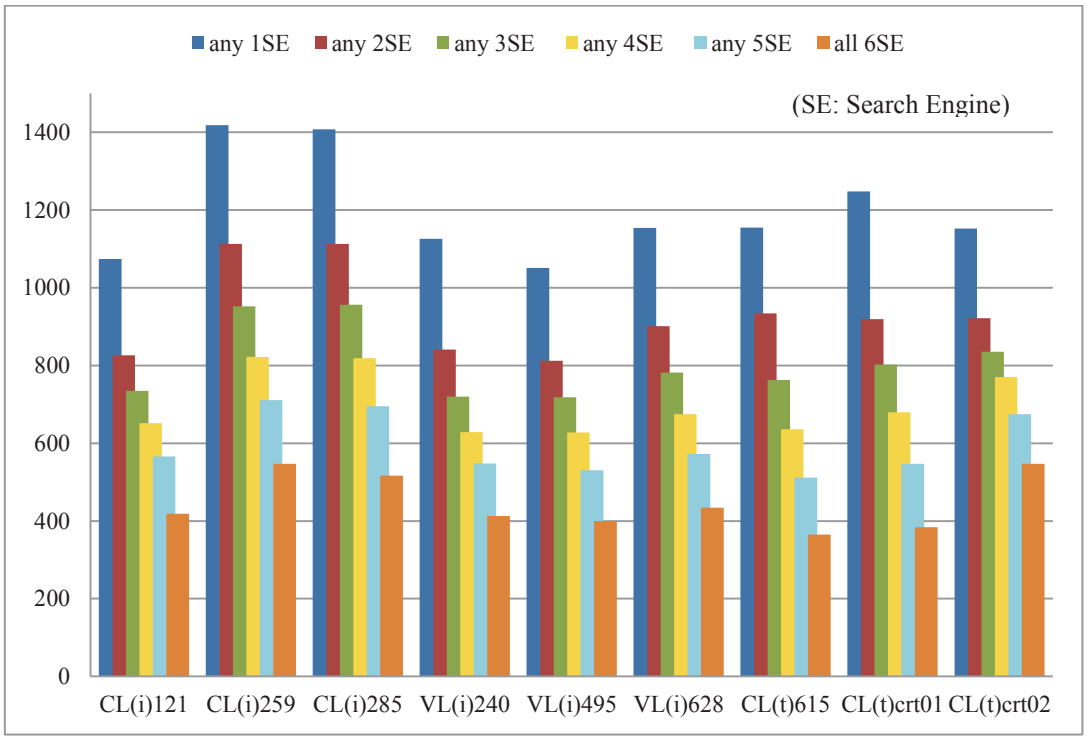


Figure 3.9. Number of shared proteins identified by more than one search engines

3.3.2 Differential Proteins Obtained From Shotgun Proteomic Analysis

In this study, we carried out ms-based qualitative proteomic analysis to determine the differential proteins between parasites causing visceral and cutaneous types of Leishmaniasis diseases. 9 shotgun proteomic results out of 15 samples which were used for gel-based proteomic analysis in chapter 2 were compared. Firstly, visceral and cutaneous comparison were evaluated mainly on 6 samples which are belong to same species, *L.infantum*. In order to enhance the credibility, results from six different bioinformatic tools were incorporated to pick the differential proteins. Later, 3 cutaneous leishmaniasis samples with *L.tropica* origin were put into other cutaneous leishmaniasis samples to widen the investigation in a manner beyond the species. In order to obtain statistically more meaningful results, database searches were performed by using two bioinformatic tools and two alternative databases. Compiled results were given as protein lists (Table 3.6 & Table 3.7) therefore proteins appearing repeatedly on these list are considered to deserve special attention. In general, protein identified by a single peptide is not considered as identified thus they are removed from the results. By searching the data with different search engines, we aimed to overcome this restriction.

In Table 3.5, differential proteins observed only in *L.infantum* samples are listed. Proteins observed twice in CL and never in VL was depicted with hollow circle, likewise with filled circle when they were observed in all three samples of CL. Proteins from the first one on the list to ORF2 (in total 39) were more reliable since they were shared by at least 4 search engines wherease the rest were shared by at least 3 out of 6. In this case, the latter may not be statistically significant compared to 4 since 4 is more than the half. The rightest column of Table 3.5 does not belong to a single search engines as is the case with others. To create this column, protein results of one sample obtained from 6 search engines are pooled into a single file first, then those proteins which were not shared by at least 3 search engines were kicked out from the identified protein list of that particular sample. Later, differential proteins list was constituted in the same way as explained above.

In fact, one single seach engine is utilized according to the common practice therefore another recognition is accepted for the validation. According to this rule, to declare the expression of a particular protein, at least two different peptide sequences matches are obliged. We did not prefer to apply this rule not to lose any information.

Table 3.5. Differential proteins compiled from the results of different search engines

Protein Name	Codes	*Mascot*	*Comet*	*Msamanda*	*Msgf+*	*Myrimatch*	*xTandem*	*shared by min. 3SE*
Differential proteins (not present in VL) in at least 2 cutaneous leishmaniasis samples								
Acidocalcisomal pyrophosphatase	A4HUT3	●	●	●	●	●	●	●
Putative glutamamyl carboxypeptidase	A4I4L8	●	●	●	●	●	●	●
Uncharacterized protein	A4ICR0	●	●	●	●	●	●	●
Uncharacterized protein	A4IB83	○	○	○	○	○	○	○
Flavoprotein subunit-like protein	A4HTA6	●	●	○	●	●	●	●
Uncharacterized protein	A4HWG6	○	○	○	○	○	○	○
Putative TPR-repeat protein	A4I340	●	●	○	○	-	●	●
Putative PUF1	A4ICY3	○	○	○	○	-	○	○
Protein phosphatase type 1 regulator-like protein	A4HSP2	○	○	○	-	○	○	○
Putative eukaryotic release factor 3	A4HV24	●	●	●	-	●	●	●
Putative N-acyl-L-amino acid amidohydrolase	A4I6N8	○	○	○	-	○	○	○
Putative NADH dehydrogenase	A4IDV2	○	○	○	-	○	○	○
Uncharacterized protein	A4I1Q4	●	●	○	○	○	-	○
Putative lanosterol 14-alpha-demethylase	A2TEF2	●	-	●	●	●	●	●
ATPase ASNA1 homolog	ASNA	●	-	●	○	○	●	-
Uncharacterized protein	A4HV02	-	●	●	○	○	●	●
Uncharacterized protein	A4I4S1	-	○	○	○	○	○	○
Uncharacterized protein	A4HTX5	-	●	●	●	○	●	●
Proline dehydrogenase	A4I294	-	●	●	○	○	●	●
Cation-transporting ATPase	A4HTD0	○	○	○	-	-	○	○
Uncharacterized protein	E9AH64	●	○	●	-	-	●	●
Uncharacterized protein	A4HU32	○	○	○	-	-	○	○
UDP-galactopyranose mutase	A4HXP6	○	○	○	-	-	○	○
Putative ubiquitin hydrolase	A4I519	○	●	-	○	-	●	●
Uncharacterized protein	E9AHC9	○	○	-	○	-	○	○
Putative deoxyribodipyrimidine photolyase	A4I8R0	○	○	-	○	-	○	○
Uncharacterized protein	A4HUZ7	●	○	-	○	-	●	○
Putative translation initiation factor IF-2	A4I9B4	○	○	-	○	-	○	○
Uncharacterized protein	A4ID44	○	○	○	-	○	-	○
Putative kinesin	A4I5H9	○	-	○	○	-	○	○
Uncharacterized protein	A4IBW6	○	-	○	-	○	○	○
Putative X-pro, dipeptidyl-peptidase,serine peptidase,Clan SC, family S15	A4I3S2	○	-	○	-	●	●	●
Putative seryl-tRNA synthetase	A4HUS2	●	-	●	○	○	-	●
S-adenosylmethionine transporter	A4HUE5	-	●	○	○	-	○	○
Uncharacterized protein	A4I741	-	○	○	○	-	○	○
Putative cytochrome b5	A4HUW0	-	○	●	-	○	○	●
Putative ribose-phosphate pyrophosphokinase	A4HTJ9	-	○	○	-	○	○	○
COP-coated vesicle membrane protein erv25	A4IDI9	-	-	●	●	●	●	●

(cont. on next page)

Table 3.5 (cont.)

ORF2	Q25299	-	-	○	●	●	●	●
Putative DNA-directed RNA polymerases II	A4HXV2	○	○	-	-	-	○	○
D-lactate dehydrogenase-like protein	A4I481	○	●	-	-	-	○	○
Uncharacterized protein	A4HZE7	●	●	-	-	-	○	○
Guanine nucleotide-binding protein beta subunit-like protein	E9AH24	○	○	-	-	-	○	○
Putative 5-oxoprolinase	A4HXX8	○	●	-	-	-	●	●
Putative pteridine transporter	E9AG72	○	○	-	-	-	○	○
Uncharacterized protein	A4HUC3	○	○	-	-	-	○	○
Putative haloacid dehalogenase-like hydrolase	A4I3L6	●	○	○	-	-	-	○
Putative vesicular-fusion ATPase-like protein	A4I2Q9	●	●	-	○	-	-	●
Putative ribonuclease II-like protein	A4I3A3	○	○	-	○	-	-	○
Putative cytochrome c oxidase subunit VI	A4HZH7	○	○	-	○	-	-	○
Putative 2-oxoisovalerate dehydrogenase alpha subunit	A4HZE0	○	○	-	-	○	-	○
Putative ATP-dependent RNA helicase	A4HXH0	○	○	-	-	○	-	○
S-methyl-5'-thioadenosine phosphorylase	A4HSK5	●	○	-	-	●	-	○
Probable methylthioribulose-1-phosphate dehydratase	MTNB	○	-	○	-	-	●	-
Putative ribose 5-phosphate isomerase	A4I3S4	○	-	○	○	-	-	○
Bifunctional NAD(P)H-hydrate repair enzyme	NNR	○	-	○	-	○	-	-
Transcription elongation factor-like protein	A4I9C2	●	-	●	-	●	-	-
ATP-dependent Clp protease subunit, heat shock protein 78 (HSP78)	A4HRR6	●	-	●	-	○	-	○
1-alkyl-2-acetylglycerophosphocholine esterase	A4IBJ0	-	○	○	-	-	○	○
Uncharacterized protein	A4I3Z9	-	○	-	○	-	○	○
Fibrillarin	A4ID95	-	○	-	○	-	○	○
RNA-binding protein-like protein	A4IAK8	-	○	-	-	○	○	○
Putative chaperone protein DNAj	A4IC84	-	○	-	-	○	○	○
Uncharacterized protein	A4HXZ9	-	○	○	-	○	-	●
Uncharacterized protein	A4HRH1	-	●	○	-	○	-	○
Lysine--tRNA ligase	A4HWC4	-	-	○	○	-	●	●
Putative DNAJ domain protein	A4I0P2	-	-	○	-	○	○	○
Putative N-acyl-L-amino acid amidohydrolase	A4I5G9	-	-	○	-	○	○	○
ER membrane protein complex subunit 3	A4ICV1	-	-	-	○	●	●	-
Histone H2B	A4I394	-	-	○	○	○	-	○
Putative ribonucleoside-diphosphate reductase small chain	A4I2X5	-	-	●	●	○	-	●
Uncharacterized protein	A4I6P7	-	-	●	○	●	-	●
Putative ribosomal protein L27	A4I890	-	-	○	○	○	-	
Uncharacterized protein	A4HRP2	-	-	●	●	●	-	●
Putative phosphatase 2C	A4I565	-	-	●	○	●	-	●
Differential proteins (not present in CL) in at least 2 visceral leishmaniasis								
Histone H2A	E9AGW0	●	●	●	○	○	●	○
Uncharacterized protein	A4I4K9	○	○	○	○	○	○	○
Zinc transporter-like protein	A4I745	●	○	●	-	●	●	○
I/6 autoantigen-like protein	A4I005	●	●	●	-	○	●	○
Uncharacterized protein	A4HYQ3	○	○	-	○	-	○	-
Uncharacterized protein	A4IDE8	-	○	○	○	-	○	○
Putative kinetoplast-associated protein	A4IDZ7	-	-	○	●	●	●	○

Among 39 proteins abovementioned, 19 protein are involved in either one searches containing *L.tropica* samples. These proteins are as follows Putative glutamamyl carboxypeptidase (A4I4L8), Uncharacterized protein (A4ICR0), Acidocalcisomal pyrophosphatase (A4HUT3), Flavoprotein subunit-like protein (A4HTA6), Uncharacterized protein (A4HWG6), Putative TPR-repeat protein (A4I340), Putative eukaryotic release factor 3 (A4HV24), Putative N-acyl-L-amino acid amidohydrolase (A4I6N8), Putative NADH dehydrogenase (A4IDV2), Putative lanosterol 14-alpha-demethylase (A2TEF2), Uncharacterized protein (A4I4S1), Proline dehydrogenase (A4I294), Uncharacterized protein (E9AH64), Uncharacterized protein (A4HUZ7), Uncharacterized protein (A4ID44), Putative X-pro, dipeptidyl-peptidase,serine peptidase,Clan SC, family S15 (A4I3S2), Putative seryl-tRNA synthetase (A4HUS2), S-adenosylmethionine transporter (A4HUE5), COP-coated vesicle membrane protein erv25 (A4IDI9). From these 19 protein, 9 proteins are standed out by being shared by both Mascot search and multiple search (with 4 search engines; Comet, xTandem, MS Amanada, MS-GF+). Therefore these 9 proteins including Putative glutamamyl carboxypeptidase (A4I4L8), Proline dehydrogenase (A4I294), Flavoprotein subunit-like protein (A4HTA6), Uncharacterized protein (A4ICR0), Uncharacterized protein (E9AH64), Putative haloacid dehalogenase-like hydrolase (A4I3L6), Putative lanosterol 14-alpha-demethylase (A2TEF2), Putative NADH dehydrogenase (A4IDV2), Uncharacterized protein (A4HUZ7) are attracted our attention more than the others. Moreover, some of these proteins are also appeared in the results of searches against NCBI-Prot of *L.genus*. Results are given in same way with a table in Appendix C. Actually, searching highthroughput shotgun data against a large database is not recommended due to elevated false positives. Besides, database containing all species causes great complexity and comparison through the accession codes is mostly misleading due to the tight discrimination. Nevertheless, a few significant proteins are revealed by this cross validation despite the problems mentioned above. In deed, acidocalcisomal pyrophosphatase, flavoprotein subunit-like protein, conserved hypothetical protein (A4ICR0) and putative lanosterol 14-alpha-demethylase almost each time in the results thus their expressions become unquestionable compared to other proteins in the lists. So far only differential proteins in cutaneous type disease were explored because differential proteins in visceral type are very limited. Histone 2A (E9AGW0), Zinc transporter-like protein (A4I745) and I/6 autoantigen-like protein (A4I005) proteins are worth to considering.

Table 3.6 Differential proteins according to Mascot search results of all samples

Group	Samples	Acession	Protein Name
All Cutaneous Samples Linf.+Ltrop	121 259 285 615 Crt01 Crt02	A4I824	Cop-coated vesicle membrane protein p24
		A4I4L8	Putative glutamamyl carboxypeptidase
		A4HRR6	ATP-dependent Clp protease subunit, heat shock protein 78 (HSP78)
		A4I294	Proline dehydrogenase
		A4HUY0	ATPase ASNA1 homolog
		A4HTA6	Flavoprotein subunit-like protein
		A4I330	Obg-like ATPase 1
		A4ICR0	Uncharacterized protein
		A4HUS2	Putative seryl-tRNA synthetase
5-CL L.inf+Ltrop.	121-259-285 615-Crt01	**E9AH64**	Uncharacterized protein
		A4I7H1	Glycylpeptide N-tetradecanoyltransferase
		A4I629	Putative CAS/CSE/importin domain protein
		A4I3L6	Putative haloacid dehalogenase-like hydrolase
	121-259-285 615-Crt02	A4HX58	Putative eukaryotic translation initiation factor 4 gamma
		A4IDI9	COP-coated vesicle membrane protein erv25
		A4IC14	Cyclophilin 40
	121 259 285 Crt01 Crt02	**A2TEF2**	Putative lanosterol 14-alpha-demethylase
		A4I3S2	Putative X-pro, dipeptidyl-peptidase,serine peptidase,Clan SC, family S15
		A4HTB1	Putative ubiquitin-conjugating enzyme e2
		A4IBT4	Protein kinase A catalytic subunit isoform 2
	121-259-615-Crt01-Crt02	A4ID44	Uncharacterized protein
		A4I621	Uncharacterized protein
	121-285 615-Crt01-Crt02	A4HXX2	UDP-glucose pyrophosphorylase
		A4IB27	Uncharacterized protein
		A4HZE7	Uncharacterized protein
	259 285 615 Crt01 Crt02	A4HVH9	Putative RNA helicase
		A4IC10	Uncharacterized protein
		A4I946	Putative small GTP-binding protein Rab18
		A4HW39	Uncharacterized protein
		A4I7R1	Uncharacterized protein
		A4I5T2	Uncharacterized protein
		A4IDV2	Putative NADH dehydrogenase
3-CL L.inf.	121 259 285	A4HTQ0	Uncharacterized protein
		A4I647	Putative mago nashi-like protein
		A4HUZ7	Uncharacterized protein
		A4I340	Putative TPR-repeat protein
		A4HW64	Putative poly(A) polymerase
		A4I5W9	Putative aldehyde dehydrogenase
		A4I2L1	Putative methylmalonyl-coenzyme a mutase
3-CL L.trop.	615 Crt01 Crt02	**A4HYX0**	Uncharacterized protein
		A4HZX1	Uncharacterized protein
		A4IAE6	Phosphomannomutase-like protein
		A4HSS6	NAD(P)-dependent steroid dehydrogenase-like protein
3-VL L.inf.	240 495 628	**A4I005**	I/6 autoantigen-like protein
		E9AGW0	Histone H2A
		A4I9B3	Uncharacterized protein

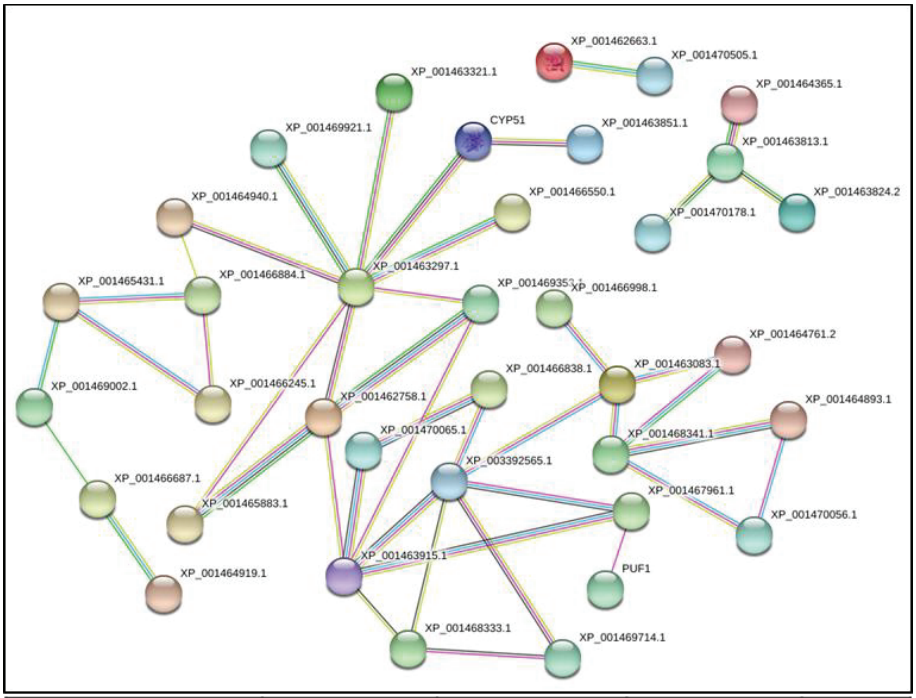
Table 3.7 Differential proteins according to 4-combined search results of all samples

Group	Samples	Accession	Protein Name
All Cutaneous Samples Linf.+Ltrop.	121 259 285 615 crt01 crt02	A4HRT4	Putative long chain fatty Acyl CoA synthetase
		C5MKJ9	Heat shock protein 70 (Fragment)
		A4HV24	Putative eukaryotic release factor 3
		A2TEF2	Putative lanosterol 14-alpha-demethylase
		E9AH64	Uncharacterized protein
		A4HSS6	NAD(P)-dependent steroid dehydrogenase-like protein
		A4I294	Proline dehydrogenase
		A4HTA6	Flavoprotein subunit-like protein
		A4I4S1	Uncharacterized protein
		A4I4L8	Putative glutamamyl carboxypeptidase
		A4HRH1	Uncharacterized protein
		A4ICR0	Uncharacterized protein
		A4HT47	Uncharacterized protein
5-CL L.inf+Ltrop.	121 259-285 615-Crt01	A4I2Q9	Putative vesicular-fusion ATPase-like protein
		A4I481	D-lactate dehydrogenase-like protein
		A4I474	Uncharacterized protein
	121-259 285-615- Crt02	A4HUT3	Acidocalcisomal pyrophosphatase
		A4I0M2	V-type proton ATPase subunit a
	259 285 615 crt01 crt02	A4I051	V-type proton ATPase proteolipid subunit
		A4I6N8	Putative N-acyl-L-amino acid amidohydrolase
		A4HWG6	Uncharacterized protein
		A4IBJ0	1-alkyl-2-acetylglycerophosphocholine esterase
		A4IDV2	Putative NADH dehydrogenase
		A4HW39	Uncharacterized protein
3-CL L.inf.	121 259 285	**A4I3L6**	Putative haloacid dehalogenase-like hydrolase
		A4I2G4	Putative sec1 family transport protein
		A4HUE5	S-adenosylmethionine transporter
		A4HXX8	Putative 5-oxoprolinase
		A4HUZ7	Uncharacterized protein
		A4I647	Putative mago nashi-like protein
3-CL L.trop.	615 Crt01-Crt02	A4IC24	Uncharacterized protein
		A4HYX0	Uncharacterized protein
3-VL L.inf.	240 495 628	**A4I005**	I/6 autoantigen-like protein
		E9AGW0	Histone H2A
		A4HTI0	Superoxide dismutase

As it is understood, inconsistences can occur easily in the results of database searching by changing the database search settings. By increasing the sequence coverage in other words, number of identified pieces in a protein, it can be prevented or cut down. It can only be possible with the improved quality in data acquirement. For this purpose, high resolution mass spectrometers and HPLC separations with long columns at nanolitre flow have become indispensable for proteomic study. Alternative data collection mode called data-independent acquisition is also being applied and quantitative analysis can be attined as well when statistically significant repeat is performed. Taking into account all of these, these results which were acquired from an old fashioned electrospray mass spectrometer coupled to relatively short column seem

narrow and ambiguous, still some distinct and reliable proteins were determined for a start. Last but not least, it should be bear in mind that sources of all samples were obtained from different individual patients and it is known that these different parasite strains exhibit virulence effect in a broad spectrum even the ones that cause same disease form. It can be concluded from their growth rate in the same medium and complications after the injection to mice therefore each clinical sample should be considered as unique. Since our aim was to detect the most relevant protein candidates which are responsible for the occurance of disease forms, this fact was disregarded and results were evaluated in groups on purpose

Finally, possible protein-protein interactions between the proteins in Table 3.5 were inspected through the internet-based software, string-db[138]. High confidence interactions (>0.7) were quite limited nevertheless we get idea about hypothetical protein (E9AH64) though the interaction with putative eukaryotic release factor 3.



Node1	Accession code	Node2	Acession code	Score
XP_003392565.1	E9AH64	XP_001463915.1	A4HV24	0.981
XP_001468341.1	A4I9C2	XP_001464893.1	A4HXV2	0.977
XP_001470236.1	A4I3S4	XP_001463390.1	A4HTJ9	0.924
XP_001469353.1	A4IC84	XP_001462758.1	A4HRR6	0.917
XP_001465883.1	A4I0P2	XP_001462758.1	A4HRR6	0.874
XP_003392565.1	E9AH64	XP_001466838.1	A4I519	0.786

Figure 3.10 Protein-protein interaction network and significant interactions

3.4. Conclusion

In shotgun proteomic analysis, high throughput data obtained from mass spectrometry is searched against a defined database which contains all possible proteins' sequences according to the genomic material. Results of this search provide a collection of protein expressed in that organism. However amount and content of this collection which contains from hundreds to thousands elements alter according to the search settings. The degree of error window for the numerical outputs of the precursor and fragments ions are adjusted with MS and MS/MS tolerance parameters. The most reasonable values for these two determinants were optimized as 1.0Da and 0.5Da for mascot searches and they were applied to all searches. Besides, search results from six different search engines were utilized for the validation of differential proteins. Differential proteins were compiled and listed by considering different aspects, yet some particular proteins appeared every time which ensures their correctness. Among them, four proteins were striking than the other and encountered in the literature.

Elucidation of signal pathways and biochemical mechanisms in parasite cell is of critical importance for the accurate treatment of leishmaniasis disease. As is well known, parasite survival is actualized by the transformation from promastigote form to non-flagellated amastigote in host's macrophages. In addition to regulation and activation of other inflammatory cells, defense mechanism of macrophages relies on phagolysosome, the fusion process of phagosome with lysosome organelle. Macrophage aims to destroy invader parasite cells by releasing reactive oxygen species however Leishmania parasites are able to endure these activities. Apart from that, parasites survival is associated with the adaptation to new environment which was switched from 22 °C/pH 7.2 to 37 °C/pH 5.5[139]. It has been demonstrated that this shift in conditions trigger the transformation through cAMP-mediated mechanism and acidocalcisomal pyrophosphatase enzyme has a crucial role in the production of cAMP from ATP[140]. In our study, expression of this enzyme only in promastigote isolates from cutaneous leishmaniasis is proven persistently with several searches. Macrophage diversity aside, appearance of an enzyme related with intra-macrophase survival only in the cutaneous leishmaniasis samples implies that parasites causing visceral leishmaniasis prefer an alternative mechanism to cAMP regulation despite the fact that all samples were promastigote form and cultivated in growth medium. Nevertheless, these findings

should be supported with more samples of visceral leishmaniasis since there only three samples of visceral leishmaniasis were involved to the research compared to six samples of cutaneous leishmaniasis. In addition, proteomic profiling of amastigote form will elaborate the argument.

As can be understood from the name, flavoprotein subunit-like protein is likely a similar protein to succinate dehydrogenase [ubiquinone] flavoprotein subunit which involves in the methabolic pathways of citrate cycle and oxidative phosphorylation and string-db confirms it with high scoring. In TCA cycle, it catalyse the reaction of fumarate production from succinate and in oxidative phosphorylation, it transfers electrons from succinate to ubiquinone. Accordingly, parasites causing to cutaneous leishmaniasis either include an extra protein to boost the carbohydrate methabolism and energy production to use it alternatively.

Putative lanosterol 14-alpha-demethylase is an enzyme involved in steroid biosynthesis. Major antifungal and antileishmanial treatments generally affect on membrane integrity by targeting the sterol metabolism through the ergosterol, a C24 alkylated sterol synthesized in the plasma membrane of these parasites and fungi. Plasma membrane of vertebrates embodies cholesterol as the major sterol which may provide selectivity for the drug in binding to ergosterol however studies show that *Leishmania spp.* also obtain and accumulate cholesterol from the host. In a recent study, reduction of ergesterol and concomitant formation of alternative sterols are demonstrated by metabolomic analysis. The following genome sequencing reveals many single nucleotide polymorphisms (SNP) on the gene of sterol 14α-demethylase in the resistant cells[141]. In our all searches, putative lanosterol 14-alpha-demethylase protein is not observed in the samples of visceral leishmaniasis (VL) group while cutaneous leishmaniasis (CL) samples of both *L. infantum* and *L. tropica* contain this protein in their proteome. Deficiency of ergesterol enables amphotericin B reistant parasites to avoid the drug binding. Same principle may work on VL samples via changing the sterol variety of the membrane to hide from host's immune system.

Another remarkable protein, zinc transporter-like protein is detected only VL samples. In a recent study, effects of zinc on parasite viability has been investigated by adding zinc chelators into the medium and apoptosis like deaths of both drug sensetive and drug resistant parasites has been observed due to the zinc depletion[142]. This transporter protein may be an indicator of a mechanism for zinc uptake.

CHAPTER 4

A METHODICAL APPROACH TO COMBINE ON-TISSUE DIGESTION WITH MATRIX APPLICATION FOR MALDI MASS SPECTROMETRY IMAGING

4.1. Mass Spectrometry Imaging

Clinical imaging techniques of histology, such as magnetic resonance spectroscopy imaging (MRI), positron emission tomography (PET), autoradiography (X-ray), near-infrared fluorescent (NIRF) or bioluminescence imaging are used for the spatial localization of substances in a tissue, after the introduction of a radioactive ligand (tracer). Unlike these techniques, mass spectrometry imaging provides label-free chemical imaging either targetted manner as drug distribution or nontargetted analyses of multiple classes of biomolecules. According to the purpose of the study, molecular species can be peptides, lipids, parent drug, drug metabolite or other existing small metabolites. In short, today mass spectrometry imaging (MSI) has become a distinguished tool for visualizing molecules in a tissue section, cell culture or other surface-based analysis like fingerprint through the mass differences of detected analytes. MSI methods can be listed as matrix-assisted laser desorption ionization mass spectrometry imaging (MALDI-MSI), desorption electrospray ionization mass spectrometry imaging (DESI-MSI), laser ablation electrospray ionizaion mass spectrometry imaging (LAESI-MSI) and imaging secondary ion mass spectrometry (SIMS)[143]. Among them, MALDI-TOF instrumentation has the largest popularity and market share. As the achievements of MSI in the field of tissue characterization, drug monitoring, and clinical diagnosis increase, importance of the resolution and mass accuracy have been well understood thus popularity and high demand on MALDI-TOF instrumentation for MSI incresed in time. Although it may not be applicable to all tissue or analyte type, maldi mass spectrometry imaging may be rendered compatible with other imaging techniques in histology[144]. In the mainstream staining which utilizes hematoxylin and eosin (H&E), blue stained nuclei is contrasted on the pink cytoplasm.

If coated matrix can be removed with caution, H&E staining can be applied to the same tissue. Alternatively, serial sections are used although it does not serve the exact purpose[145]. Serial sections are actually being benefitted from the constraction of three-dimensional image[146].

Imaging mass spectrometry using Matrix-Assisted Laser Desorption/Ionization allows the detailed mapping of biomolecules directly from tissue section. Matrix deposition is the key step for highly resolved neat imaging since appropriate concentration and deposition of matrix is critical for extraction, desorption and ionization of molecules from tissue without losing molecular localization. In total, imaging process involves sequential steps of tissue sectioning, matrix deposition, data acquisition, and image construction.

From the beginning of sample acquisition to mass spectrometric analysis, great care should be taken to preserve the morphology of the tissue section and spatila distribution of analytes. Rapid cooling of the samples is prefered by snap freezing to reduce the crystal ice formation of the water however MALDI imaging of the samples prepared with the most widespread concervation technique, formalin fixation followed by paraffin embedding can also be employed. Cryostats allow to cut the tissue at desirable thickness (5-15μm) under low temperature. Electrical conductivity may be adversely affected by the excessive thickness and drying process takes longer time. On the other hand, thin sections are prone to be torn easily. Although stainless steel and gold-coated MALDI targets are applicable, their opaque nature pose an obstacle for the microscopic visualization therefore conductive glass slides which are coated with indium tin oxide (ITO) are developed. Prior to the deposition of the matrix, rinsing procedure is recommended for the fixation of the tissue onto the support and removal of interfering analytes such as lipids or salts as well. Between the steps of rinsing, matrix application or on-tissue digestion, slides should be kept in the vacuum desiccator for faster drying and protection[147]. On-tissue digestion is required for protein analysis due to the low detection limits however activity conditions of the enzymes involving elevated temperature and humidity conflict with dryness. Risk of diffusion increases by the treatments containing excess liquid of aqueous solutions[148]. For matrix application, several strategies based on either coating or spotting are being implemented. Minute amount of matrix solution is dispensed from the ejectors using either acoustic transfer or piezoelectric technology. Coating is provided by spraying of fine matrix droplets through pneumatic pressure or electrospray deposition.

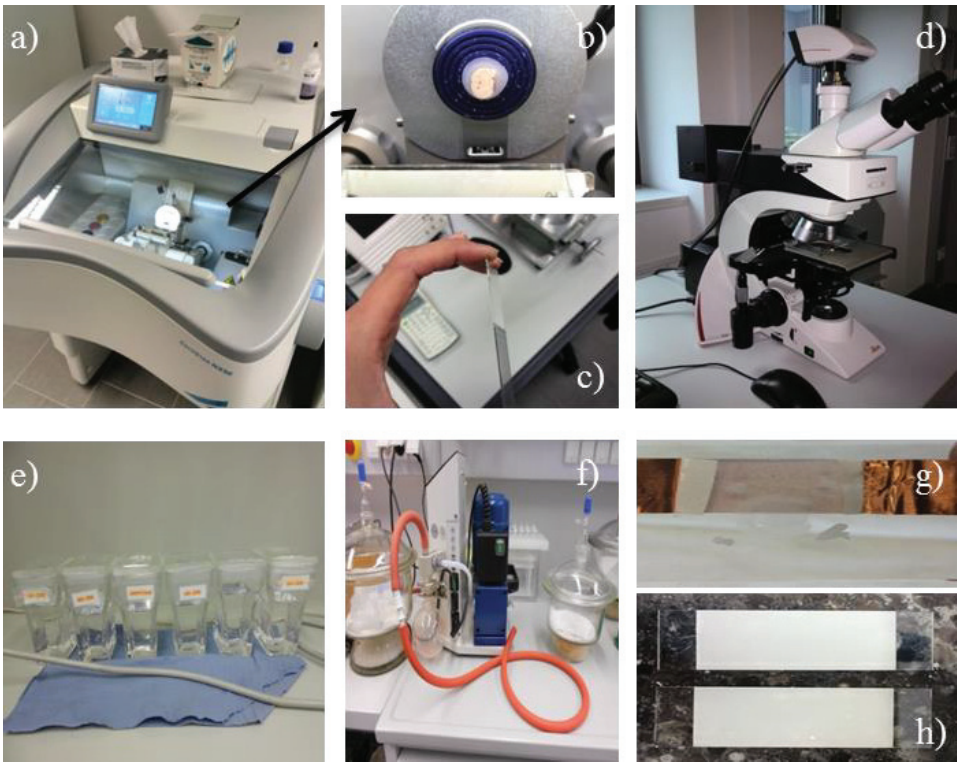


Figure 4.1. Sample preparation steps in MSI: a) Cryostat b) Tissue cutting c) ITO coated glass slide, d) Microscope, e)Rinsing flasks, f) Vacuum desiccator, g) Tissue section before matrix application, h) Tissue section after matrix application by matrix sublimation

Matrix deposition by sublimation is very favorable method in terms of surface homogenity therefore home-built sublimation apparatus has started to be preferred in MALDI imaging laboratories. Certain amount of matrix solution is prepared in volatile organic solvents, poured onto a planar plate area. A thin crystal surface is obtained via rapid evaporation of solvent. Thin layer of matrix is sublimated onto glass slide where tissue was mounted in a vacuum sealed chamber. Subsequent hydration/recrystallization under the vapour of acidic solution improves the signal intensity provides high-spatial resolution images[149].

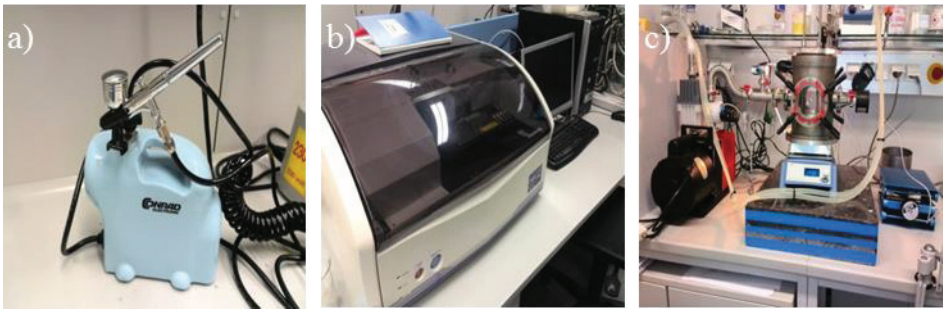


Figure 4.2. Matrix deposition a) Airbrush, b) Ink-jet printer, c) Sublimation device

4.2. Experimental Methods

All experiments and measurements were carried out in Institute of Chemical Technologies and Analytics in Technische Universitat Wien which is directed by Prof. Dr. Günter Allmaier. In addition, this project was conducted during the term of Erasmus+ mobility program under the supervision of Assoc. Prof. Dr. Martina Marchetti-Deschmann. Fresh-frozen mouse brains were obtained from Dr. Balazs Hegedüs in Department of Surgery located in Anna Spiegel Center of Translational Research. First, standard proteins including bovine serum albumin, ovalbumin and cytochrome c were prepared in 0,1% TFA-water solution. To test the effect of organic solvent, protein solution was mixed with 0,1% TFA containing 20%, 30%, 40%, 50%, 60% and 80% acetonitrile. Pepsin was dissolved in water to prevent the autolysis and released to acidic protein solution to activate the enzyme. Additionally, 2,5, and 10 fold diluted BSA solutions prepared from 0,1 mg/ml BSA were spotted onto a glass slide and enzyme with DHB matrix solution (20mg/ml) in 20% and 50% ethanol was applied by printing with 300µm raster to observe the reflection of relative quantity.

Fresh frozen mouse brain tissue was cut by scarpel and sectioned at 12 µm thickness using Thermo Scientific Cryostar NX50 at -12°C. Subsequently they were thaw-mounted onto cold ITO slides and put in desiccator until needed. Slides were washed with six step rinse protocol by immersing into 70% EtOH, 100% EtOH, Carnoy's fluid, 100% EtOH, H_2O, 100% EtOH respectively. Optical images were taken by microscope Leica DM2500M. Image composite editor (Microsoft) was used to stitch the images obtained from microscope. Alternatively regular flatbed scanner (HP Scanjet G4050) was used when resolution was not important.

DHB matrix containing pepsin was either applied by SHIMADZU ChIP 1000 Chemical Inkjet Printer or Conrad Profi Airbrush Set with 0.3mm screw nozzle. Printing parameters were as the following:

Quantity of drop: 100pL, interval: 4, total amont: 30nL, 250 µl center-to-center.

Spraying parameters were as the following:

200µl in every 6 minutes;15-20 cycles

Bruker UltrafleXtreme was used for the measurements, data was analyzed by FlexControl and FlexImaging softwares. Images were normalized according to total ion current. All the devices and materials are exactly same as in Figure 4.1 and Figure 4.2.

4.3. Results and Discussion

4.3.1. Feasibility Tests for Enzyme-Matrix Association

First, pepsin digestion at different conditions was carried out on standard proteins like ovalbumin, BSA, cytochrome c. Protein digestion solutions were kept simple throughout the pretesting that no protein denaturation or disulfide breakage was involved to the procedure to adapt it on-surface digestion. Briefly protein dissolved in ultrapure water and pepsin enzyme activated with acidic solution (0,1% TFA) were mixed and incubated for 30 minutes nevertheless many peptides signals were obtained and the both spectra shown similar pattern. Although optimum temperature for pepsin digestion is 37 °C as in the other digestive enzymes, pepsin's digestion ability at room temperature was first tested to eliminate the heating chamber. Most of the MALDI matrices are prepared at low pH for a good crystallization moreover proton donor requirement for the analysis at positive mode is solved concomitantly. However solubility of matrices (especially alpha-cyano cinnamic acid and sinapinic acid) in aqueous solution is rather low therefore organic solvents like acetonitrile, ethanol, methanol or acetone is added up to 70%. While organic solvent addition to an enzyme medium up to 10% has positive effect on catalytic activity of the enzyme, it has been demonstrated that this effect is dramatically reversed at high concentrations of organic solvent. It can be seen that acetonitrile content in the ovalbumin digestion medium which ranges from 0% to 40% indeed affects the digestion ability of the enzyme. On the other hand, digestion selectivity was seemingly altered judging by the changes of base peaks on the spectra, yet the enzyme was not entirely inactivated. At this point, it is unfortunately unavoidable to compromise either the enzyme digestion or matrix solubility in order to fuse these two sample preparation steps. In spite of the attenuated yield in peptide diversity, pepsin enzyme showed a lot promise thanks to its resistant nature. Lastly, possible inhibitory effect of the matrices to the catalytic activity of the enzyme was tested by adding DHB and HCCA matrices to the protein-enzyme solution and no significant inhibition was observed. Application of matrix containing pepsin method was later tested on mass spectrometry imaging glass slides displaying the relative quantitation of protein spots with increasing concentrations. All results regarding feasibility tests are given in Appendix E.

4.3.2. On-Surface Digestion Together with Matrix Application

In this study, two matrices, alpha-cyano-4-hydroxycinnamic acid (HCCA) and 2,5-dihydroxybenzoic acid (DHB) which are widely-used in peptide analysis were chosen. Among them, DHB matrix was more proper to combine with pepsin due to its solubility in water. The application of the matrix was carried out either with spraying or inkjet printing. Matrix application by spraying with an air brush is cheaper, simpler and more common method among the mass spectrometry imaging sample preparation methods however reproducibility is not as good as inkjet printing. On the other hand, inkjet printing takes longer time and suffers from clogging. HCCA matrix could not be used for inkjet printing because of clogging. Apart from that, sublimation method with DHB matrix was utilized for once in external control to eliminate the previously existing signals (Figure 4.3). Actually matrix sublimation is the superior matrix application method among all in terms of surface homogeneity however it was not feasible for enzyme incorporation since enzyme is quite a big molecule, approximately 35kDa. Main purpose of this study was to investigate the applicability of the suggested method therefore resolution or shot-to-shot distance were chosen moderate since we were not searching a particular analyte. Partial tissue area instead of the whole section was preferred to keep the analysis time short.

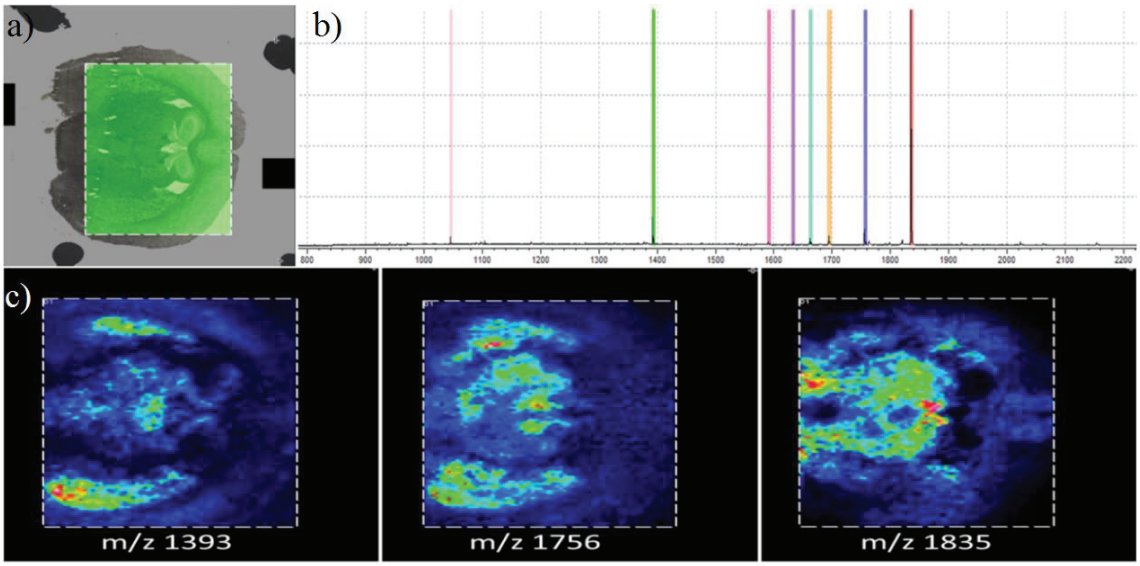


Figure 4.3. DHB matrix deposited with sublimation for the external control (no pepsin .....addition). a) MALDI imaging analysis frame (green area) on the optical .....image, b) Representative spectrum of the all accumulated spectra, c) .....Distribution of m/z values 1393, 1756 and 1835 respectively on the tissue

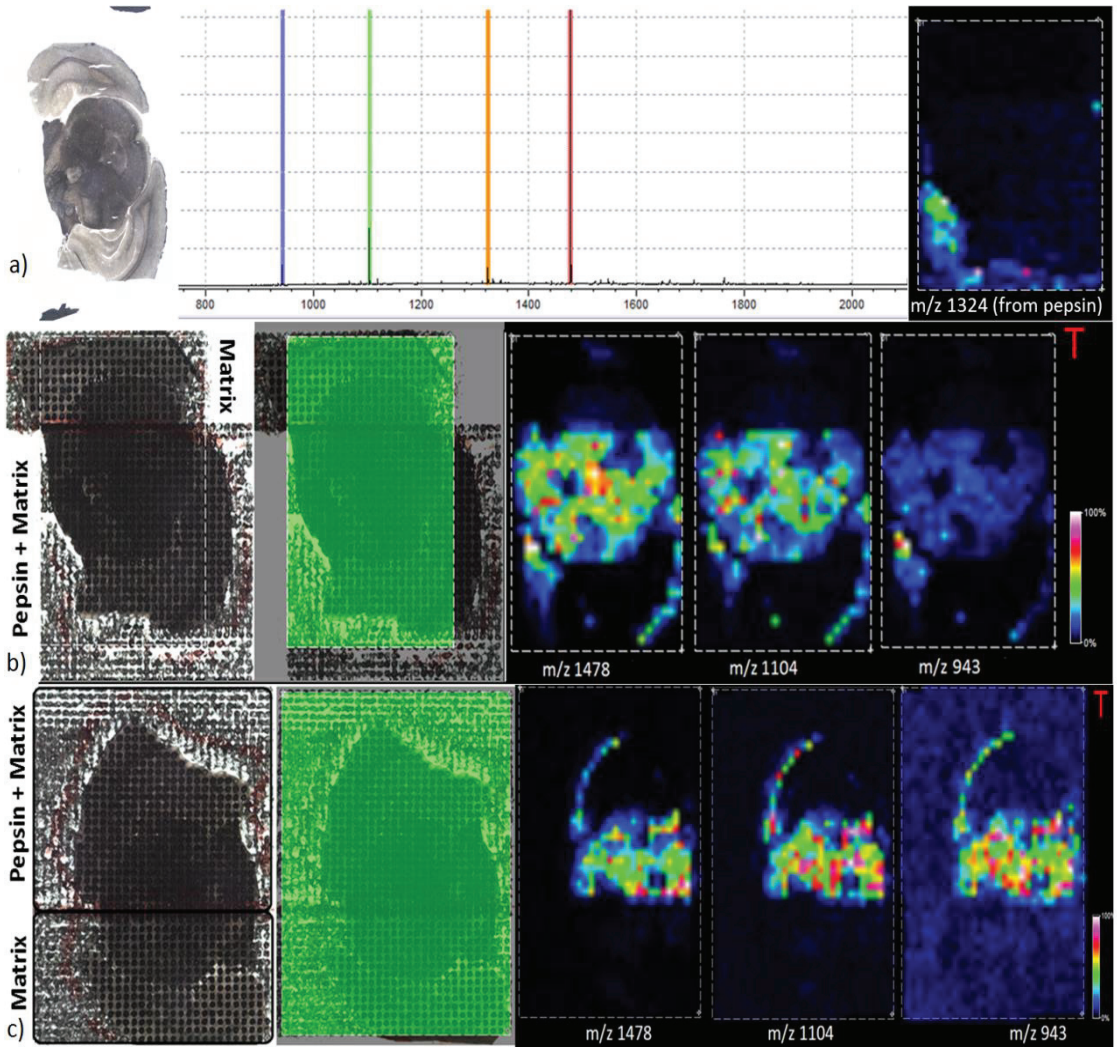


Figure 4.4. Deposition with inkjet printing. a) From left to right: Optical image before .....matrix printing, representative spectrum and distribution of m/z 1324 on the .....tissue, c and d) DHB+Pepsin application: (from left to right) Optical image .....after matrix printing, MALDI imaging analysis area (green) on the optical .....image and distribution of specific m/z values on the tissue section

Results acquired from DHB matrix application with inkjet printing is given in Figure 4.4. Stable depositing condition in a reasonable time was obtained with 25 mg/ml in 50% EtOH and 0.1%TFA by trial and error. Method was tested twice on two different tissue sections. In order to create an internal control, some part of the tissue was printed only by matrix and other part by matrix and pepsin. Both samples gave 3 significant signals at m/z 943, 1104 and 1478 Da showing the same image shape. Those signals were detected only on the part pepsin and matrix part. These signals give a particular shape on the image of the tissue which implies it might correspond to a distinct protein however source of these signal could not be identified by mass spectrometry or in-situ protein digestion of expected protein techniques.

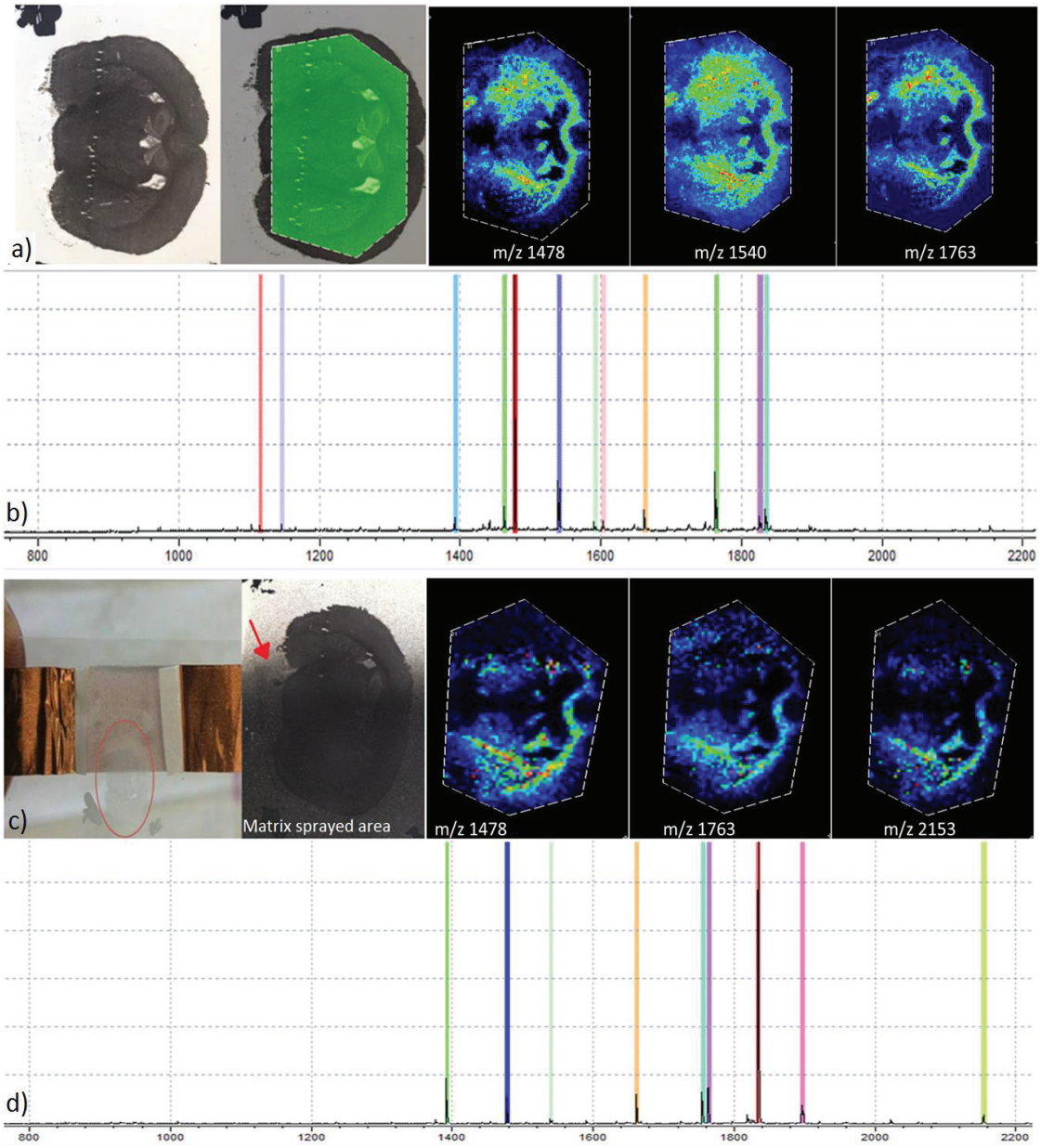


Figure 4.5. Deposition with spray coating. a) HCCA+Pepsin application: (from left to .....right) Optical image, MSI analysis frame, distribution of specific m/z .....signals, c) DHB+Pepsin application: camera image, optical image and .....distribution of specific m/z signals, b&d) Representative spectrum for .....accumulated spectra

Figure 4.5 shows the spectra and images obtained from matrix-enzyme applications by spraying with air brush. In order to create an internal control, some portion of the tissue section was covered by sticky tape in DHB matrix application. First matrix containing pepsin solution was sprayed then only matrix solution was sprayed to the rest of the tissue. Diffusion linage (red arrow in Figure 4.5, c) indicates that new signals are definitely arisen from the addition of pepsin however pre-existing signals detected earlier were also appeared on the spectrum.

4.4. Conclusion

In addition to detection and identification, localization of molecules can be analyzed with MALDI imaging technique by putting together the automatically acquired spectra from each spot position of a planar sample such as a thin tissue section. Low detection limit of high mass molecules including proteins generate problem on sensitivity as well as resolution therefore a digestive enzyme is introduced to tissue surface from the top layer before matrix deposition thus proteins are reduced to peptides and become more favorable analyte for the mass spectrometric measurement. However, requirement of an aqueous medium for the enzymatic reaction contradicts with the protection of tissue integrity due to diffusional effect. The essential enzyme of the mainstream protocols in proteomics, trypsin cleaves satisfactorily selective after the C-terminus of basic amino acids, arginine and lysine however its kinetic values do not allow to fast catalysis that longer incubation durations from 4 hours to overnight is applied. In addition, it is quite costly preotease and requires high pH around 8 and 37 °C temperature to work optimally. In this study, an enzyme which is catalytically active in acidic environment is mixed with matrix solution in order to drop two separate processes, enzyme and matrix applications to one step for simpler sample preparation of on-tissue digestion. Although pepsin does not digest very selectively as trypsin does, its low price and digestion speed even at room temperature make this enzyme quite convenient for on-tissue digestion. Besides, unlike other enzymes, pepsin has unique feature of digesting at low pH which enable us to interfuse it with MALDI matrix in the same solution. To keep the method more practical for on-tissue digestion, conditions for the enzyme were compromised to rather extreme for the enzyme such as high organic solvent content and room temperature incubation. Two alternative matrix deposition methods, printing and spraying were utilized. Furthermore two different matrices, 2,5-dihydroxybenzoic acid and α-cyano-4-hydroxycinnamic acid could be tested with spraying method and gave parallel results. Only a few possible signals were detected but they were reproducible. Particullarly, an unexceptional intense peak at m/z 1478 which has the potential to reflect a particular brain region was detected each time when pepsin was involved. This method holds the potential to be improved with further optimizations which might be utilized as a complementary approach to other enzymes or quantitative analysis in targeted proteomics.

CHAPTER 5

CONCLUSION

Traditional peptide sequencing method, Edman degradation offers notably accurate determination for the amino acid order in a peptide chain however it is not applicable to peptide mixtures. Additionally, it does not allow to the analysis of low abundance proteins and post-translational modifications. Thanks to the soft ionization sources and HPLC, an analytical technique mass spectrometry could meet the requirements of sensetive, specific and rapid analysis as well as characterization of PTMs and assignment of a peptide in a complex mixture as a result, scientists could attain to the protein lists containing from hundred to thousands proteins thus large-scale protein investigations have merged under the name of mass spectrometry-based proteomics. General topic of this thesis is MS-based proteomics. Each three experimental stages in the workflow including sample preparation, mass spectrometric data acquisition and computational calculations in bioinformatics individually affect to the final result. In chapter 1, procedures and principles of each stage are given and deduction of amino acid sequence from the peptide fragmentation is explained in detail. Although various sample preparation techniques and different type of mass spectrometers are available, two methodical approaches mentioned as gel-based and gel-free have gained wide acceptance. In this study, both methods were applied to same sample set in chapter 2 and chapter 3 respectively thus we investigated the problem from two aspects whose experimental designs are completely different. Accordingly, same topic were addressed in these two chapters which is a serious public health problem encountered spread by a vector namely sandfly. The causative agent of this infectious disease is a unicellular parasite from *Leishmania* genus therefore the disease takes the name of leishmaniasis referring to the parasite however it is classified in three clinical forms due to the different symptoms and outcomes. Visceral leishmaniasis has fatal consequences if untreated whereas cutaneous form is characterized by skin lesions. Variety in species is the main reason for clinical manifestations however this discrimination seems to disappear recently. In this context, it is being observed that species causing cutaneous leishmaniasis has started to cause visceral leishmaniasis too

and vice versa. Therefore mechanisms underlying this transition should be unraveled for the disease treatment and eradication.

In the gel-based part, 2D-SDS PAGE images of total protein mixture extracted from promastigote isolates of 15 different patients were compared by using 17 cm long, pH 3-10 IPG strips. 10 of these samples were obtained from patients infected with cutaneous leishmaniasis (CL) and the rest 5 samples were obtained from patients suffered from visceral leishmaniasis (VL). Samples of all visceral leishmaniasis and 5 samples of cutaneous leishmaniasis belong to *Leishmania infantum* and the other 5 samples of cutaneous leishmaniasis were originated from *Leishmania tropica* thus three different groups were generated. Comparison of the gels was carried out by eye detection on single run of fifteen different biological replicates. Striking changes as newly appeared spot or conversion from pale to dark color and vice versa between visceral and cutaneous groups were taken into consideration. For ease of seeing, some particular areas from each gel were brought close together creating the region of interests (ROIs). Only spots identified with in-gel digestion and MALDI-TOF/TOF MS with subsequent Mascot search were indicated with a number on the figures. From a broad perspective, gel images of CL isolates, belonging to either *L. infantum* or *L. tropica* group, were quite similar whereas those of VL isolates were distinguishable from them. Main differentiation between CL and VL isolates was observed as down regulation in many proteins on the gels of VL samples. Accordingly, proteins identified as putative dihydrolipoamide dehydrogenase, pyruvate kinase, putative succinly-coA:3-ketoacid-coenzyme A transferase and aspartate aminotransferase are observed explicitly down regulated in that specific region. In addition, putative aldose 1-epimerase and trypanothione reductase proteins seem down regulated in VL samples although not as notable as the formers. Finally, there is a noticeably new spot visible on the gels of CL samples. This distinct new spot was identified as superoxide dismutase and detected like a biomarker only on the samples of CL either caused by *L. infantum* or *L. tropica*. However this protein was determined on the gels of VL samples at different location as well as in shotgun results of all VL samples. Although it may not be counted as biomarker protein, apparently this protein underwent a structural change therefore further analysis targeting the superoxide dismutase should be considered.

In shotgun proteomics, number of identified proteins fluctuates due to the user defined nature of peptide database searching. In the introduction of chapter 3, the reasons and consequences are discussed on the results of alternative settings and

bioinformatics tool options. For the gel-free part, 9 aliquots out of 15 samples were again sorted to three groups thus each group contained 3 samples. Six of these samples were belong to the same species, *Leishmania infantum* therefore comparison in these samples were more meaningful since expressed proteins were transcribed from the same gene source. Accordingly, we focused on the discrimination between cutaneous and visceral leishmaniasis samples of *Leishmania infantum*. For the differential protein recognition, we stipulated the presence in at least two samples of either cutaneous or visceral group. Since presence in all three samples of the group is more pronounced, differential proteins in *Leishmania infantum* were graded to certain and possible differential proteins. In order to obtain an overall perspective, all nine samples in the shotgun proteomics part were searched by using six different bioinformatics tools thus certainty of any particular differential protein could be validated too (Table 3.5). In this context, acidocalcisomal pyrophosphatase, putative glutamamyl carboxy peptidase and an uncharacterized protein with the code A4ICR0 exceptionally stand out from the others by holding certainty and complete validation. Flavoprotein subunit-like protein, putative eukaryotic release factor 3, putative lanosterol 14-alpha-demethylase, uncharacterized protein (A4HTX5), cop-coated vesicle membrane protein erv25, ORF2 and proline dehydrogenase also grab attention. Remarkably, these proteins and also some others show up when we involve the cutaneous leishmaniasis samples of *Leishmania tropica*. In conclusion, expression of these proteins must be crucial for the parasites causing to cutaneous leishmaniasis regardless the species therefore promising response may be received if these proteins are targeted for the treatment of only cutaneous leishmaniasis. On the other side, there are only a few differential proteins detected for the visceral leishmaniasis samples. I/6 autoantigen-like protein, zinc transporter-like protein and a particular Histon H2A deserve consideration. Apart from that, differential proteins from the searches with the database NCBI_L.genus are also determined and provided however discrimination based on the species is beyond the scope of this study. Since this database includes the protein sequences of all species, we do not expect a good correlation with the small database of UniProt_L.infantum, yet there are a few common differential proteins in the samples of cutaneous leishmaniasis. Although, results of gel-based method are not coherent with the results of gel-free method, a combinatorial interpretation can provide deeper perception and insight. For instance, expression of the superoxide dismutase protein in the visceral leishmaniasis is approved by gel-free analysis thus that distinct spot observed only on the gels of

cutaneous samples should be considered as alteration rather than presence. As a future direction, increasing the number of samples by including new patients, of course, renders this investigation more comprehensive. Particularly, application of the same experiments to the isolates obtained from patients with visceral leishmaniasis which is caused by *Leishmania tropica* will enable us to precisely associate those differential proteins to tropism in leishmanial.

Unlike previous chapters, subject is not related with proteomic analysis in chapter 4, yet it is about mass spectrometric analysis of proteins. Mass spectrometry imaging (MSI) is a confirmation to the versatility of mass spectrometry in which spatial distribution of unknown compounds can be attained without the need of any tracer molecule. In short, accumulated mass spectra of a two dimensional thin surface such as tissue section is used to generate an image thus any m/z signal can be mapped on this framed area where relative intensity is reflected as color scale. Sensitivity problem in the protein analysis is an obstacle to construct a well resolved image therefore proteins are converted to more detectable smaller units, peptides with proteases. Although this application help to overcome this issue to some extent, an additional digestion step prior to matrix application contains the risk of losing the genuine localization of analytes. Therefore we suggested a sample preparation method which combines consecutive applications of enzyme and matrix. For this purpose, we added and exceptional enzyme which is catalytically active at low pH is added to the matrix solution and applied onto tissue in a single step. Pepsin is an enzyme both affordable and highly reactive even under the poor conditions. In our preliminary results, we obtained new m/z signals whose occurrences were associated with pepsin addition however they could not be identified. In conclusion, this combined procedure has potential for further development.

REFERENCES

1. de Hoffmann, E.; Stroobant, V., *Mass Spectrometry: Principles and Applications, 3rd Edition*. Wiley-Interscience: 2007.

2. Thomson, J. J., Cathode Rays. *Philosophical Magazine* **2010,** *90* (sup1), 25-29.

3. Millikan, R. A., On the Elementary Electrical Charge and the Avogadro Constant. *Physical Review* **1913,** *2* (2), 109-143.

4. Dempster, A. J., A new Method of Positive Ray Analysis. *Physical Review* **1918,** *11* (4), 316-325.

5. Aston, F. W., LXXIV. A positive ray spectrograph. *The London, Edinburgh, and Dublin Philosophical Magazine and Journal of Science* **1919,** *38* (228), 707-714.

6. Cameron, A. E.; Jr., D. F. E., An Ion ``Velocitron''. *Review of Scientific Instruments* **1948,** *19* (9), 605-607.

7. Hipple, J. A.; Sommer, H.; Thomas, H. A., A Precise Method of Determining the Faraday by Magnetic Resonance. *Physical Review* **1949,** *76* (12), 1877-1878.

8. Johnson, E. G.; Nier, A. O., Angular Aberrations in Sector Shaped Electromagnetic Lenses for Focusing Beams of Charged Particles. *Physical Review* **1953,** *91* (1), 10-17.

9. Paul, W. S., Helmut, Notizen: Ein neues Massenspektrometer ohne Magnetfeld. *Zeitschrift für Naturforschung A* **1953,** *8* (7), 448-450.

10. McLafferty, F. W., Mass Spectrometry in Chemical Research and Production. *Appl. Spectrosc.* **1957,** *11* (4), 148-156.

11. Gohlke, R. S., Time-of-Flight Mass Spectrometry and Gas-Liquid Partition Chromatography. *Analytical Chemistry* **1959,** *31* (4), 535-541.

12. Munson, M. S. B.; Field, F. H., Chemical Ionization Mass Spectrometry. I. General Introduction. *Journal of the American Chemical Society* **1966,** *88* (12), 2621-2630.

13. McLafferty, F. W.; Bryce, T. A., Metastable-ion characteristics: characterization of isomeric molecules. *Chemical Communications (London)* **1967,** (23), 1215-1217.

14. Jennings, K. R., Collision-induced decompositions of aromatic molecular ions. *International Journal of Mass Spectrometry and Ion Physics* **1968,** *1* (3), 227-235.

15. Beckey, H. D., Field desorption mass spectrometry: A technique for the study of thermally unstable substances of low volatility. *International Journal of Mass Spectrometry and Ion Physics* **1969,** *2* (6), 500-502.

16. Mamyrin, B. A.; Karatev, V. I.; Shmikk, D. V.; Zagulin, V. A., The mass-reflectron, a new nonmagnetic time-of-flight mass spectrometer with high resolution. *Experimental and Theoretical Physics* **1972,** *37* (1).

17. Horning, E. C.; Carroll, D. I.; Dzidic, I.; Haegele, K. D.; Horning, M. G.; Stillwell, R. N., Atmospheric pressure ionization (API) mass spectrometry. Solvent-mediated ionization of samples introduced in solution and in a liquid chromatograph effluent stream. *J Chromatogr Sci* **1974,** *12* (11), 725-9.

18. Arpino, P.; Baldwin, M. A.; McLafferty, F. W., Liquid chromatography-mass spectrometry. II—continuous monitoring. *Biomedical Mass Spectrometry* **1974,** *1* (1), 80-82.

19. Comisarow, M. B.; Marshall, A. G., Fourier transform ion cyclotron resonance spectroscopy. *Chemical Physics Letters* **1974,** *25* (2), 282-283.

20. Macfarlane, R.; Torgerson, D., Californium-252 plasma desorption mass spectroscopy. *Science* **1976,** *191* (4230), 920-925.

21. Yost, R. A.; Enke, C. G., Selected ion fragmentation with a tandem quadrupole mass spectrometer. *Journal of the American Chemical Society* **1978,** *100* (7), 2274-2275.

22. Houk, R. S.; Fassel, V. A.; Flesch, G. D.; Svec, H. J.; Gray, A. L.; Taylor, C. E., Inductively coupled argon plasma as an ion source for mass spectrometric determination of trace elements. *Analytical Chemistry* **1980,** *52* (14), 2283-2289.

23. Karas, M.; Bachmann, D.; Bahr, U.; Hillenkamp, F., Matrix-assisted ultraviolet laser desorption of non-volatile compounds. *International Journal of Mass Spectrometry and Ion Processes* **1987,** *78*, 53-68.

24. Tanaka, K.; Waki, H.; Ido, Y.; Akita, S.; Yoshida, Y.; Yoshida, T.; Matsuo, T., Protein and polymer analyses up to m/z 100 000 by laser ionization time-of-flight mass spectrometry. *Rapid Communications in Mass Spectrometry* **1988,** *2* (8), 151-153.

25. Olivares, J. A.; Nguyen, N. T.; Yonker, C. R.; Smith, R. D., On-line mass spectrometric detection for capillary zone electrophoresis. *Analytical Chemistry* **1987,** *59* (8), 1230-1232.

26. Fenn, J.; Mann, M.; Meng, C.; Wong, S.; Whitehouse, C., Electrospray ionization for mass spectrometry of large biomolecules. *Science* **1989,** *246* (4926), 64-71.

27. Wasinger, V. C.; Cordwell, S. J.; Cerpa-Poljak, A.; Yan, J. X.; Gooley, A. A.; Wilkins, M. R.; Duncan, M. W.; Harris, R.; Williams, K. L.; Humphery-Smith, I., Progress with gene-product mapping of the Mollicutes: Mycoplasma genitalium. *Electrophoresis* **1995,** *16* (7), 1090-4.

28. Wilm, M.; Mann, M., Analytical Properties of the Nanoelectrospray Ion Source. *Analytical Chemistry* **1996,** *68* (1), 1-8.

29. Makarov, A., Electrostatic Axially Harmonic Orbital Trapping: A High-Performance Technique of Mass Analysis. *Analytical Chemistry* **2000,** *72* (6), 1156-1162.

30. Hosseini, S.; Martinez-Chapa, S. O., Principles and Mechanism of MALDI-ToF-MS Analysis. In *Fundamentals of MALDI-ToF-MS Analysis: Applications in Bio-diagnosis, Tissue Engineering and Drug Delivery*, Springer Singapore: Singapore, 2017; pp 1-19.

31. Konermann, L.; Vahidi, S.; Sowole, M. A., Mass Spectrometry Methods for Studying Structure and Dynamics of Biological Macromolecules. *Analytical Chemistry* **2014,** *86* (1), 213-232.

32. Wilm, M., Principles of Electrospray Ionization. *Molecular & Cellular Proteomics : MCP* **2011,** *10* (7), M111.009407.

33. Wysocki, V. H.; Resing, K. A.; Zhang, Q.; Cheng, G., Mass spectrometry of peptides and proteins. *Methods* **2005,** *35* (3), 211-222.

34. Griffiths, W. J., Nanospray mass spectrometry in protein and peptide chemistry. In *Proteomics in Functional Genomics: Protein Structure Analysis*, Jollès, P.; Jörnvall, H., Eds. Birkhäuser Basel: Basel, 2000; pp 69-79.

35. Wasinger, V. C.; Zeng, M.; Yau, Y., Current Status and Advances in Quantitative Proteomic Mass Spectrometry. *International Journal of Proteomics* **2013,** *2013*, 12.

36. Watson, J. T.; Sparkman, O. D., *Introduction to Mass Spectrometry: Instrumentation, Applications and Strategies for Data Interpretation*. Fourth ed.; John Wiley & Sons, Ltd: 2008.

37. Johnson, J. V.; Yost, R. A.; Kelley, P. E.; Bradford, D. C., Tandem-in-space and tandem-in-time mass spectrometry: triple quadrupoles and quadrupole ion traps. *Analytical Chemistry* **1990,** *62* (20), 2162-2172.

38. Strupat, K., Molecular weight determination of peptides and proteins by ESI and MALDI. *Methods Enzymol* **2005,** *405*, 1-36.

39. O'Farrell, P. H., High Resolution Two-Dimensional Electrophoresis of Proteins. *The Journal of biological chemistry* **1975,** *250* (10), 4007-4021.

40. Pomastowski, P.; Buszewski, B., Two-dimensional gel electrophoresis in the light of new developments. *TrAC Trends in Analytical Chemistry* **2014,** *53*, 167-177.

41. Rabilloud, T.; Lelong, C., Two-dimensional gel electrophoresis in proteomics: A tutorial. *Journal of Proteomics* **2011,** *74* (10), 1829-1841.

42. Clevett, K. J., High-performance liquid chromatography. *Bioprocess Technol* **1990,** *6*, 47-73.

43. Klein, J. B.; Thongboonkerd, V., Overview of proteomics. *Contrib.Nephrol.* **2004,** *141*, 1-10.

44. Liebler, D., *Introduction to Proteomics Tools for the New Biology*. Humana Press: 2002.

45. Tsiatsiani, L.; Heck, A. J. R., Proteomics beyond trypsin. *The FEBS Journal* **2015,** *282* (14), 2612-2626.

46. León, I. R.; Schwämmle, V.; Jensen, O. N.; Sprenger, R. R., Quantitative Assessment of In-solution Digestion Efficiency Identifies Optimal Protocols for Unbiased Protein Analysis. *Molecular & Cellular Proteomics : MCP* **2013,** *12* (10), 2992-3005.

47. Gundry, R. L.; White, M. Y.; Murray, C. I.; Kane, L. A.; Fu, Q.; Stanley, B. A.; Van Eyk, J. E., Preparation of Proteins and Peptides for Mass Spectrometry Analysis in a Bottom-Up Proteomics Workflow. *Current protocols in molecular biology / edited by Frederick M. Ausubel ... [et al.]* **2009,** *CHAPTER*, Unit10.25-Unit10.25.

48. Messing, J.; Crea, R.; Seeburg, P. H., A System For Shotgun DNA Sequencing. *Nucleic Acids Res.* **1981,** *9* (2), 309-321.

49. Hunt, D. F.; Yates, J. R.; Shabanowitz, J.; Winston, S.; Hauer, C. R., Protein sequencing by tandem mass spectrometry. *Proceedings of the National Academy of Sciences* **1986,** *83* (17), 6233-6237.

50. Zhang, Y.; Fonslow, B. R.; Shan, B.; Baek, M.-C.; Yates, J. R., Protein Analysis by Shotgun/Bottom-up Proteomics. *Chemical Reviews* **2013,** *113* (4), 2343-2394.

51. Pappin, D. J. C.; Hojrup, P.; Bleasby, A. J., Rapid identification of proteins by peptide-mass fingerprinting. *Current Biology* **1993,** *3* (6), 327-332.

52. Nesvizhskii, A. I.; Aebersold, R., Interpretation of shotgun proteomic data: the protein inference problem. *Mol Cell Proteomics* **2005,** *4* (10), 1419-40.

53. Nagele, E.; Vollmer, M.; Horth, P.; Vad, C., 2D-LC/MS techniques for the identification of proteins in highly complex mixtures. *Expert Rev Proteomics* **2004,** *1* (1), 37-46.

54. Yamana, R.; Iwasaki, M.; Wakabayashi, M.; Nakagawa, M.; Yamanaka, S.; Ishihama, Y., Rapid and Deep Profiling of Human Induced Pluripotent Stem Cell Proteome by One-shot NanoLC–MS/MS Analysis with Meter-scale Monolithic Silica Columns. *Journal of Proteome Research* **2013,** *12* (1), 214-221.

55. Iwasaki, M.; Miwa, S.; Ikegami, T.; Tomita, M.; Tanaka, N.; Ishihama, Y., One-Dimensional Capillary Liquid Chromatographic Separation Coupled with Tandem Mass Spectrometry Unveils the Escherichia coli Proteome on a Microarray Scale. *Analytical Chemistry* **2010,** *82* (7), 2616-2620.

56. Johnson, R. S.; Martin, S. A.; Biemann, K.; Stults, J. T.; Watson, J. T., Novel fragmentation process of peptides by collision-induced decomposition in a tandem mass spectrometer: differentiation of leucine and isoleucine. *Analytical Chemistry* **1987,** *59* (21), 2621-2625.

57. http://www.matrixscience.com/help/fragmentation_help.html.

58. Edman, P. H., Erik; Sillén, Lars Gunnar; Kinell, Per-Olof, Method for determination of the amino acid sequence in peptides. *Acta Chemica Scandinavica* **1950,** *4*, 283–293.

59. Aebersold, R.; Mann, M., Mass spectrometry-based proteomics. *Nature* **2003,** *422*, 198.

60. Hjernø, K., Protein Identification by Peptide Mass Fingerprinting. In *Mass Spectrometry Data Analysis in Proteomics*, Matthiesen, R., Ed. Humana Press: Totowa, NJ, 2007; pp 61-75.

61. Elias, J. E.; Gygi, S. P., Target-Decoy Search Strategy for Mass Spectrometry-Based Proteomics. *Methods in molecular biology (Clifton, N.J.)* **2010,** *604*, 55-71.

62. Ishihama, Y.; Oda, Y.; Tabata, T.; Sato, T.; Nagasu, T.; Rappsilber, J.; Mann, M., Exponentially modified protein abundance index (emPAI) for estimation of absolute protein amount in proteomics by the number of sequenced peptides per protein. *Mol Cell Proteomics* **2005,** *4* (9), 1265-72.

63. Organization, W. W. H., A global brief on vector-borne diseases. **2014**.

64. Organization, W. W. H., Control of the leishmaniasis: report of a meeting of the WHO Expert Committee on the Control of Leishmaniases. 2010.

65. Organization, W. W. H., *Integrating neglected tropical diseases into global health and development: fourth WHO report on neglected tropical diseases*. 2017.

66. Rose, K.; Curtis, J.; Baldwin, T.; Mathis, A.; Kumar, B.; Sakthianandeswaren, A.; Spurck, T.; Low Choy, J.; Handman, E., Cutaneous leishmaniasis in red kangaroos: isolation and characterisation of the causative organisms. *International Journal for Parasitology* **2004,** *34* (6), 655-664.

67. Sources: ECDC, W., Claborn, David M. 2010. J Glob Infect Dis.2(2): 127–134. Usual Suspects: 004 Phlebotomine Sand Fly. https://www.isglobal.org/en/publication/-/asset_publisher/ljGAMKTwu9m4/content/sospechosos-habituales-005-flebotomo.

68. Lindoso, J. A. L.; Cunha, M. A.; Queiroz, I. T.; Moreira, C. H. V., Leishmaniasis–HIV coinfection: current challenges. *HIV/AIDS (Auckland, N.Z.)* **2016,** *8*, 147-156.

69. Organization, W. H. Clinical forms of the leishmaniases. http://www.who.int/leishmaniasis/disease/clinical_forms_leishmaniases/en/.

70. Chappuis, F.; Sundar, S.; Hailu, A.; Ghalib, H.; Rijal, S.; Peeling, R. W.; Alvar, J.; Boelaert, M., Visceral leishmaniasis: what are the needs for diagnosis, treatment and control? *Nat Rev Microbiol* **2007,** *5* (11), 873-82.

71. Meireles, C. B.; Maia, L. C.; Soares, G. C.; Teodoro, I. P. P.; Gadelha, M. d. S. V.; da Silva, C. G. L.; de Lima, M. A. P., Atypical presentations of cutaneous leishmaniasis: A systematic review. *Acta Tropica* **2017,** *172*, 240-254.

72. Hashemi, S. A.; Badirzadeh, A.; Sabzevari, S.; Nouri, A.; Seyyedin, M., First case report of atypical disseminated cutaneous leishmaniasis in an opium abuser in Iran. *Revista do Instituto de Medicina Tropical de São Paulo* **2018,** *60*.

73. Coelho, A. C.; Leprohon, P.; Ouellette, M., Generation of Leishmania Hybrids by Whole Genomic DNA Transformation. *PLOS Neglected Tropical Diseases* **2012,** *6* (9), e1817.

74. Peacock, C. S.; Seeger, K.; Harris, D.; Murphy, L.; Ruiz, J. C.; Quail, M. A.; Peters, N.; Adlem, E.; Tivey, A.; Aslett, M.; Kerhornou, A.; Ivens, A.; Fraser, A.; Rajandream, M.-A.; Carver, T.; Norbertczak, H.; Chillingworth, T.; Hance, Z.; Jagels, K.; Moule, S.; Ormond, D.; Rutter, S.; Squares, R.; Whitehead, S.; Rabbinowitsch, E.; Arrowsmith, C.; White, B.; Thurston, S.; Bringaud, F.; Baldauf, S. L.; Faulconbridge, A.; Jeffares, D.; Depledge, D. P.; Oyola, S. O.; Hilley, J. D.; Brito, L. O.; Tosi, L. R. O.; Barrell, B.; Cruz, A. K.; Mottram, J. C.; Smith, D. F.; Berriman, M., Comparative genomic analysis of three Leishmania species that cause diverse human disease. *Nature genetics* **2007,** *39* (7), 839-847.

75. Zhang, W. W.; Ramasamy, G.; McCall, L.-I.; Haydock, A.; Ranasinghe, S.; Abeygunasekara, P.; Sirimanna, G.; Wickremasinghe, R.; Myler, P.; Matlashewski, G., Genetic Analysis of Leishmania donovani Tropism Using a Naturally Attenuated Cutaneous Strain. *PLOS Pathogens* **2014,** *10* (7), e1004244.

76. Prieto Barja, P.; Pescher, P.; Bussotti, G.; Dumetz, F.; Imamura, H.; Kedra, D.; Domagalska, M.; Chaumeau, V.; Himmelbauer, H.; Pages, M.; Sterkers, Y.; Dujardin, J. C.; Notredame, C.; Spath, G. F., Haplotype selection as an adaptive mechanism in the protozoan pathogen Leishmania donovani. *Nat Ecol Evol* **2017,** *1* (12), 1961-1969.

77. Mehmet Salih Gürel, Y. Y., M. Kirami Ölgen, Yusuf Özbel, Türkiye'de Kutanöz Leishmaniasisin Durumu. *Turkiye Parazitol Derg* **2012**.

78. Özkeklikçi, A.; Karakuş, M.; Özbel, Y.; Töz, S., The new situation of cutaneous leishmaniasis after Syrian civil war in Gaziantep city, Southeastern region of Turkey. *Acta Tropica* **2017,** *166*, 35-38.

79. Ahmet, Ö.; Gülnaz, Ç.; Soner, U.; Mehmet, H.; Günaştı, T. S.; Fulya, O.; Fadile, Z.; Cumhur, G.; İpek, Ö.; Mehmet, K.; Seray, T.; Özgür, K.; Işın, A.; Ayşegül, E.; Dilek, G.; Çağla, K.; İbrahim, Ç.; Patrick, B.; Francine, P.; Tanıl, K.; Yusuf,

Ö., Leishmaniasis in Turkey: first clinical isolation of Leishmania major from 18 autochthonous cases of cutaneous leishmaniasis in four geographical regions. *Tropical Medicine & International Health* **2016,** *21* (6), 783-791.

80. Ok, Ü. Z.; Balcıoğlu, İ. C.; Taylan Özkan, A.; Özensoy, S.; Özbel, Y., Leishmaniasis in Turkey. *Acta Tropica* **2002,** *84* (1), 43-48.

81. Özbilgin, A.; Harman, M.; Karakuş, M.; Bart, A.; Töz, S.; Kurt, Ö.; Çavuş, İ.; Polat, E.; Gündüz, C.; Van Gool, T.; Özbel, Y., Leishmaniasis in Turkey: Visceral and cutaneous leishmaniasis caused by Leishmania donovani in Turkey. *Acta Tropica* **2017,** *173*, 90-96.

82. Ghosh, S.; Goswami, S.; Adhya, S., Role of superoxide dismutase in survival of Leishmania within the macrophage. *Biochemical Journal* **2003,** *369* (Pt 3), 447-452.

83. Mukherjee, S.; Bandyapadhyay, R.; Basu, M. K., Leishmania donovani: Superoxide dismutase level in infected macrophages. *Bioscience Reports* **1988,** *8* (2), 131-137.

84. Mittra, B.; Laranjeira-Silva, M. F.; Miguel, D. C.; Perrone Bezerra de Menezes, J.; Andrews, N. W., The iron-dependent mitochondrial superoxide dismutase SODA promotes Leishmania virulence. *Journal of Biological Chemistry* **2017,** *292* (29), 12324-12338.

85. Yao, C.; Li, Y.; Donelson, J. E.; Wilson, M. E., Proteomic examination of Leishmania chagasi plasma membrane proteins: contrast between avirulent and virulent (metacyclic) parasite forms. *Proteomics. Clinical applications* **2010,** *4* (1), 4-16.

86. Kulshrestha, A.; Sharma, V.; Singh, R.; Salotra, P., Comparative transcript expression analysis of miltefosine-sensitive and miltefosine-resistant Leishmania donovani. *Parasitology Research* **2014,** *113* (3), 1171-1184.

87. Hajjaran, H.; Mousavi, P.; Burchmore, R.; Mohebali, M.; Mohammadi Bazargani, M.; Hosseini Salekdeh, G.; Kazemi-Rad, E.; Khoramizadeh, M. R., Comparative Proteomic Profiling of Leishmania tropica: Investigation of a Case Infected with Simultaneous Cutaneous and Viscerotropic Leishmaniasis by 2-Dimentional Electrophoresis and Mass Spectrometry. *Iranian Journal of Parasitology* **2015,** *10* (3), 366-380.

88. Vincent, I. M.; Racine, G.; Légaré, D.; Ouellette, M., Mitochondrial Proteomics of Antimony and Miltefosine Resistant Leishmania infantum. *Proteomes* **2015,** *3* (4), 328-346.

89. Daifalla, N. S.; Bayih, A. G.; Gedamu, L., Leishmania donovani recombinant iron superoxide dismutase B1 protein in the presence of TLR-based adjuvants induces partial protection of BALB/c mice against Leishmania major infection. *Experimental Parasitology* **2012,** *131* (3), 317-324.

90. Danesh-Bahreini, M. A.; Shokri, J.; Samiei, A.; Kamali-Sarvestani, E.; Barzegar-Jalali, M.; Mohammadi-Samani, S., Nanovaccine for leishmaniasis:

preparation of chitosan nanoparticles containing Leishmania superoxide dismutase and evaluation of its immunogenicity in BALB/c mice. *International Journal of Nanomedicine* **2011,** *6*, 835-842.

91. Daifalla, N. S.; Bayih, A. G.; Gedamu, L., Immunogenicity of Leishmania donovani iron superoxide dismutase B1 and peroxidoxin 4 in BALB/c mice: The contribution of Toll-like receptor agonists as adjuvant. *Experimental Parasitology* **2011,** *129* (3), 292-298.

92. Santos, L. O.; Garcia-Gomes, A. S.; Catanho, M.; Sodre, C. L.; Santos, A. L. S.; Branquinha, M. H.; d'Avila-Levy, C. M., Aspartic Peptidases of Human Pathogenic Trypanosomatids: Perspectives and Trends for Chemotherapy. *Current Medicinal Chemistry* **2013,** *20* (25), 3116-3133.

93. Promastigote. https://msu.edu/course/zol/316/lsppscope.htm.

94. Amastigote. http://www.cdc.gov/parasites/leishmaniasis/health_professionals/index.html.

95. El Fakhry, Y.; Ouellette, M.; Papadopoulou, B., A proteomic approach to identify developmentally regulated proteins in Leishmania infantum. *Proteomics* **2002,** *2* (8), 1007-1017.

96. Tsigankov, P.; Gherardini, P. F.; Helmer-Citterich, M.; Zilberstein, D. A. N., What has proteomics taught us about Leishmania development? *Parasitology* **2012,** *139* (9), 1146-1157.

97. Nugent, P. G.; Karsani, S. A.; Wait, R.; Tempero, J.; Smith, D. F., Proteomic analysis of Leishmania mexicana differentiation. *Molecular and Biochemical Parasitology* **2004,** *136* (1), 51-62.

98. Bente, M.; Harder, S.; Wiesgigl, M.; Heukeshoven, J.; Gelhaus, C.; Krause, E.; Clos, J.; Bruchhaus, I., Developmentally induced changes of the proteome in the protozoan parasite Leishmania donovani. *Proteomics* **2003,** *3* (9), 1811-1829.

99. Walker, J.; Vasquez, J.-J.; Gomez, M. A.; Drummelsmith, J.; Burchmore, R.; Girard, I.; Ouellette, M., Identification of developmentally-regulated proteins in Leishmania panamensis by proteome profiling of promastigotes and axenic amastigotes. *Molecular and Biochemical Parasitology* **2006,** *147* (1), 64-73.

100. A H Fairlamb, a.; Cerami, A., Metabolism and Functions of Trypanothione in the Kinetoplastida. *Annual Review of Microbiology* **1992,** *46* (1), 695-729.

101. Armin, S.; Krauth-Siegel, R. L., Enzymes of the Trypanothione Metabolism as Targets for Antitrypanosomal Drug Development. *Current Topics in Medicinal Chemistry* **2002,** *2* (11), 1239-1259.

102. Biyani, N.; Singh, A. K.; Mandal, S.; Chawla, B.; Madhubala, R., Differential expression of proteins in antimony-susceptible and -resistant isolates of Leishmania donovani. *Molecular and Biochemical Parasitology* **2011,** *179* (2), 91-99.

103. Chakravarty, J.; Sundar, S., Drug Resistance in Leishmaniasis. *Journal of Global Infectious Diseases* **2010,** *2* (2), 167-176.

104. Alcolea, P. J.; Tuñón, G. I. L.; Alonso, A.; García-Tabares, F.; Ciordia, S.; Mena, M. C.; Campos, R. N. S.; Almeida, R. P.; Larraga, V., Differential protein abundance in promastigotes of nitric oxide-sensitive and resistant Leishmania chagasi strains. *PROTEOMICS – Clinical Applications* **2016,** *10* (11), 1132-1146.

105. Moreira, D. d. S.; Pescher, P.; Laurent, C.; Lenormand, P.; Späth, G. F.; Murta, S. M. F., Phosphoproteomic analysis of wild-type and antimony-resistant Leishmania braziliensis lines by 2D-DIGE technology. *Proteomics* **2015,** *15* (17), 2999-3019.

106. Brotherton, M.-C.; Bourassa, S.; Légaré, D.; Poirier, G. G.; Droit, A.; Ouellette, M., Quantitative proteomic analysis of amphotericin B resistance in Leishmania infantum. *International Journal for Parasitology: Drugs and Drug Resistance* **2014,** *4* (2), 126-132.

107. de Rezende, E.; Kawahara, R.; Peña, M. S.; Palmisano, G.; Stolf, B. S., Quantitative proteomic analysis of amastigotes from Leishmania (L.) amazonensis LV79 and PH8 strains reveals molecular traits associated with the virulence phenotype. *PLOS Neglected Tropical Diseases* **2017,** *11* (11), e0006090.

108. Kumar, A.; Misra, P.; Sisodia, B.; Shasany, A. K.; Sundar, S.; Dube, A., Proteomic analyses of membrane enriched proteins of Leishmania donovani Indian clinical isolate by mass spectrometry. *Parasitology International* **2015,** *64* (4), 36-42.

109. Santarém, N.; Racine, G.; Silvestre, R.; Cordeiro-da-Silva, A.; Ouellette, M., Exoproteome dynamics in Leishmania infantum. *Journal of Proteomics* **2013,** *84*, 106-118.

110. Morales, M. A.; Watanabe, R.; Dacher, M.; Chafey, P.; Osorio y Fortéa, J.; Scott, D. A.; Beverley, S. M.; Ommen, G.; Clos, J.; Hem, S.; Lenormand, P.; Rousselle, J.-C.; Namane, A.; Späth, G. F., Phosphoproteome dynamics reveal heat-shock protein complexes specific to the Leishmania donovani infectious stage. *Proceedings of the National Academy of Sciences of the United States of America* **2010,** *107* (18), 8381-8386.

111. Tsigankov, P.; Gherardini, P. F.; Helmer-Citterich, M.; Späth, G. F.; Zilberstein, D., Phosphoproteomic Analysis of Differentiating Leishmania Parasites Reveals a Unique Stage-Specific Phosphorylation Motif. *Journal of Proteome Research* **2013,** *12* (7), 3405-3412.

112. Sundar, S.; Singh, B., Understanding Leishmania parasites through proteomics and implications for the clinic. *Expert Review of Proteomics* **2018,** *15* (5), 371-390.

113. Veras, P. S. T.; Bezerra de Menezes, J. P., Using Proteomics to Understand How Leishmania Parasites Survive inside the Host and Establish Infection. *International Journal of Molecular Sciences* **2016,** *17* (8), 1270.

114. Alcolea, P. J.; Alonso, A.; García-Tabares, F.; Mena, M. d. C.; Ciordia, S.; Larraga, V., Proteome profiling of the growth phases of Leishmania pifanoi promastigotes in axenic culture reveals differential abundance of immunostimulatory proteins. *Acta Tropica* **2016,** *158*, 240-247.

115. Alves-Ferreira, E. V. C.; Toledo, J. S.; De Oliveira, A. H. C.; Ferreira, T. R.; Ruy, P. C.; Pinzan, C. F.; Santos, R. F.; Boaventura, V.; Rojo, D.; López-Gonzálvez, Á.; Rosa, J. C.; Barbas, C.; Barral-Netto, M.; Barral, A.; Cruz, A. K., Differential Gene Expression and Infection Profiles of Cutaneous and Mucosal Leishmania braziliensis Isolates from the Same Patient. *PLOS Neglected Tropical Diseases* **2015,** *9* (9), e0004018.

116. Lye, L.-F.; Owens, K.; Shi, H.; Murta, S. M. F.; Vieira, A. C.; Turco, S. J.; Tschudi, C.; Ullu, E.; Beverley, S. M., Retention and Loss of RNA Interference Pathways in Trypanosomatid Protozoans. *PLOS Pathogens* **2010,** *6* (10), e1001161.

117. Cohen-Freue, G.; Holzer, T. R.; Forney, J. D.; McMaster, W. R., Global gene expression in Leishmania. *International Journal for Parasitology* **2007,** *37* (10), 1077-1086.

118. Eng, J. K.; Jahan, T. A.; Hoopmann, M. R., Comet: An open-source MS/MS sequence database search tool. *Proteomics* **2013,** *13* (1), 22-24.

119. Dorfer, V.; Pichler, P.; Stranzl, T.; Stadlmann, J.; Taus, T.; Winkler, S.; Mechtler, K., MS Amanda, a Universal Identification Algorithm Optimized for High Accuracy Tandem Mass Spectra. *Journal of Proteome Research* **2014,** *13* (8), 3679-3684.

120. Tabb, D. L.; Fernando, C. G.; Chambers, M. C., MyriMatch: highly accurate tandem mass spectral peptide identification by multivariate hypergeometric analysis. *J Proteome Res* **2007,** *6* (2), 654-61.

121. Kim, S.; Pevzner, P. A., MS-GF+ makes progress towards a universal database search tool for proteomics. *Nat Commun* **2014,** *5* (5277).

122. Craig, R.; Beavis, R. C., TANDEM: matching proteins with tandem mass spectra. *Bioinformatics* **2004,** *20* (9), 1466-7.

123. Vaudel, M.; Barsnes, H.; Berven, F. S.; Sickmann, A.; Martens, L., SearchGUI: An open-source graphical user interface for simultaneous OMSSA and X!Tandem searches. *Proteomics* **2011,** *11* (5), 996-999.

124. Perkins, D. N.; Pappin, D. J. C.; Creasy, D. M.; Cottrell, J. S., Probability-based protein identification by searching sequence databases using mass spectrometry data. *Electrophoresis* **1999,** *20* (18), 3551-3567.

125. Brosch, M.; Yu, L.; Hubbard, T.; Choudhary, J., Accurate and Sensitive Peptide Identification with Mascot Percolator. *Journal of Proteome Research* **2009,** *8* (6), 3176-3181.

126. Uniprot, UniProt: the universal protein knowledgebase. *Nucleic Acids Res.* **2017,** *45* (D1), D158-D169.

127. Vaudel, M.; Venne, A. S.; Berven, F. S.; Zahedi, R. P.; Martens, L.; Barsnes, H., Shedding light on black boxes in protein identification. *Proteomics* **2014,** *14* (9), 1001-5.

128. Craig, R.; Cortens, J. P.; Beavis, R. C., Open source system for analyzing, validating, and storing protein identification data. *J Proteome Res* **2004,** *3* (6), 1234-42.

129. Ncbi Resource Coordinators, Database Resources of the National Center for Biotechnology Information. *Nucleic Acids Res.* **2017,** *45* (Database issue), D12-D17.

130. Berman, H. M.; Westbrook, J.; Feng, Z.; Gilliland, G.; Bhat, T. N.; Weissig, H.; Shindyalov, I. N.; Bourne, P. E., The Protein Data Bank. *Nucleic Acids Res.* **2000,** *28* (1), 235-242.

131. O'Leary, N. A.; Wright, M. W.; Brister, J. R.; Ciufo, S.; Haddad, D.; McVeigh, R.; Rajput, B.; Robbertse, B.; Smith-White, B.; Ako-Adjei, D.; Astashyn, A.; Badretdin, A.; Bao, Y.; Blinkova, O.; Brover, V.; Chetvernin, V.; Choi, J.; Cox, E.; Ermolaeva, O.; Farrell, C. M.; Goldfarb, T.; Gupta, T.; Haft, D.; Hatcher, E.; Hlavina, W.; Joardar, V. S.; Kodali, V. K.; Li, W.; Maglott, D.; Masterson, P.; McGarvey, K. M.; Murphy, M. R.; O'Neill, K.; Pujar, S.; Rangwala, S. H.; Rausch, D.; Riddick, L. D.; Schoch, C.; Shkeda, A.; Storz, S. S.; Sun, H.; Thibaud-Nissen, F.; Tolstoy, I.; Tully, R. E.; Vatsan, A. R.; Wallin, C.; Webb, D.; Wu, W.; Landrum, M. J.; Kimchi, A.; Tatusova, T.; DiCuccio, M.; Kitts, P.; Murphy, T. D.; Pruitt, K. D., Reference sequence (RefSeq) database at NCBI: current status, taxonomic expansion, and functional annotation. *Nucleic Acids Res* **2016,** *44* (D1), 8.

132. Kodama, Y.; Mashima, J.; Kosuge, T.; Kaminuma, E.; Ogasawara, O.; Okubo, K.; Nakamura, Y.; Takagi, T., DNA Data Bank of Japan: 30th anniversary. *Nucleic Acids Res* **2018,** *46* (D1), D30-D35.

133. Kanz, C.; Aldebert, P.; Althorpe, N.; Baker, W.; Baldwin, A.; Bates, K.; Browne, P.; van den Broek, A.; Castro, M.; Cochrane, G.; Duggan, K.; Eberhardt, R.; Faruque, N.; Gamble, J.; Diez, F. G.; Harte, N.; Kulikova, T.; Lin, Q.; Lombard, V.; Lopez, R.; Mancuso, R.; McHale, M.; Nardone, F.; Silventoinen, V.; Sobhany, S.; Stoehr, P.; Tuli, M. A.; Tzouvara, K.; Vaughan, R.; Wu, D.; Zhu, W.; Apweiler, R., The EMBL Nucleotide Sequence Database. *Nucleic Acids Res.* **2005,** *33* (Database Issue), D29-D33.

134. Benson, D. A.; Karsch-Mizrachi, I.; Lipman, D. J.; Ostell, J.; Wheeler, D. L., GenBank. *Nucleic Acids Res.* **2005,** *33* (Database Issue), D34-D38.

135. Wu, C. H.; Yeh, L.-S. L.; Huang, H.; Arminski, L.; Castro-Alvear, J.; Chen, Y.; Hu, Z.; Kourtesis, P.; Ledley, R. S.; Suzek, B. E.; Vinayaka, C. R.; Zhang, J.; Barker, W. C., The Protein Information Resource. *Nucleic Acids Res.* **2003,** *31* (1), 345-347.

136. Vaudel, M.; Burkhart, J. M.; Zahedi, R. P.; Oveland, E.; Berven, F. S.; Sickmann, A.; Martens, L.; Barsnes, H., PeptideShaker enables reanalysis of MS-derived proteomics data sets. *Nature biotechnology* **2015,** *33*, 22.

137. Calculate and draw custom Venn diagrams. http://bioinformatics.psb.ugent.be/webtools/Venn/.

138. Szklarczyk, D.; Franceschini, A.; Wyder, S.; Forslund, K.; Heller, D.; Huerta-Cepas, J.; Simonovic, M.; Roth, A.; Santos, A.; Tsafou, K. P.; Kuhn, M.; Bork, P.; Jensen, L. J.; von Mering, C., STRING v10: protein-protein interaction networks, integrated over the tree of life. *Nucleic Acids Res* **2015,** *43* (Database issue), 28.

139. Biswas, A.; Bhattacharya, A.; Das, P. K., Role of cAMP Signaling in the Survival and Infectivity of the Protozoan Parasite, Leishmania donovani. *Mol Biol Int* **2011,** *782971* (10), 5.

140. Biswas, A.; Bhattacharya, A.; Vij, A.; Das, P. K., Role of leishmanial acidocalcisomal pyrophosphatase in the cAMP homeostasis in phagolysosome conditions required for intra-macrophage survival. *The International Journal of Biochemistry & Cell Biology* **2017,** *86*, 1-13.

141. Mwenechanya, R.; Kovářová, J.; Dickens, N. J.; Mudaliar, M.; Herzyk, P.; Vincent, I. M.; Weidt, S. K.; Burgess, K. E.; Burchmore, R. J. S.; Pountain, A. W.; Smith, T. K.; Creek, D. J.; Kim, D.-H.; Lepesheva, G. I.; Barrett, M. P., Sterol 14α-demethylase mutation leads to amphotericin B resistance in Leishmania mexicana. *PLOS Neglected Tropical Diseases* **2017,** *11* (6), e0005649.

142. Saini, S.; Bharati, K.; Shaha, C.; Mukhopadhyay, C. K., Zinc depletion promotes apoptosis-like death in drug-sensitive and antimony-resistance Leishmania donovani. *Scientific Reports* **2017,** *7* (1), 10488.

143. Bodzon-Kulakowska, A.; Suder, P., Imaging mass spectrometry: Instrumentation, applications, and combination with other visualization techniques. *Mass Spectrometry Reviews* **2016,** *35* (1), 147-169.

144. Kaya, I.; Michno, W.; Brinet, D.; Iacone, Y.; Zanni, G.; Blennow, K.; Zetterberg, H.; Hanrieder, J., Histology-Compatible MALDI Mass Spectrometry Based Imaging of Neuronal Lipids for Subsequent Immunofluorescent Staining. *Analytical Chemistry* **2017,** *89* (8), 4685-4694.

145. Seeley, E. H.; Caprioli, R. M., MALDI imaging mass spectrometry of human tissue: method challenges and clinical perspectives. *Trends in Biotechnology* **2011,** *29* (3), 136-143.

146. Crecelius, A. C.; Cornett, D. S.; Caprioli, R. M.; Williams, B.; Dawant, B. M.; Bodenheimer, B., Three-Dimensional Visualization of Protein Expression in Mouse Brain Structures Using Imaging Mass Spectrometry. *Journal of the American Society for Mass Spectrometry* **2005,** *16* (7), 1093-1099.

147. Nimesh, S.; Mohottalage, S.; Vincent, R.; Kumarathasan, P., Current Status and Future Perspectives of Mass Spectrometry Imaging. *International Journal of Molecular Sciences* **2013,** *14* (6), 11277-11301.

148. Stauber, J.; MacAleese, L.; Franck, J.; Claude, E.; Snel, M.; Kaletas, B. K.; Wiel, I. M. V. D.; Wisztorski, M.; Fournier, I.; Heeren, R. M. A., On-Tissue Protein Identification and Imaging by MALDI-Ion Mobility Mass Spectrometry. *Journal of the American Society for Mass Spectrometry* **2010,** *21* (3), 338-347.

149. Yang, J.; Caprioli, R. M., Matrix Sublimation/Recrystallization for Imaging Proteins by Mass Spectrometry at High Spatial Resolution. *Analytical Chemistry* **2011,** *83* (14), 5728-5734.

APPENDIX A

SOLUTION COMPONENTS FOR 2D-SDS PAGE

- *Acrylamide/bis-acrylamide mixture (30%):* 29,2 gram acrylamide and 0,8 gram N,N'-methylenebisacylamide are dissolved in a total volume of 100 ml with dH_2O.
- *1,5 M Tris/HCl; pH 8.8:* 18,15 gram Tris base is dissolved in ~80 ml dH_2O and pH is adjusted to 8.8 with 6N HCl. Volume is made up to 100 ml with dH_2O.
- *10% SDS:* 10 gram SDS is dissolved in a total volume of 100 ml dH_2O.
- *10% APS*: 100 mg APS is dissolved in 1 ml dH_2O (should be fresh).
- *Overlay agarose*: 0,5% low melting point agarose including 0,003% bromophenol blue dye is prepared in 1X TGS buffer.
- *Tris/Glycine/SDS (TGS) Running Buffer (5X)*: 15 gram tris base, 72 gram glycine and 5 gram SDS are dissolved in total volume of 1 litre dH_2O. It should be diluted to 1X for electrophoretic run.

Isoelectric Focusing Steps		Focus temp: 20 °C
Active rehydration: fixed 50 V for 16 hours		
Step#1	Voltage slope: Linear-200 V	300 Vhours
Step#2	Voltage slope: Linear-500 V	500 Vhours
Step#3	Voltage slope: Linear-1000 V	1000 Vhours
Step#4	Voltage slope: Linear-4000 V	4000 Vhours
Step#5	Voltage slope: Rapid-8000 V	24000 Vhours
Step#6	Voltage slope: Rapid 8000 V	30000 Vhours

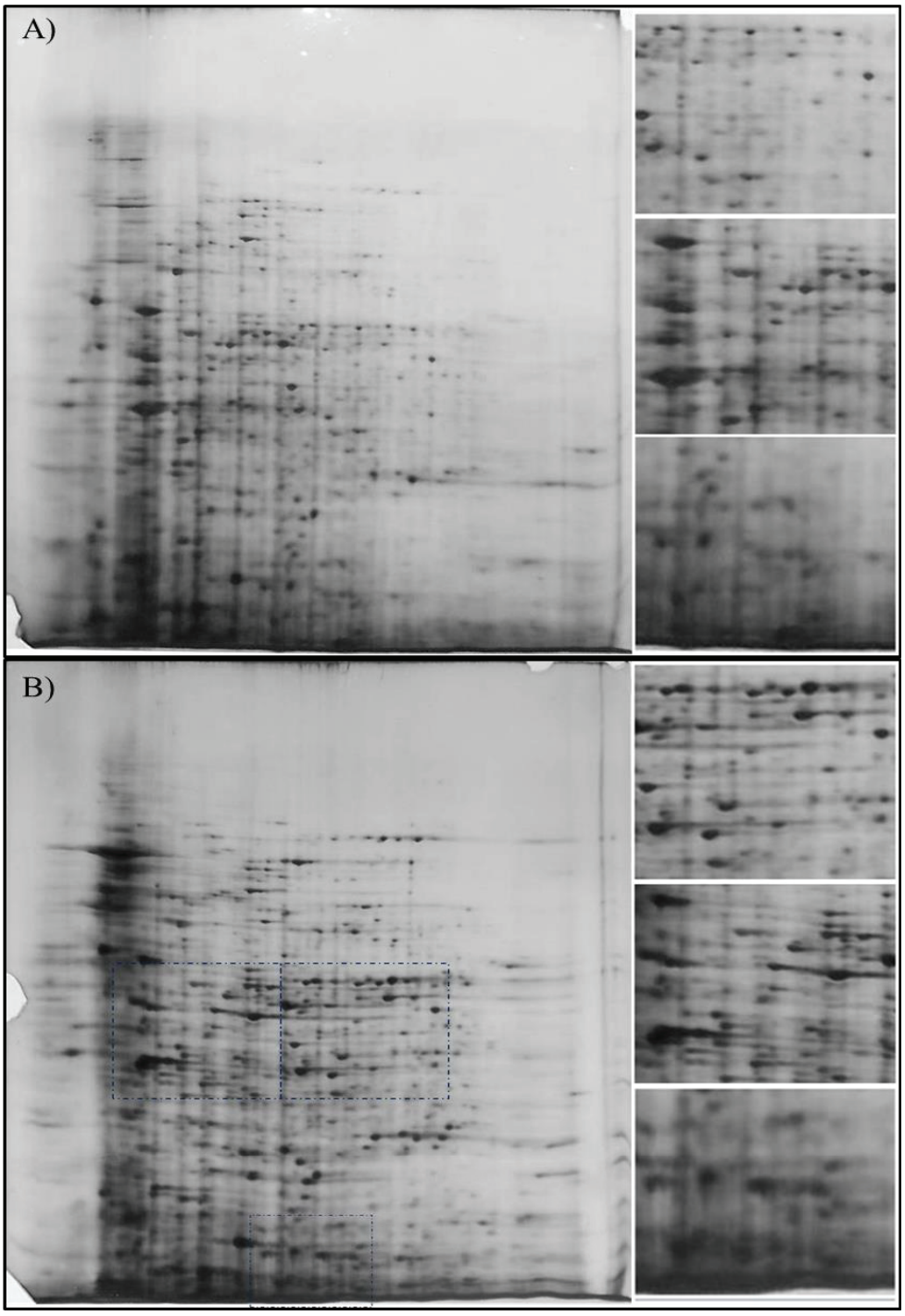
Equilibration base buffer	**Reagents**	**Amount (Final Concentration)**
	Urea (6M)	1.81 gram
	SDS (2%)	0.1 gram
	0.375 M Tris/HCl, pH 8.8	1.25 ml (from 1.5 M)
	Glycerol	1 ml
	Water	Adjust to final volume of 5 ml
to equilibration buffer-1	2% DTT	0.1 gram
to equilibration buffer-2	2.5 % Iodoacetamide	0.125 gram

Solution components for 12%	**30 ml**	**40 ml**	**Solution components for 15%**	**30 ml**	**40 ml**
H2O	9,9	13,2	H2O	6,9	9,2
30% acrylamide mix	12,0	16,0	30% acrylamide mix	15,0	20,0
1.5 M Tris/HCl (pH 8.8)	7,5	10,0	1.5 M Tris/HCl (pH 8.8)	7,5	10,0
10% SDS	0,3	0,4	10% SDS	0,3	0,4
10%Ammonium persulfate	0,3	0,4	10%Ammonium persulfate	0,3	0,4
TEMED	12	16	TEMED	12	16

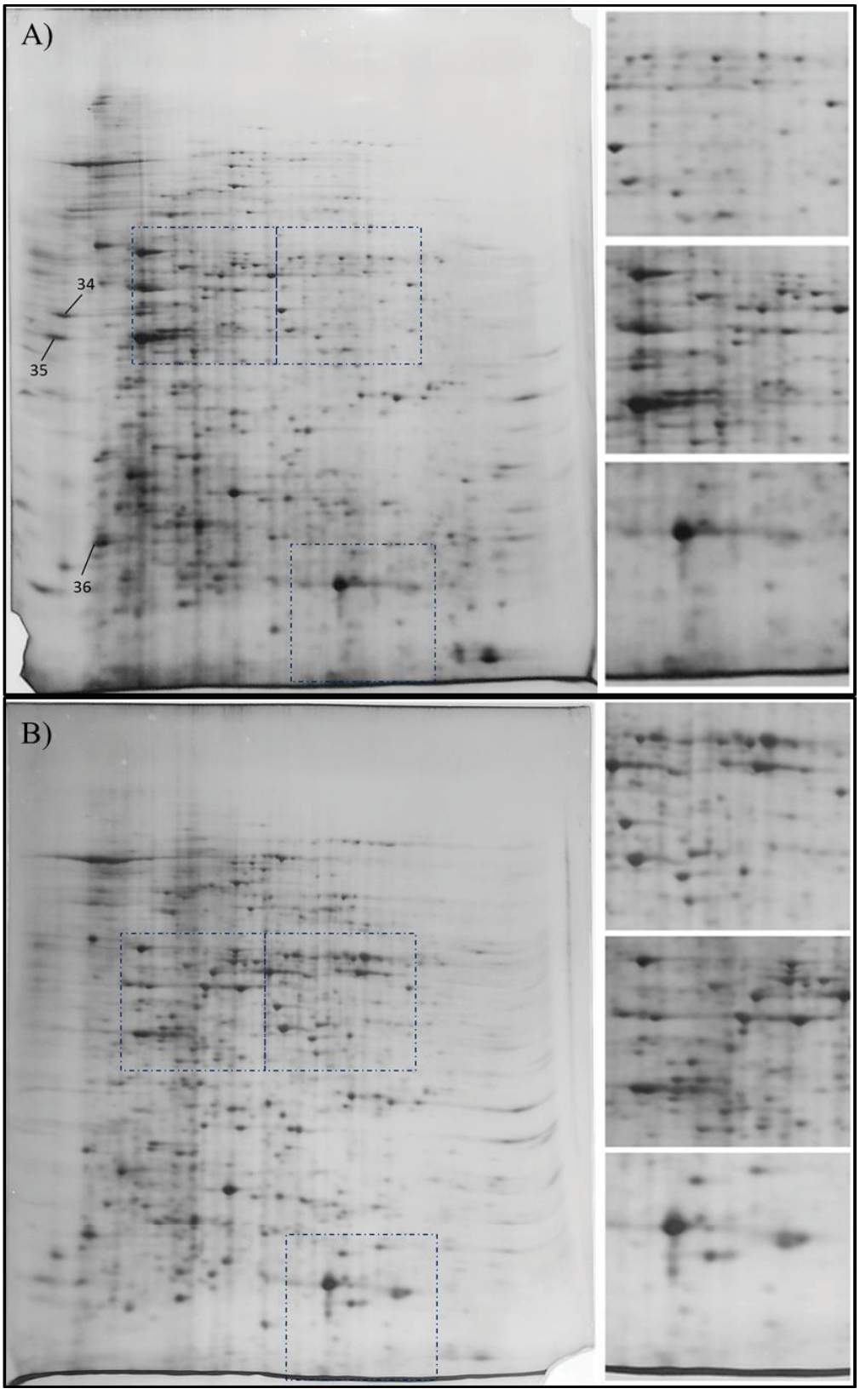
APPENDIX B

2D-SDS-PAGE IMAGES AND IDENTIFIED SPOTS

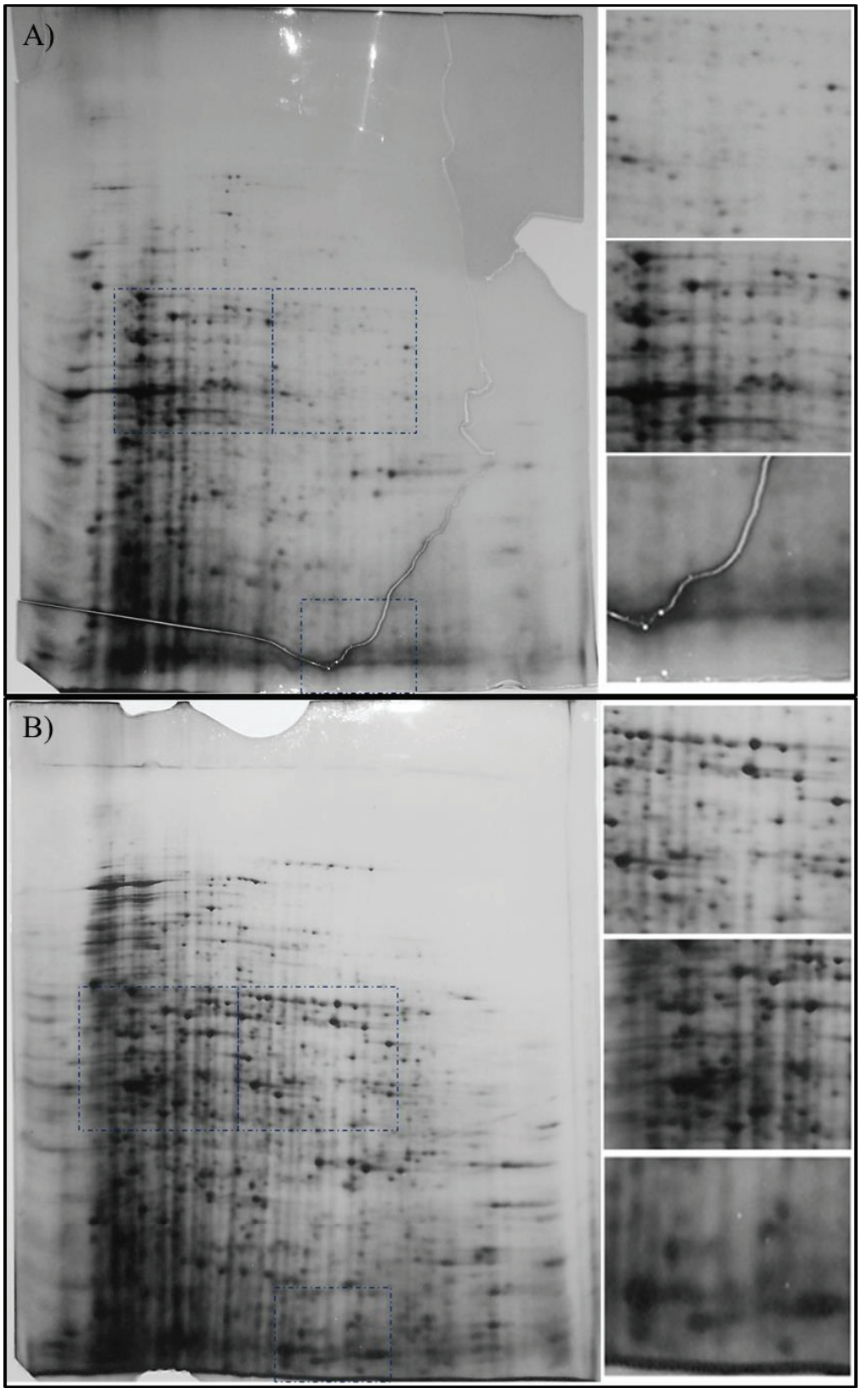
A) *Leishmania infantum*-Visceral sample-239 (pH 3-10)
B) *Leishmania infantum*-Cutaneous sample-121 (pH 3-10)



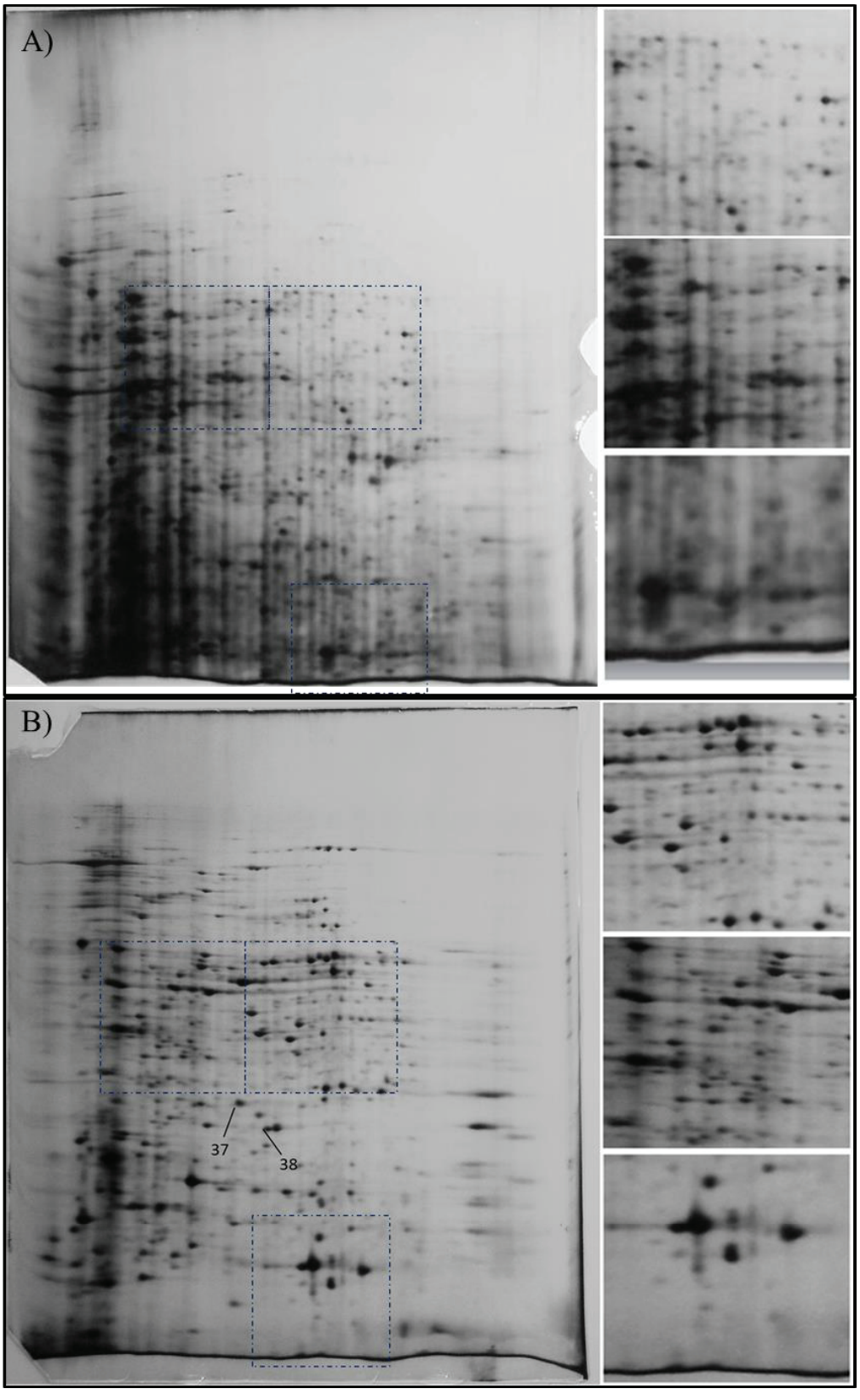
A) *Leishmania infantum*-Visceral sample-240 (pH 3-10)
B) *Leishmania infantum*-Cutaneous sample-250 (pH 3-10)



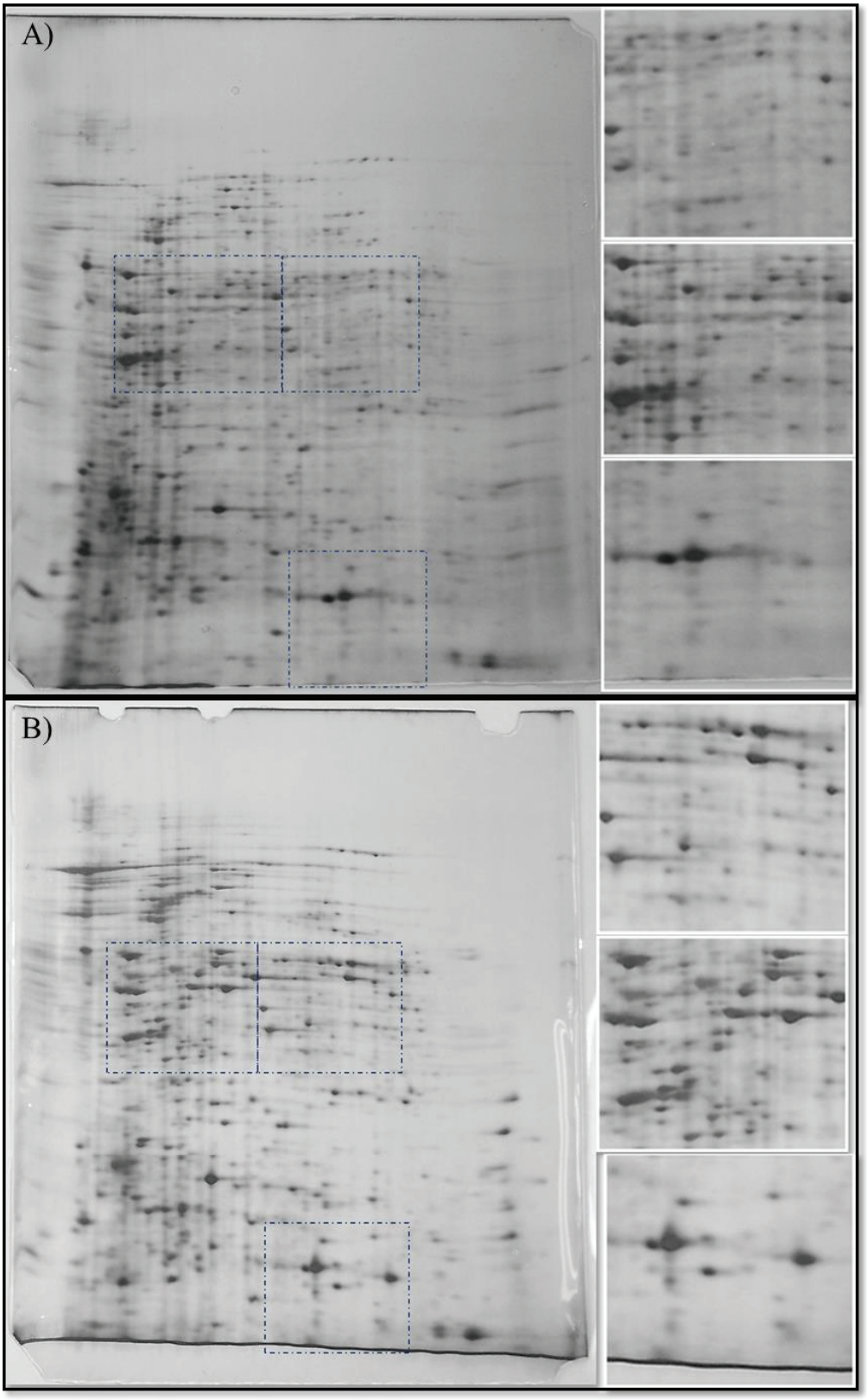
A) *Leishmania infantum*-Visceral sample-495 (pH 3-10)
B) *Leishmania infantum*-Cutaneous sample-285 (pH 3-10)



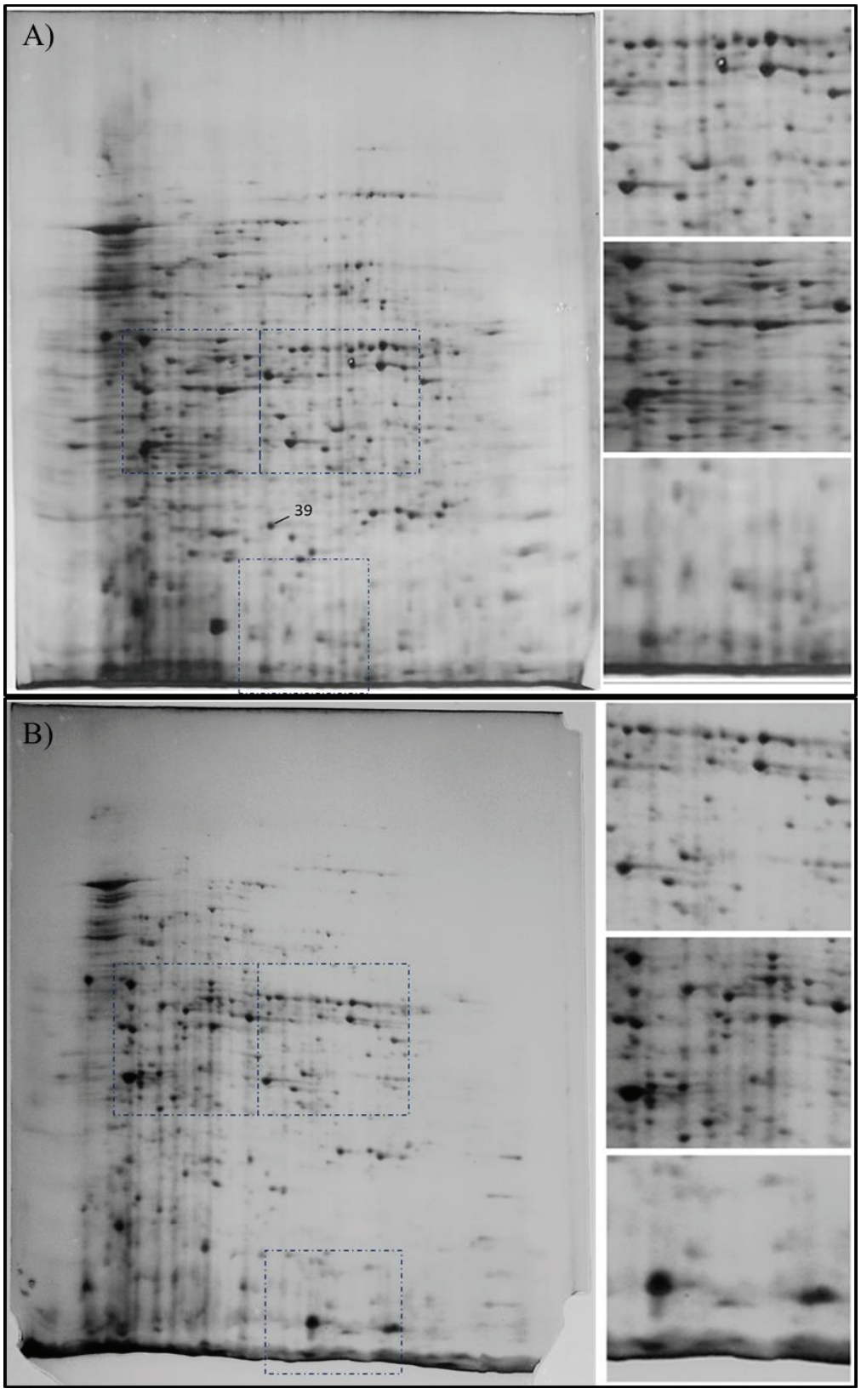
A) *Leishmania infantum*-Visceral sample-628 (pH 3-10)
B) *Leishmania infantum*-Cutaneous sample-259 (pH 3-10)



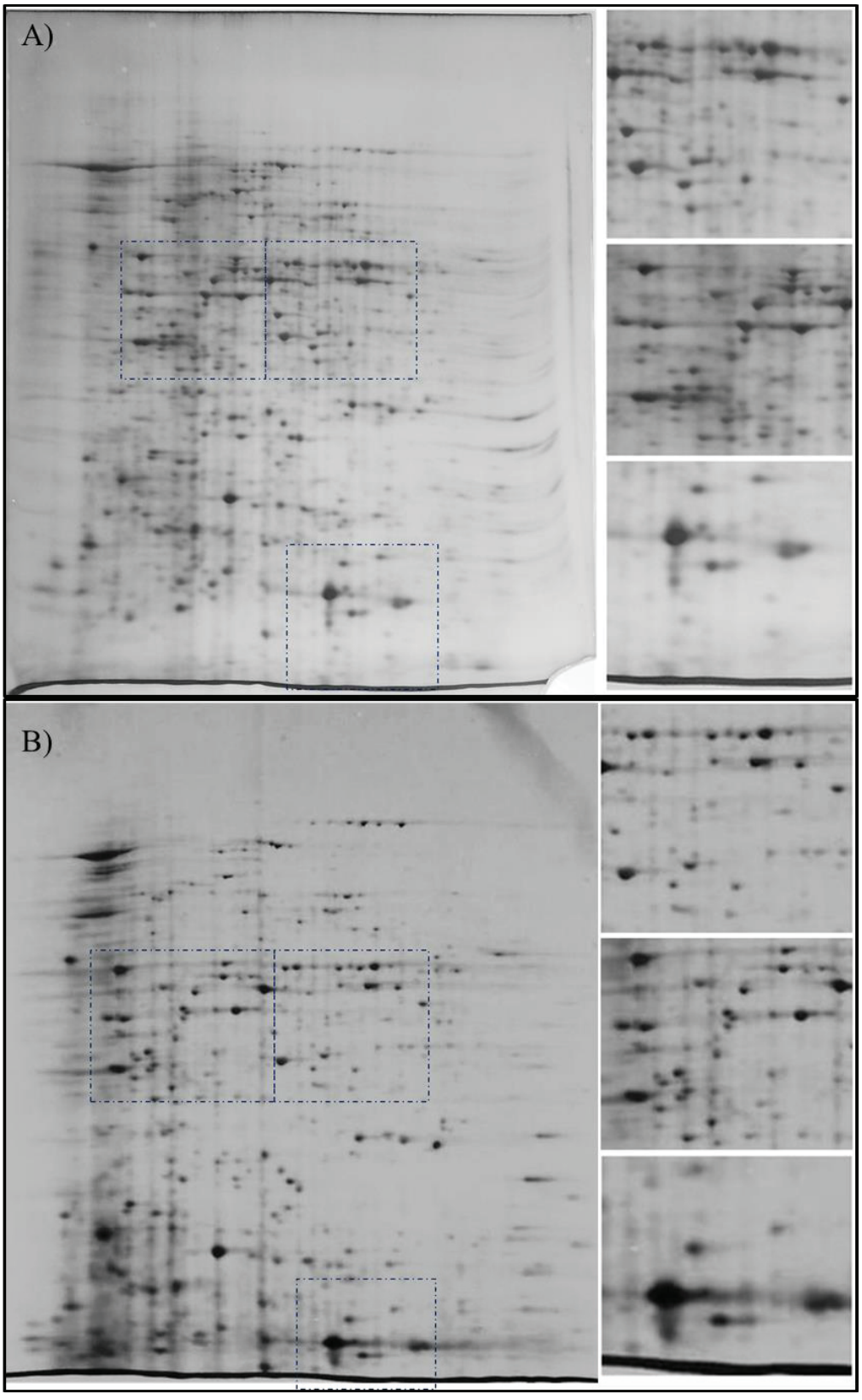
A) *Leishmania infantum* Visceral sample-Crt07 (pH 3-10)
B) *Leishmania infantum*-Cutaneous sample-Cri08 (pH 3-10)



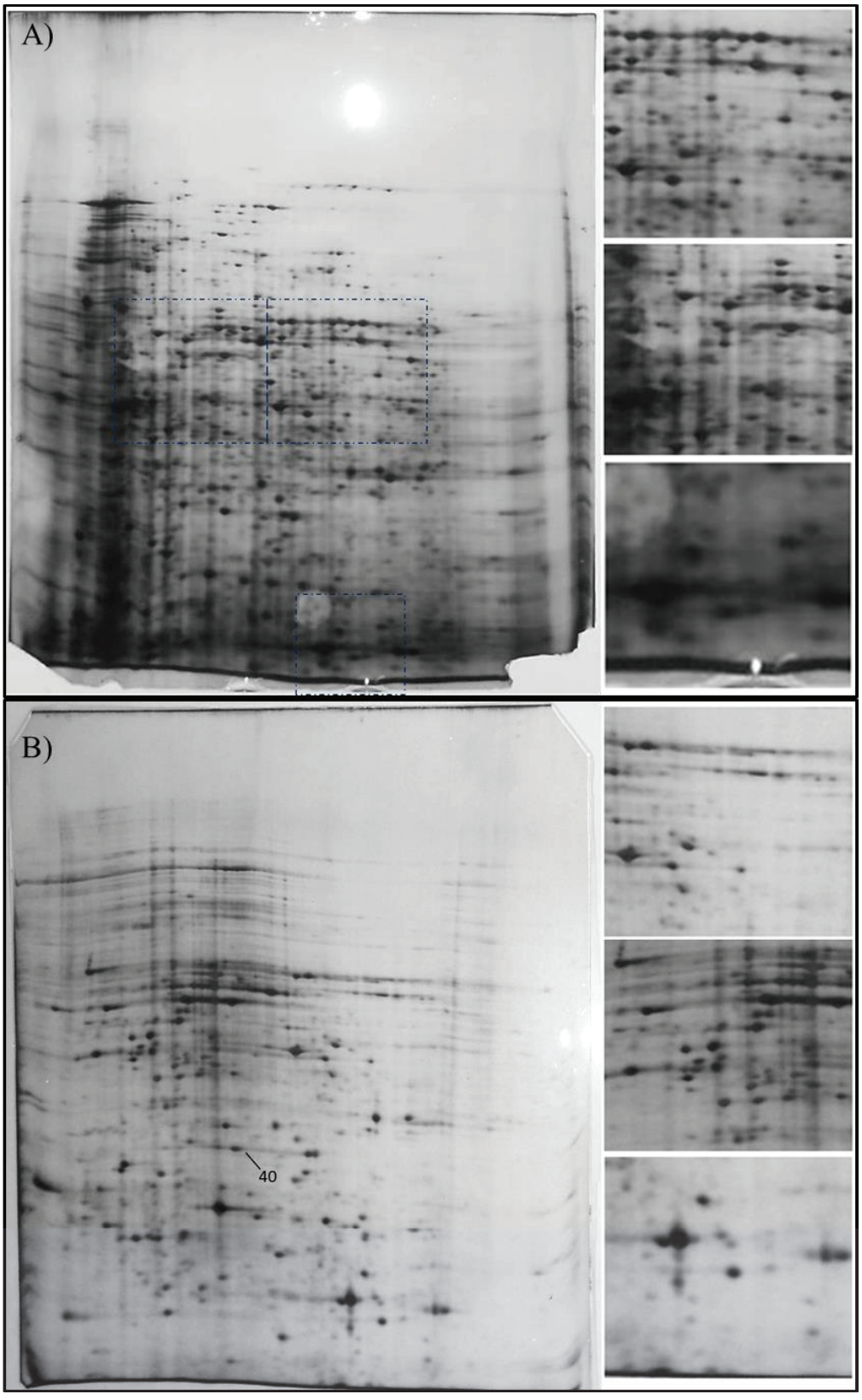
A) *Leishmania tropica*-Cutaneous sample-Crt02 (pH 3-10)
B) *Leishmania tropica*-Cutaneous sample-615 (pH 310)



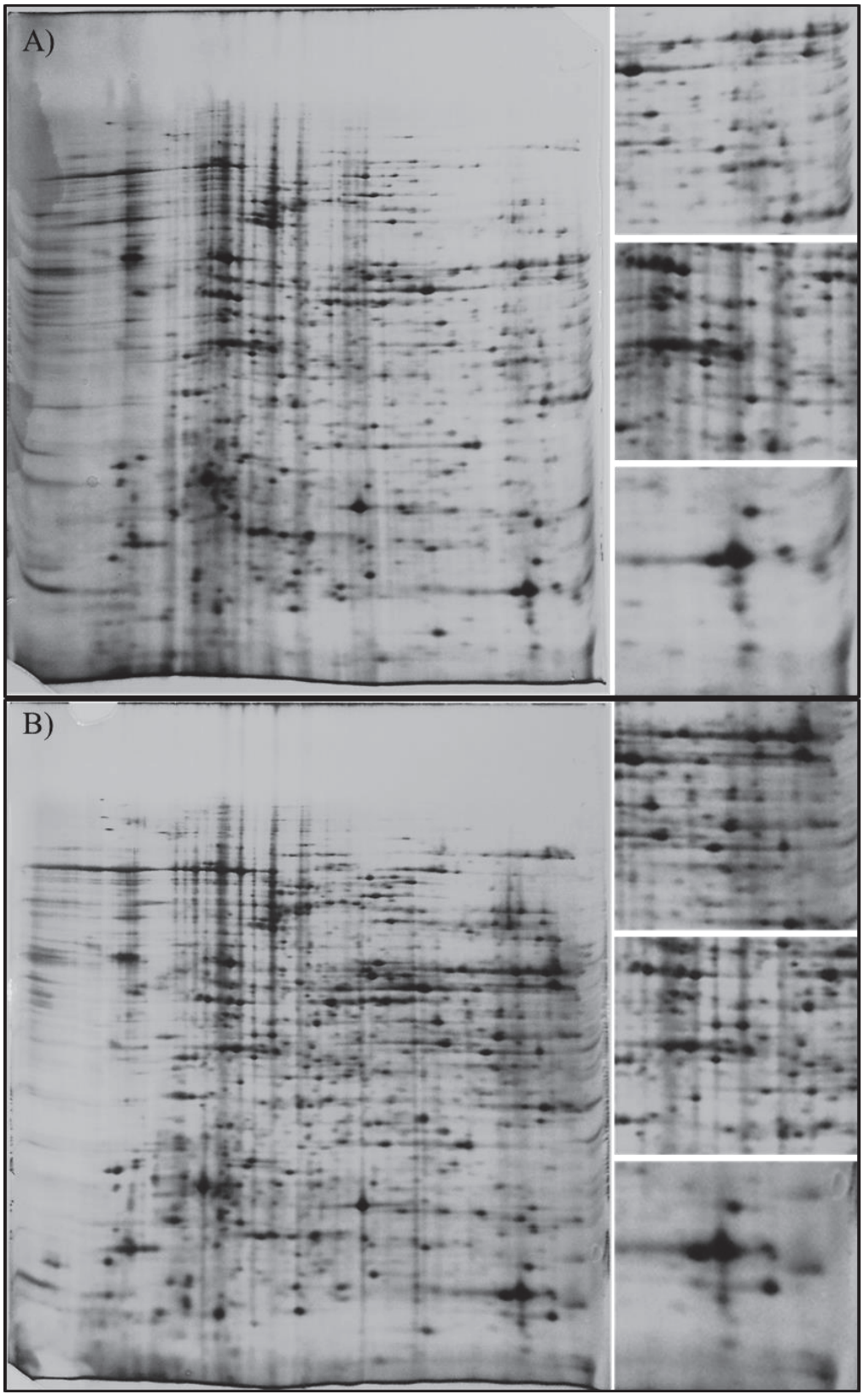
A) *Leishmania tropica*-Cutaneous sample-Crt03 (pH3-10)
B) *Leishmania tropica*-Cutaneous sample-2FA (pH 3-10)



A) *Leishmania tropica*-Cutaneous sample-Crt01 (pH 3-10)
B) *Leishmania infantum*-Cutaneous sample-285 (pH 5-8)



A) *Leishmania infantum*-Visceral sample-628 (pH 4-7)
B) *Leishmania infantum*-Cutaneous sample-259 (pH 4-7)



LIST OF IDENTIFIED SPOTS

Spot #	Protein Name	MW	pI	PMF	MS/MS
1	Putative iron superoxide dismutase	21	6,4	-	2
2	Peroxidoxin	21	6,3	11	-
3	Peroxidoxin	21	6,3	8	2
4	Superoxide dismutase	21	6,4	-	2
5	Peroxidoxin	21	6,3	8	2
6	Pyruvate kinase	56	6,4	14	3
7	Putative succinyl-coA:3-ketoacid-coenzyme A transferase, mitochondrial precursor	53	7,9	6	4
8	Putative dihydrolipoamide dehydrogenase	51	6,4	-	2
9	Trypanothione reductase	53	5,8	-	1
10	Aspartate aminotransferase	46	6,5	8	2
11	Putative dihydrolipoamide dehydrogenase	51	6,4	9	-
12	Putative dihydrolipoamide dehydrogenase	51	6,4	9	-
13	Putative aldose 1-epimerase	42	5,9	-	2
14	Putative aldose 1-epimerase	42	5,9	6	-
15	Putative eukaryotic initiation factor 4a	45	5,8	9	2
16	Putative aldose 1-epimerase	42	5,8	9	1
17	Arginase	36	6,3	-	1
18	Putative aldose 1-epimerase	42	5,9	11	2
19	Putative dihydrolipoamide dehydrogenase	51	6,4	-	1
20	Pyruvate kinase	56	6,4	8	2
21	Metallo-peptidase, Clan ME, Family M16	55	6,6	10	-
22	Metallo-peptidase, Clan MA(E), Family M32	57	5,4	12	2
23	Putative eukaryotic initiation factor 4a	45	5,8	9	2
24	Heat shock protein 83-1	81	5,0	-	3
25	Putative N-acyl-L-amino acid amidohydrolase	44	5,1	5	2
26	Putative N-acyl-L-amino acid amidohydrolase	44	5,1	4	-
27	Enolase	46	5,3	13	3
28	Metallo-peptidase, Clan MA(E), family 32	57	5,4	12	1
29	Cytochrome c oxidase subunit IV	39	5,5	6	2
30	Metallo-peptidase, Clan MA(E), Family M32	57	5,4	11	2
31	Heat shock protein 83-1	81	5,0	-	2
32	Metallo-peptidase, Clan MA(E), family 32	57	5,4	16	2
33	Heat shock protein 83-1	81	5,0	-	3
34	Heat shock protein 83-1	81	5,0	-	2
35	Heat shock protein 83-2	80	5,0	10	2
36	Heat shock protein 70	50	5,2	4	2
37	Inosine-uridine preferring nucleoside hydrolase	34	5,8	4	-
38	Prostaglandin f2-alpha synthase, partial	30	6,7	5	1
39	Inosine-uridine preferring nucleoside hydrolase	34	5,8	5	-
40	Coproporphyrinogen III oxidase	35	5,3	6	2

APPENDIX C

DIFFERENTIAL PROTEINS LIST OF THE SEARCH AGAINST NCBI_L.GENUS DATABASE

Differential Proteins List of Mascot Search against NCBI Database

Groups	Accession code	Protein Identity
All cutaneous 121 259 285 615 Crt01 Crt02	gi|1201119579	iron superoxide dismutase
	gi|259437281	unnamed protein product
	gi|68124282	flavoprotein subunit-like protein
	gi|134073997	putative eukaryotic translation initiation factor 3 subunit
	gi|322488686	conserved hypothetical protein (90% →A4I4F2_LEIIN)
	gi|321438672	conserved hypothetical protein (90% →A4I4F2_LEIIN)
	gi|68126423	putative (H+)-ATPase G subunit
	gi|134073899	conserved hypothetical protein (100% →A4ICR0_LEIIN)
	gi|157872750	glyceraldehyde 3-phosphate dehydrogenase,glycosomal
	gi|9858999	putative protein kinase A catalytic subunit
	gi|134072679	metallo-peptidase, Clan MA(E), Family M32
	gi|134073430	galactokinase-like protein
	gi|68223979	putative 2-oxoisovalerate dehydrogenase alpha subunit
	gi|68124875	putative lanosterol 14-alpha-demethylase
	gi|321438278	conserved hypothetical protein (90% →A4HRV7_LEIIN)
	gi|134061940	putative proteasome alpha 2 subunit
121 259 285 615 crt01	gi|68124755	conserved hypothetical protein
	gi|68128322	2-oxoglutarate dehydrogenase, e3 component,lipoamidedehydrogenase-like protein
	gi|68126666	triosephosphate isomerase
	gi|68124837	anion-transporting ATPase-like protein
	gi|68128018	putative CAS/CSE/importin domain protein
	gi|68126740	conserved hypothetical protein
121-259 285-615 crt02	gi|94469839	glucose-6-phosphate dehydrogenase
	gi|321438704	putative serine peptidase
	gi|322490788	chaperonin HSP60, mitochondrial precursor
121 259 285 crt01 crt02	gi|321438725	conserved hypothetical protein
	gi|68127546	putative X-pro, dipeptidyl-peptidase,serine peptidase, Clan SC, family S15
	gi|321398956	metallo-peptidase, Clan ME, Family M16C
	gi|134068630	putative fucose kinase
	gi|134072673	conserved hypothetical protein
	gi|134070867	putative proteasome regulatory non-ATP-ase subunit 3
121-259 615-crt01 crt02	gi|68126596	beta transducin-like protein
	gi|134071050	metallo-peptidase, Clan MH, Family M18
	gi|68130009	conserved hypothetical protein
121-285	gi|134067273	putative deoxyribose-phosphate aldolase

615-crt01 crt02	gi|68223986	conserved hypothetical protein
	gi|11022577	hypothetical protein L7845.03
	gi|32329336	acidocalcisomal pyrophosphatase
259 285 615 crt01 crt02	gi|321438357	putative calpain-like cysteine peptidase
	gi|68126620	putative aspartate aminotransferase
	gi|68125093	putative RNA helicase
	gi|321438474	conserved hypothetical protein
	gi|68223941	putative 40S ribosomal protein S23
	gi|134073884	putative eukaryotic translation initiation factor 6 (eIF-6)
Cutaneous *L.infantum* 121 259 285	gi|134074178	COP-coated vesicle membrane protein erv25 precursor
	gi|134068133	putative protein kinase A regulatory subunit
	gi|68124450	beta tubulin
	gi|68128036	putative mago nashi-like protein
	gi|134070835	putative TPR-repeat protein
	gi|134071352	putative N-acyl-L-amino acid amidohydrolase
Cutaneous *L.tropica.* 615 Crt01 Crt02	gi|68126335	conserved hypothetical protein (90% →A4HYX0_LEIIN)
	gi|68224133	conserved hypothetical protein (90% → A4HZX1_LEIIN)
	gi|159393	pyruvate kinase, partial
	gi|134071629	putative 5-methyltetrahydropteroyltriglutamate-homocystein e S-methyltransferase
	gi|68127346	putative glutamate 5-kinase
	gi|68128232	nucleosome assembly protein-like protein
Visceral *L.infantum* 240 495 628	gi|68129116	conserved hypothetical protein (50% →A4I9C1_LEIIN)
	gi|459866	ribosomal protein P0
	gi|68124075	putative 20S proteasome beta 6 subunit
	gi|134068758	elongation factor 1-alpha
	gi|134067952	metallo-peptidase, Clan MF, Family M17
	gi|27434608	glucose-6-phosphate dehydrogenase
	gi|134067632	conserved hypothetical protein (100% → A4HU05_LEIIN)
	gi|68130046	putative eukaryotic translation initiation factor 3 subunit
	gi|117617445	superoxide dismutase B1
	gi|68126142	putative phenylalanyl-tRNA synthetase

Differential Proteins List of Comet Search against NCBI Database

Groups	**Accession code**	**Protein Identity**
All cutaneous 121 259 285 615 Crt01 Crt02	gi|157873843	putative dynein
	gi|1099051280	proteasome alpha 2 subunit, putative
	gi|157865503	conserved hypothetical protein
	gi|134070910	putative thymine-7-hydroxylase
	gi|321438278	conserved hypothetical protein
	gi|134066495	putative 40S ribosomal protein SA
	gi|1099045425	protein transport protein Sec13, putative
	gi|134073899	conserved hypothetical protein
	gi|134070533	trifunctional enzyme alpha subunit, mitochondrial precursor-like protein
	gi|134062977	proline oxidase, mitochondrial precursor-like protein
	gi|1099058444	major vault protein, putative
	gi|322488686	conserved hypothetical protein
	gi|259437281	unnamed protein product
	gi|321438725	conserved hypothetical protein
	gi|134067381	flavoprotein subunit-like protein
	gi|213508900	unnamed protein product
	gi|134066741	conserved hypothetical protein
121	gi|134071248	putative phosphatase 2C

259 285 615 crt01	gi|134071050	metallo-peptidase, Clan MH, Family M18
	gi|134069085	putative C-terminal motor kinesin
	gi|1099044115	heat shock protein, putative
	gi|134071823	diphthine synthase-like protein
	gi|134072950	putative asparaginyl-tRNA synthetase
	gi|1099057709	hypothetical protein, conserved
121 259 285 615 crt02	gi|134069751	putative 40S ribosomal protein L14
	gi|134060875	inosine-5'-monophosphate dehydrogenase
	gi|134069281	conserved hypothetical protein
	gi|157865564	acidocalcisomal pyrophosphatase
	gi|157873476	putative calreticulin
	gi|322495258	putative pyruvate kinase
	gi|134073652	cyclophilin 40
	gi|1099052766	asparaginyl-tRNA synthetase, putative
	gi|134072680	putative isocitrate dehydrogenase
121 259 285 crt01 crt02	gi|134069536	conserved hypothetical protein
	gi|259655018	unnamed protein product
	gi|134064331	putative 5-methyltetrahydropteroyltriglutamate-homocystein e S-methyltransferase
	gi|134065642	60S ribosomal protein L32
121 259 615 crt01 crt02	gi|321438728	conserved hypothetical protein
	gi|321438728	conserved hypothetical protein
	gi|1099052581	hypothetical protein, conserved
	gi|145411496	mitochondrial tryparedoxin peroxidase
	gi|134068978	conserved hypothetical protein
	gi|134072596	putative map kinase
	gi|134084922	putative ribonucleoside-diphosphate reductase large chain
121-285-615 crt01-crt02	gi|157869818	putative coronin
	gi|157877272	putative glycine synthase
259 285 615 crt01 crt02	gi|134068886	conserved hypothetical protein
	gi|134067200	NAD(p)-dependent steroid dehydrogenase-like protein
	gi|1099049837	elongation factor, putative
	gi|134071126	putative cysteine peptidase, Clan CA, family C19
	gi|134065708	galactokinase-like protein
	gi|321438474	conserved hypothetical protein
	gi|1099058414	nucleolar RNA helicase II, putative
259 259 285 615 crt01 crt02	gi|134070935	ADF/Cofilin
	gi|134068777	cystathionine beta-synthase
	gi|157867709	putative mevalonate-diphosphate decarboxylase
	gi|321398624	putative succinyl-coA:3-ketoacid-coenzyme A transferase,mitochondrial precursor
	gi|322490614	conserved hypothetical protein
	gi|157869012	putative 40S ribosomal protein S23
Cutaneous *L.infantum* 121 259 285	gi|1099049424	hypothetical protein, conserved
	gi|134060760	putative cytochrome c
	gi|134067981	conserved hypothetical protein
	gi|134067773	folate/biopterin transporter, putative
Cutaneous *L.tropica.* 615 Crt01 Crt02	gi|134069358	conserved hypothetical protein
Visceral *L.infantum* 240 495 628	gi|134072699	conserved hypothetical protein
	gi|322491767	conserved hypothetical protein
	gi|321399401	histone H2A
	gi|134074163	chaperonin HSP60, mitochondrial precursor
	gi|134071929	zinc transporter-like protein
	gi|134068758	elongation factor 1-alpha

APPENDIX D

PEPTIDESHAKER VALIDATION REPORT DATA

	Comet			X! Tandem		
	121	**259**	**285**	**121**	**259**	**285**
Proteins_#Validated	840	1179	1082	787	1044	1121
Proteins_Total Possible TP	996,37	1347,89	1355,98	1087,04	1353,73	1354,39
Proteins_FDR Limit [%]	0,83	0,93	0,92	0,89	0,96	0,98
Proteins_FNR Limit [%]	16,43	13,25	20,89	28,31	23,7	18,02
Proteins_Confidence Limit [%]	85	79,41	91,89	84	73,24	89,32
Proteins_PEP Limit [%]	15	20,59	8,11	16	26,76	10,68
Proteins_Confidence Accuracy [%]	2,94	3,03	2,78	4,17	1,41	0,97
Peptides_#Validated	2324	3338	2967	2665	3633	3561
Peptides_Total Possible TP	3453,1	4499,67	4299,22	4501,07	6579,17	6136,7
Peptides_FDR Limit [%]	0,99	0,99	0,98	0,98	0,99	0,98
Peptides_FNR Limit [%]	33,37	26,53	31,68	41,41	45,37	42,61
Peptides_Confidence Limit [%]	94,34	91,43	94,74	90,79	92,26	90,38
Peptides_PEP Limit [%]	5,66	8,57	5,26	9,21	7,74	9,62
Peptides_Confidence Accuracy [%]	0,96	1,01	0,53	0,22	0,17	0,14
PSMs (Charge 2)_#Validated PSM	6032	8389	8307	6870	8920	9715
PSMs (Charge 1)_#Validated PSM	918	1109	929	1160	1400	1109
PSMs (Charge 3)_#Validated PSM	104	111	136	116	100	120
PSMs (Charge 2)_Total Possible TP	8559,7	11977,2	13001,23	12300,5	15577,3	16438,74
PSMs (Charge 1)_Total Possible TP	2052,1	2035,65	1901,43	2182,87	2186,39	2104,89
PSMs (Charge 3)_Total Possible TP	197,53	240,92	234,45	669,84	745,63	985,84
PSMs (Charge 2)_FDR Limit [%]	0,99	0,99	1	0,98	0,95	1
PSMs (Charge 1)_FDR Limit [%]	0,98	0,99	0,97	0,95	1	0,99
PSMs (Charge 3)_FDR Limit [%]	0,96	0,9	0,74	0,86	1	0,83
PSMs (Charge 2)_FNR Limit [%]	30,32	30,67	36,78	44,75	43,3	41,53
PSMs (Charge 1)_FNR Limit [%]	55,69	46,08	51,69	47,36	36,65	47,83
PSMs (Charge 3)_FNR Limit [%]	48,16	54,6	42,76	83,24	86,88	88,05
PSMs (Charge 2)_Confidence Limit [%]	88,63	90,18	92,84	87,98	92,46	89,29
PSMs (Charge 1)_Confidence Limit [%]	94,12	93,68	92,36	95,97	95,39	94,5
PSMs (Charge 3)_Confidence Limit [%]	80	84,21	80	86,32	92,47	90,8
PSMs (Charge 2)_PEP Limit [%]	11,37	9,82	7,16	12,02	7,54	10,71
PSMs (Charge 1)_PEP Limit [%]	5,88	6,32	7,64	4,03	4,61	5,5
PSMs (Charge 3)_PEP Limit [%]	20	15,79	20	13,68	7,53	9,2
PSMs (Charge 2)_Confidence Accuracy [%]	0,12	0,31	0,13	0,1	0,11	0,09
PSMs (Charge 1)_Confidence Accuracy [%]	1,22	1,14	0,64	0,83	0,77	0,52
PSMs (Charge 3)_Confidence Accuracy [%]	6,67	5,26	6,67	1,06	1,1	1,18

	MS Amanda			MSGF+		
	121	**259**	**285**	**121**	**259**	**285**
Proteins_#Validated	706	816	805	537	728	675
Proteins_Total Possible TP	885,77	1182,19	1178,72	891,5	1076,08	986,21
Proteins_FDR Limit [%]	0,99	0,98	0,99	0,93	0,96	0,89
Proteins_FNR Limit [%]	21	31,7	32,42	40,16	32,99	32,21
Proteins_Confidence Limit [%]	89,8	95,38	97,47	93,67	93,75	88,37
Proteins_PEP Limit [%]	10,2	4,62	2,53	6,33	6,25	11,63
Proteins_Confidence Accuracy [%]	2,33	1,54	1,54	1,92	2,13	2,38
Peptides_#Validated	1727	2272	2253	1619	2272	1996
Peptides_Total Possible TP	3025,23	4261,3	3996,21	2362,57	4214,72	3755,99
Peptides_FDR Limit [%]	0,98	0,97	0,98	0,99	0,97	0,95
Peptides_FNR Limit [%]	43,51	47,14	44,19	32,09	46,55	47,31
Peptides_Confidence Limit [%]	95,58	96,33	94,84	94,21	94,55	93,45
Peptides_PEP Limit [%]	4,42	3,67	5,16	5,79	5,45	6,55
Peptides_Confidence Accuracy [%]	0,99	0,44	0,32	0,83	0,32	0,3
PSMs (Charge 2)_#Validated PSM	3952	5469	5436	3224	4632	3915
PSMs (Charge 1)_#Validated PSM	547	517	280	503	602	233
PSMs (Charge 3)_#Validated PSM	22	36	111	111	87	81
PSMs (Charge 2)_Total Possible TP	6445,27	9236,61	9884,98	5371,29	7915,02	7398,17
PSMs (Charge 1)_Total Possible TP	1733,18	1610,28	1450,34	1359,21	1657,96	825,96
PSMs (Charge 3)_Total Possible TP	138,72	244,8	232,61	153,92	187,42	146,41
PSMs (Charge 2)_FDR Limit [%]	0,99	0,97	0,99	0,99	0,99	1
PSMs (Charge 1)_FDR Limit [%]	0,91	0,77	0,71	0,99	1	0,86
PSMs (Charge 3)_FDR Limit [%]	0	0	0,9	0,9	0	0
PSMs (Charge 2)_FNR Limit [%]	39,3	41,45	45,64	40,55	42,13	47,64
PSMs (Charge 1)_FNR Limit [%]	68,85	68,19	80,98	63,4	64,08	72,04
PSMs (Charge 3)_FNR Limit [%]	84,93	85,5	53,05	29,71	54,06	45,16
PSMs (Charge 2)_Confidence Limit [%]	89,39	91,66	92,77	92,57	90,92	92,23
PSMs (Charge 1)_Confidence Limit [%]	90,53	91,55	94,59	94,86	93,46	93,62
PSMs (Charge 3)_Confidence Limit [%]	93,94	97,78	91,3	83,78	89,47	88,24
PSMs (Charge 2)_PEP Limit [%]	10,61	8,34	7,23	7,43	9,08	7,77
PSMs (Charge 1)_PEP Limit [%]	9,47	8,45	5,41	5,14	6,54	6,38
PSMs (Charge 3)_PEP Limit [%]	6,06	2,22	8,7	16,22	10,53	11,76
PSMs (Charge 2)_Confidence Accuracy [%]	0,27	0,12	0,07	0,31	0,11	0,14
PSMs (Charge 1)_Confidence Accuracy [%]	1,06	1,41	0,91	0,32	0,94	2,17
PSMs (Charge 3)_Confidence Accuracy [%]	3,13	2,22	4,35	2,78	5,56	6,25

	4-Multiple Search Comet-xTandem-MSamanda-MSGF+			**MyriMatch**		
	121	**259**	**285**	**121**	**259**	**285**
Proteins_#Validated	931	1205	1218	609	752	772
Proteins_Total Possible TP	1170,11	1503,4	1499,35	835,59	1219,73	1199,12
Proteins_FDR Limit [%]	0,32	1	0,99	0,99	0,93	0,91
Proteins_FNR Limit [%]	20,67	20,66	19,54	27,9	39,04	36,41
Proteins_Confidence Limit [%]	86	83,87	86,11	90,48	86,05	90,24
Proteins_PEP Limit [%]	14	16,13	13,89	9,52	13,95	9,76
Proteins_Confidence Accuracy [%]	2,04	3,33	3,85	5	2,7	1,25
Peptides_#Validated	3022	4075	4063	1451	2210	1551
Peptides_Total Possible TP	4301,1	5601,97	5347,83	3307,16	4663,92	4130,06
Peptides_FDR Limit [%]	0,99	0,98	0,98	0,96	1	0,97
Peptides_FNR Limit [%]	30,33	27,92	24,69	56,45	53,08	62,82
Peptides_Confidence Limit [%]	90,71	93,33	93,09	90,23	95,52	95,94
Peptides_PEP Limit [%]	9,29	6,67	6,91	9,77	4,48	4,06
Peptides_Confidence Accuracy [%]	0,21	0,24	0,57	0,81	0,45	0,51
PSMs (Charge 2)_#Validated PSM	8203	10859	11420	4053	5723	5124
PSMs (Charge 1)_#Validated PSM	1522	1607	1224	525	519	161
PSMs (Charge 3)_#Validated PSM	252	255	303	36	57	77
PSMs (Charge 2)_Total Possible TP	13136,5	16492,1	15375,98	8135,65	13048,52	11931,76
PSMs (Charge 1)_Total Possible TP	2614,66	2357,56	2355,82	1697,97	1905,57	1360,1
PSMs (Charge 3)_Total Possible TP	416,15	526,78	601,19	152,09	202,34	117,9
PSMs (Charge 2)_FDR Limit [%]	1	0,99	1	0,99	1	1
PSMs (Charge 1)_FDR Limit [%]	0,99	1	0,98	0,95	0,96	0,62
PSMs (Charge 3)_FDR Limit [%]	0,79	0,39	0,99	0	0	0
PSMs (Charge 2)_FNR Limit [%]	38,23	34,82	26,48	50,7	56,62	57,5
PSMs (Charge 1)_FNR Limit [%]	42,56	32,57	48,55	69,39	73,05	88,27
PSMs (Charge 3)_FNR Limit [%]	44,41	52,93	52,43	76,65	72,07	35,78
PSMs (Charge 2)_Confidence Limit [%]	84,52	88,5	88,81	94,65	93,68	95,12
PSMs (Charge 1)_Confidence Limit [%]	90,25	92,36	95,93	95,52	98,4	97,22
PSMs (Charge 3)_Confidence Limit [%]	78,94	91,35	84,04	97,56	95,65	76,19
PSMs (Charge 2)_PEP Limit [%]	15,48	11,5	11,19	5,35	6,32	4,88
PSMs (Charge 1)_PEP Limit [%]	9,75	7,64	4,07	4,48	1,6	2,78
PSMs (Charge 3)_PEP Limit [%]	21,06	8,65	15,96	2,44	4,35	23,81
PSMs (Charge 2)_Confidence Accuracy [%]	0,13	0,09	0,14	0,27	0,07	0,14
PSMs (Charge 1)_Confidence Accuracy [%]	0,28	0,64	0,37	1,52	0,81	1,41
PSMs (Charge 3)_Confidence Accuracy [%]	0,25	0,31	0,26	2,5	4,55	4,76

	Comet			X! Tandem		
	240	**495**	**628**	**240**	**495**	**628**
Proteins_#Validated	843	759	885	832	804	837
Proteins_Total Possible TP	946,66	966,13	1078,03	1041,59	1043,4	1146,08
Proteins_FDR Limit [%]	0,95	0,79	0,9	0,96	1	0,96
Proteins_FNR Limit [%]	11,84	22,07	18,77	20,16	23,86	27,84
Proteins_Confidence Limit [%]	88,24	87,88	90,32	98,51	78,05	85,11
Proteins_PEP Limit [%]	11,76	12,12	9,68	1,49	21,95	14,89
Proteins_Confidence Accuracy [%]	3,13	3,13	4	1,18	2,5	2,17
Peptides_#Validated	2155	2059	2456	2506	2655	2728
Peptides_Total Possible TP	2973,45	3226,87	3665,03	5198,26	5742,1	4638,54
Peptides_FDR Limit [%]	0,97	0,97	0,98	1	0,98	0,99
Peptides_FNR Limit [%]	28,28	36,8	33,67	52,21	54,26	41,73
Peptides_Confidence Limit [%]	89,91	94,39	90,65	89,56	92,43	94,9
Peptides_PEP Limit [%]	10,09	5,61	9,35	10,44	7,57	5,1
Peptides_Confidence Accuracy [%]	0,97	0,97	0,93	0,16	0,16	0,14
PSMs (Charge 2)_#Validated PSM	5681	6394	7118	6578	7894	8009
PSMs (Charge 1)_#Validated PSM	790	481	962	915	753	1122
PSMs (Charge 3)_#Validated PSM	45	89	70	15	126	72
PSMs (Charge 2)_Total Possible TP	8376,64	10587,47	11643,14	11859,89	14709,88	14860,75
PSMs (Charge 1)_Total Possible TP	1636,48	1164,05	2013,23	1894,82	1443,74	2139,95
PSMs (Charge 3)_Total Possible TP	82,41	182,31	198,06	235,76	357,19	449,08
PSMs (Charge 2)_FDR Limit [%]	0,99	0,99	1	0,91	0,91	1
PSMs (Charge 1)_FDR Limit [%]	0,89	0,83	0,94	0,98	0,93	0,98
PSMs (Charge 3)_FDR Limit [%]	0	0	0	0	0,79	0
PSMs (Charge 2)_FNR Limit [%]	32,88	40,2	39,49	45,06	46,88	46,64
PSMs (Charge 1)_FNR Limit [%]	52,18	59,04	52,73	52,25	48,56	48,11
PSMs (Charge 3)_FNR Limit [%]	45,95	51,73	64,89	93,9	65,66	84,32
PSMs (Charge 2)_Confidence Limit [%]	89,25	91,18	91,66	89,62	90,93	92,23
PSMs (Charge 1)_Confidence Limit [%]	93,79	89,87	93,88	95,33	90,55	95,04
PSMs (Charge 3)_Confidence Limit [%]	92,31	76,92	93,33	94,37	89,47	95,24
PSMs (Charge 2)_PEP Limit [%]	10,75	8,82	8,34	10,38	9,07	7,77
PSMs (Charge 1)_PEP Limit [%]	6,21	10,13	6,12	4,67	9,45	4,96
PSMs (Charge 3)_PEP Limit [%]	7,69	23,08	6,67	5,63	10,53	4,76
PSMs (Charge 2)_Confidence Accuracy [%]	0,24	0,19	0,13	0,47	0,11	0,09
PSMs (Charge 1)_Confidence Accuracy [%]	0,69	1,27	0,68	0,97	0,51	0,75
PSMs (Charge 3)_Confidence Accuracy [%]	7,69	7,69	6,67	0,9	1,06	0,96

	MS Amanda			MSGF+		
	240	**495**	**628**	**240**	**495**	**628**
Proteins_#Validated	639	624	685	585	565	685
Proteins_Total Possible TP	947,76	983,13	1007,71	787,16	760,73	913,12
Proteins_FDR Limit [%]	0,78	0,96	0,88	0,85	0,88	0,88
Proteins_FNR Limit [%]	33,17	37,1	32,69	26,47	25,92	25,38
Proteins_Confidence Limit [%]	86,05	83,1	86,08	92,31	81,4	93,33
Proteins_PEP Limit [%]	13,95	16,9	13,92	7,69	18,6	6,67
Proteins_Confidence Accuracy [%]	2,33	1,43	1,28	2,44	2,33	2,27
Peptides_#Validated	1685	1524	1709	1475	1281	1735
Peptides_Total Possible TP	3241,3	3867,69	3416,86	2485,33	2421,54	2938,23
Peptides_FDR Limit [%]	0,95	0,98	0,99	0,95	0,94	0,92
Peptides_FNR Limit [%]	48,38	60,98	50,42	41,15	47,5	41,33
Peptides_Confidence Limit [%]	98,02	95,5	95,24	90,91	90,77	94,92
Peptides_PEP Limit [%]	1,98	4,5	4,76	9,09	9,23	5,08
Peptides_Confidence Accuracy [%]	1,12	0,84	0,68	0,73	1,54	1,11
PSMs (Charge 2)_#Validated PSM	3915	4237	4649	3071	3260	4032
PSMs (Charge 1)_#Validated PSM	526	261	420	306	152	333
PSMs (Charge 3)_#Validated PSM	59	114	118	61	50	61
PSMs (Charge 2)_Total Possible TP	6130,01	7943,17	8618,46	4720,11	5769,93	6747,23
PSMs (Charge 1)_Total Possible TP	1299,23	916,12	1591,59	955,53	780,55	1253,18
PSMs (Charge 3)_Total Possible TP	111,52	177,73	199,22	99,29	128,13	125,17
PSMs (Charge 2)_FDR Limit [%]	1	0,99	0,99	0,98	0,98	0,99
PSMs (Charge 1)_FDR Limit [%]	0,95	0,77	0,95	0,98	0,66	0,9
PSMs (Charge 3)_FDR Limit [%]	0	0,88	0,85	0	0	0
PSMs (Charge 2)_FNR Limit [%]	36,79	47,2	46,59	35,6	44,06	40,86
PSMs (Charge 1)_FNR Limit [%]	59,94	71,77	73,86	68,38	80,72	73,7
PSMs (Charge 3)_FNR Limit [%]	47,52	36,57	42,32	39,22	61,34	51,65
PSMs (Charge 2)_Confidence Limit [%]	91,33	93,68	94,13	95,47	94,48	90,12
PSMs (Charge 1)_Confidence Limit [%]	96,92	90,24	97,33	95,35	98,11	96,49
PSMs (Charge 3)_Confidence Limit [%]	94,74	80	82,93	88,24	93,33	95,65
PSMs (Charge 2)_PEP Limit [%]	8,67	6,32	5,87	4,53	5,52	9,88
PSMs (Charge 1)_PEP Limit [%]	3,08	9,76	2,67	4,65	1,89	3,51
PSMs (Charge 3)_PEP Limit [%]	5,26	20	17,07	11,76	6,67	4,35
PSMs (Charge 2)_Confidence Accuracy [%]	0,68	0,37	0,18	0,41	0,62	0,41
PSMs (Charge 1)_Confidence Accuracy [%]	1,56	2,5	1,33	2,33	1,89	1,79
PSMs (Charge 3)_Confidence Accuracy [%]	5,56	6,67	2,44	6,25	7,14	4,55

	4-Multiple Search Comet-xTandem-MSamanda-MSGF+			**MyriMatch**		
	240	**495**	**628**	**240**	**495**	**628**
Proteins_#Validated	915	946	988	587	589	583
Proteins_Total Possible TP	1125,92	1164,4	1240,04	980,42	979,44	1027,79
Proteins_FDR Limit [%]	0,98	0,95	0,91	0,85	0,68	0,86
Proteins_FNR Limit [%]	19,36	19,42	21,05	40,72	40,41	43,87
Proteins_Confidence Limit [%]	91,89	76,47	60,87	88,24	75,68	87,8
Proteins_PEP Limit [%]	8,11	23,53	39,13	11,76	24,32	12,2
Proteins_Confidence Accuracy [%]	2,78	3,13	4,35	2	2,78	2,5
Peptides_#Validated	2893	3034	3170	1219	1235	1074
Peptides_Total Possible TP	3996,69	4159,02	5020,42	2967,09	3139,51	3609,05
Peptides_FDR Limit [%]	0,93	0,99	0,98	0,98	0,97	0,93
Peptides_FNR Limit [%]	28,28	27,57	37,31	59,3	61,03	70,49
Peptides_Confidence Limit [%]	89,46	92,27	93,07	96,58	95,08	96,62
Peptides_PEP Limit [%]	10,54	7,73	6,93	3,42	4,92	3,38
Peptides_Confidence Accuracy [%]	0,28	0,48	0,33	0,85	1,67	0,42
PSMs (Charge 2)_#Validated PSM	7737	9624	9780	3904	3906	4588
PSMs (Charge 1)_#Validated PSM	1220	805	1273	106	158	157
PSMs (Charge 3)_#Validated PSM	135	255	241	54	27	39
PSMs (Charge 2)_Total Possible TP	12403,54	13378,11	15791,96	8402,57	10668,96	11544,42
PSMs (Charge 1)_Total Possible TP	1692,2	1341,91	2180,03	1363,55	1041,66	1594,92
PSMs (Charge 3)_Total Possible TP	313,83	456,7	464,3	83,88	93,13	100,05
PSMs (Charge 2)_FDR Limit [%]	0,94	0,99	0,98	1	1	0,98
PSMs (Charge 1)_FDR Limit [%]	0,98	0,99	0,86	0	0,63	0,64
PSMs (Charge 3)_FDR Limit [%]	0	0,78	0,83	0	0	0
PSMs (Charge 2)_FNR Limit [%]	38,28	28,87	38,7	54,01	63,76	60,66
PSMs (Charge 1)_FNR Limit [%]	28,65	40,62	42,16	92,3	85,05	90,24
PSMs (Charge 3)_FNR Limit [%]	72,03	48,02	52,11	38,58	71,87	61,88
PSMs (Charge 2)_Confidence Limit [%]	87,23	83,15	88,73	93,72	97,02	94,87
PSMs (Charge 1)_Confidence Limit [%]	93,22	92,68	90,8	98,43	96,61	95,38
PSMs (Charge 3)_Confidence Limit [%]	56,25	81,98	82,96	56,52	86,67	90,48
PSMs (Charge 2)_PEP Limit [%]	12,77	16,85	11,27	6,28	2,98	5,13
PSMs (Charge 1)_PEP Limit [%]	6,78	7,32	9,2	1,57	3,39	4,62
PSMs (Charge 3)_PEP Limit [%]	43,75	18,02	17,04	43,48	13,33	9,52
PSMs (Charge 2)_Confidence Accuracy [%]	0,09	0,15	0,16	0,42	0,34	0,15
PSMs (Charge 1)_Confidence Accuracy [%]	1,69	2,5	1,16	0,79	0,85	1,56
PSMs (Charge 3)_Confidence Accuracy [%]	0,16	0,25	0,24	4,55	6,67	5

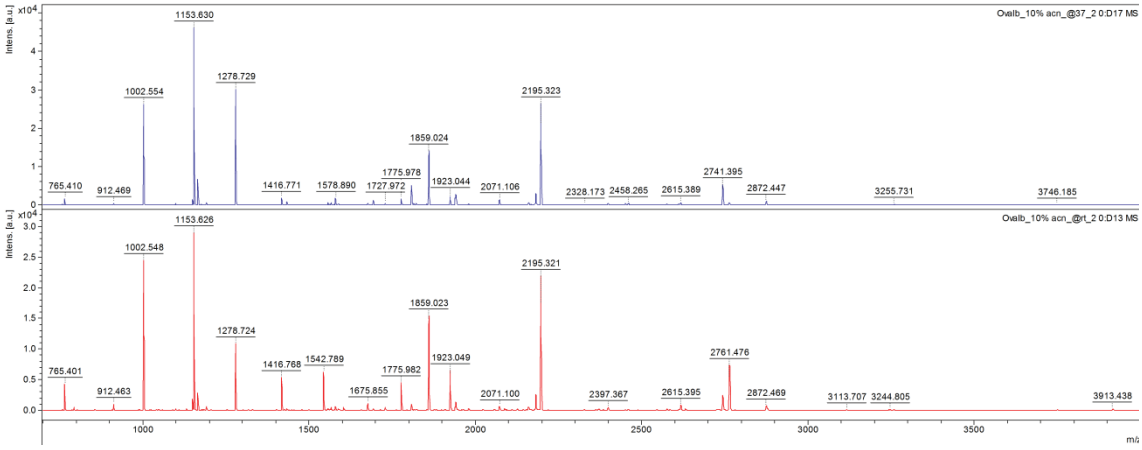
	Comet			X! Tandem		
	615	**Crt01**	**Crt02**	**615**	**Crt01**	**Crt02**
Proteins_#Validated	854	847	905	866	906	908
Proteins_Total Possible TP	1154,67	1147,16	1110,01	1255,44	1158,24	1179,84
Proteins_FDR Limit [%]	0,94	0,94	0,99	0,92	0,99	0,88
Proteins_FNR Limit [%]	26,64	26,89	19,2	31,75	22,5	23,74
Proteins_Confidence Limit [%]	83,67	91,8	90	91,59	86,42	88,68
Proteins_PEP Limit [%]	16,33	8,2	10	8,41	13,58	11,32
Proteins_Confidence Accuracy [%]	2,04	1,72	6,25	0,93	1,23	1,89
Peptides_#Validated	2305	2304	2769	2874	2505	3231
Peptides_Total Possible TP	3480,36	3547,85	3789,05	5416,69	4284,45	4603,92
Peptides_FDR Limit [%]	1	0,95	0,98	0,97	1	0,96
Peptides_FNR Limit [%]	34,47	35,81	27,64	47,52	42,13	30,58
Peptides_Confidence Limit [%]	92,5	88,31	93,6	91,1	91,49	88,13
Peptides_PEP Limit [%]	7,5	11,69	6,4	8,9	8,51	11,87
Peptides_Confidence Accuracy [%]	0,85	0,33	0,85	0,21	0,26	0,22
PSMs (Charge 2)_#Validated PSM	5149	4259	10654	6316	4807	11513
PSMs (Charge 1)_#Validated PSM	516	530	855	576	588	852
PSMs (Charge 3)_#Validated PSM	103	88	401	108	86	375
PSMs (Charge 2)_Total Possible TP	8759,56	7139,21	14471,78	11813,9	8710,01	15989,63
PSMs (Charge 1)_Total Possible TP	1150,25	843,79	1554,59	1288,16	897,81	1636,66
PSMs (Charge 3)_Total Possible TP	302,49	142,01	764,65	396,26	211,21	953,14
PSMs (Charge 2)_FDR Limit [%]	0,99	0,99	0,99	0,98	0,92	1
PSMs (Charge 1)_FDR Limit [%]	0,97	0,94	0,94	0,69	0,68	0,82
PSMs (Charge 3)_FDR Limit [%]	0,97	0	1	0,93	0	0,8
PSMs (Charge 2)_FNR Limit [%]	41,84	40,93	27,13	47,06	45,33	28,91
PSMs (Charge 1)_FNR Limit [%]	55,6	37,79	45,6	55,67	35,48	48,49
PSMs (Charge 3)_FNR Limit [%]	66,49	38,73	48,21	73,26	60,15	61,06
PSMs (Charge 2)_Confidence Limit [%]	93,37	91,93	90,43	92,28	94,87	85,47
PSMs (Charge 1)_Confidence Limit [%]	97,02	92,59	91,6	97,49	92,9	96,18
PSMs (Charge 3)_Confidence Limit [%]	95,74	85,71	97,5	93,42	87,93	92,39
PSMs (Charge 2)_PEP Limit [%]	6,63	8,07	9,57	7,72	5,13	14,53
PSMs (Charge 1)_PEP Limit [%]	2,98	7,41	8,4	2,51	7,1	3,82
PSMs (Charge 3)_PEP Limit [%]	4,26	14,29	2,5	6,58	12,07	7,61
PSMs (Charge 2)_Confidence Accuracy [%]	0,18	0,11	0,12	0,12	0,41	0,05
PSMs (Charge 1)_Confidence Accuracy [%]	0,6	1,89	0,85	0,43	0,3	0,75
PSMs (Charge 3)_Confidence Accuracy [%]	2,13	4,76	1,27	1,37	1,75	1,1

	MS Amanda			MSGF+		
	615	**Crt01**	**Crt02**	**615**	**Crt01**	**Crt02**
Proteins_#Validated	675	728	778	464	510	719
Proteins_Total Possible TP	985,15	875,73	1041,54	829,24	854,59	932,49
Proteins_FDR Limit [%]	0,89	0,96	0,9	0,86	0,98	0,97
Proteins_FNR Limit [%]	32,25	17,61	26,03	44,46	42,23	23,67
Proteins_Confidence Limit [%]	79,07	82,46	70,27	92,93	88,14	88,41
Proteins_PEP Limit [%]	20,93	17,54	29,73	7,07	11,86	11,59
Proteins_Confidence Accuracy [%]	2,63	2,04	2,7	1,02	0,13	2
Peptides_#Validated	1526	1631	2052	1358	1403	2415
Peptides_Total Possible TP	3300,16	3132,67	3578,45	3048,74	2992,29	3625,46
Peptides_FDR Limit [%]	0,98	0,98	0,97	0,88	0,93	0,99
Peptides_FNR Limit [%]	54,18	48,38	43,18	55,83	53,51	33,96
Peptides_Confidence Limit [%]	95,35	93,52	94,44	95,74	94,53	91,98
Peptides_PEP Limit [%]	4,65	6,48	5,56	4,26	5,47	8,02
Peptides_Confidence Accuracy [%]	0,74	0,34	0,75	0,43	0,35	0,54
PSMs (Charge 2)_#Validated PSM	2850	2840	7612	2374	2164	6610
PSMs (Charge 1)_#Validated PSM	28	129	415	230	217	205
PSMs (Charge 3)_#Validated PSM	113	43	217	105	75	282
PSMs (Charge 2)_Total Possible TP	6608,02	5323,04	11393,74	5076,73	3898,34	10095,56
PSMs (Charge 1)_Total Possible TP	898,92	711,64	1272,29	741,52	519,94	1045,45
PSMs (Charge 3)_Total Possible TP	254,24	201,11	643,57	179,42	154,71	481,9
PSMs (Charge 2)_FDR Limit [%]	0,98	0,99	1	0,97	0,97	1
PSMs (Charge 1)_FDR Limit [%]	0	0,78	0,96	0,87	0,92	0,98
PSMs (Charge 3)_FDR Limit [%]	0,88	0	0,92	0,95	0	0,71
PSMs (Charge 2)_FNR Limit [%]	57,3	47,2	33,87	53,7	45,06	35,2
PSMs (Charge 1)_FNR Limit [%]	96,93	82,16	67,77	69,37	58,82	80,61
PSMs (Charge 3)_FNR Limit [%]	56,12	78,86	66,61	42,39	51,99	41,9
PSMs (Charge 2)_Confidence Limit [%]	95,96	93,13	92,64	94,12	92,24	91,24
PSMs (Charge 1)_Confidence Limit [%]	98,85	98,52	95,92	96,83	87,1	96,36
PSMs (Charge 3)_Confidence Limit [%]	96	97,14	92,16	89,47	90,48	93,55
PSMs (Charge 2)_PEP Limit [%]	4,04	6,87	7,36	5,88	7,76	8,76
PSMs (Charge 1)_PEP Limit [%]	1,15	1,48	4,08	3,17	12,9	3,64
PSMs (Charge 3)_PEP Limit [%]	4	2,86	7,84	10,53	9,52	6,45
PSMs (Charge 2)_Confidence Accuracy [%]	0,24	0,34	0,23	0,26	0,28	0,2
PSMs (Charge 1)_Confidence Accuracy [%]	0,85	0,75	0,41	1,61	3,23	1,82
PSMs (Charge 3)_Confidence Accuracy [%]	4,17	2,94	2	5,56	5	3,23

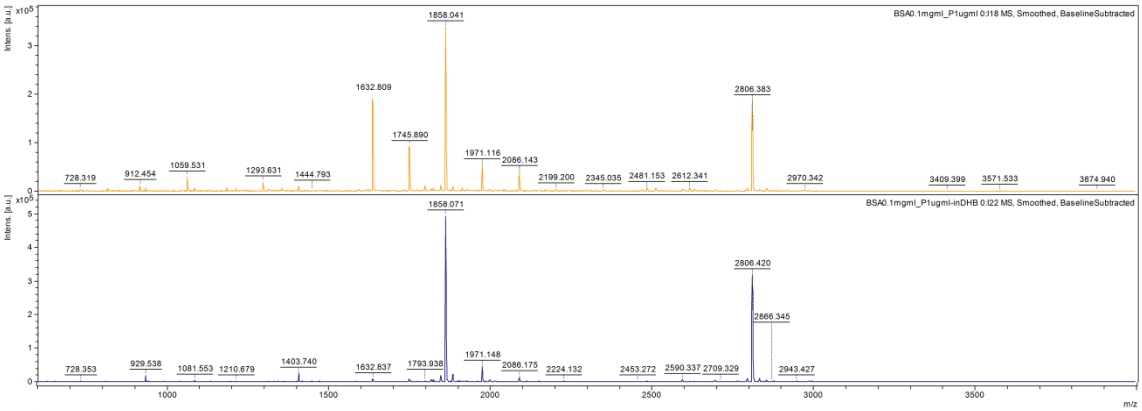
	4-Multiple Search Comet-xTandem-MSamanda-MSGF+			**MyriMatch**		
	615	**Crt01**	**Crt02**	**615**	**Crt01**	**Crt02**
Proteins_#Validated	1009	1068	1055	595	637	693
Proteins_Total Possible TP	1265,96	1318,98	1204,12	997	1047,3	1042,73
Proteins_FDR Limit [%]	0,99	0,94	0,95	0,84	0,94	0,87
Proteins_FNR Limit [%]	21,05	19,78	13,32	40,85	39,79	34,08
Proteins_Confidence Limit [%]	84,62	87,18	81,25	87,88	95,65	69,57
Proteins_PEP Limit [%]	15,38	12,82	18,75	12,12	4,35	30,43
Proteins_Confidence Accuracy [%]	2,33	2,56	3,7	3,03	4,35	4,35
Peptides_#Validated	3242	3330	3544	1175	1382	1971
Peptides_Total Possible TP	4551,01	4382,89	4356,61	3063,55	3121,99	4125
Peptides_FDR Limit [%]	0,99	0,99	0,99	0,94	0,94	0,96
Peptides_FNR Limit [%]	29,42	24,71	19,28	61,93	56,05	52,64
Peptides_Confidence Limit [%]	92,17	90,87	91,38	93,65	92,45	93,48
Peptides_PEP Limit [%]	7,83	9,13	8,62	6,35	7,55	6,52
Peptides_Confidence Accuracy [%]	0,62	0,38	0,93	1,61	1,01	0,74
PSMs (Charge 2)_#Validated PSM	7422	7520	13027	2567	2624	7456
PSMs (Charge 1)_#Validated PSM	752	768	1216	112	143	232
PSMs (Charge 3)_#Validated PSM	300	237	728	43	63	171
PSMs (Charge 2)_Total Possible TP	10978,21	12036,39	16120,92	7171,36	6120,15	14680,33
PSMs (Charge 1)_Total Possible TP	1342,84	1143,78	1748,95	818,95	649,94	1394,86
PSMs (Charge 3)_Total Possible TP	613,29	495,07	998,41	133,89	135	368,21
PSMs (Charge 2)_FDR Limit [%]	1	1	1	0,97	0,99	0,99
PSMs (Charge 1)_FDR Limit [%]	0,93	0,91	0,99	0,89	0,7	0,86
PSMs (Charge 3)_FDR Limit [%]	1	0,84	0,96	0	0	0,58
PSMs (Charge 2)_FNR Limit [%]	33,1	38,19	20	64,56	57,56	49,73
PSMs (Charge 1)_FNR Limit [%]	44,64	33,48	31,23	86,5	78,36	83,59
PSMs (Charge 3)_FNR Limit [%]	52,23	54,11	27,86	68,25	54,12	54,08
PSMs (Charge 2)_Confidence Limit [%]	89	88,44	83,92	94,12	92,25	94,87
PSMs (Charge 1)_Confidence Limit [%]	93,79	93,1	93,25	98,04	96,4	97,09
PSMs (Charge 3)_Confidence Limit [%]	90,48	82,19	85,71	97,14	80	90,48
PSMs (Charge 2)_PEP Limit [%]	11	11,56	16,08	5,88	7,75	5,13
PSMs (Charge 1)_PEP Limit [%]	6,21	6,9	6,75	1,96	3,6	2,91
PSMs (Charge 3)_PEP Limit [%]	9,52	17,81	14,29	2,86	20	9,52
PSMs (Charge 2)_Confidence Accuracy [%]	0,24	0,12	0,14	0,85	0,71	0,15
PSMs (Charge 1)_Confidence Accuracy [%]	0,63	1,16	0,62	2	0,9	0,97
PSMs (Charge 3)_Confidence Accuracy [%]	0,37	0,41	1,1	2,86	7,14	5

APPENDIX E

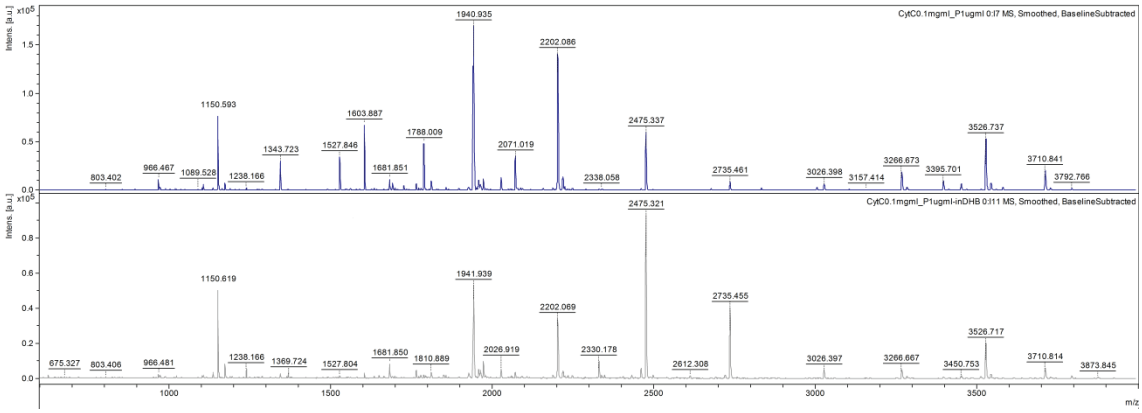
TEST RESULTS OF ENZYME-MATRIX ASSOCIATION



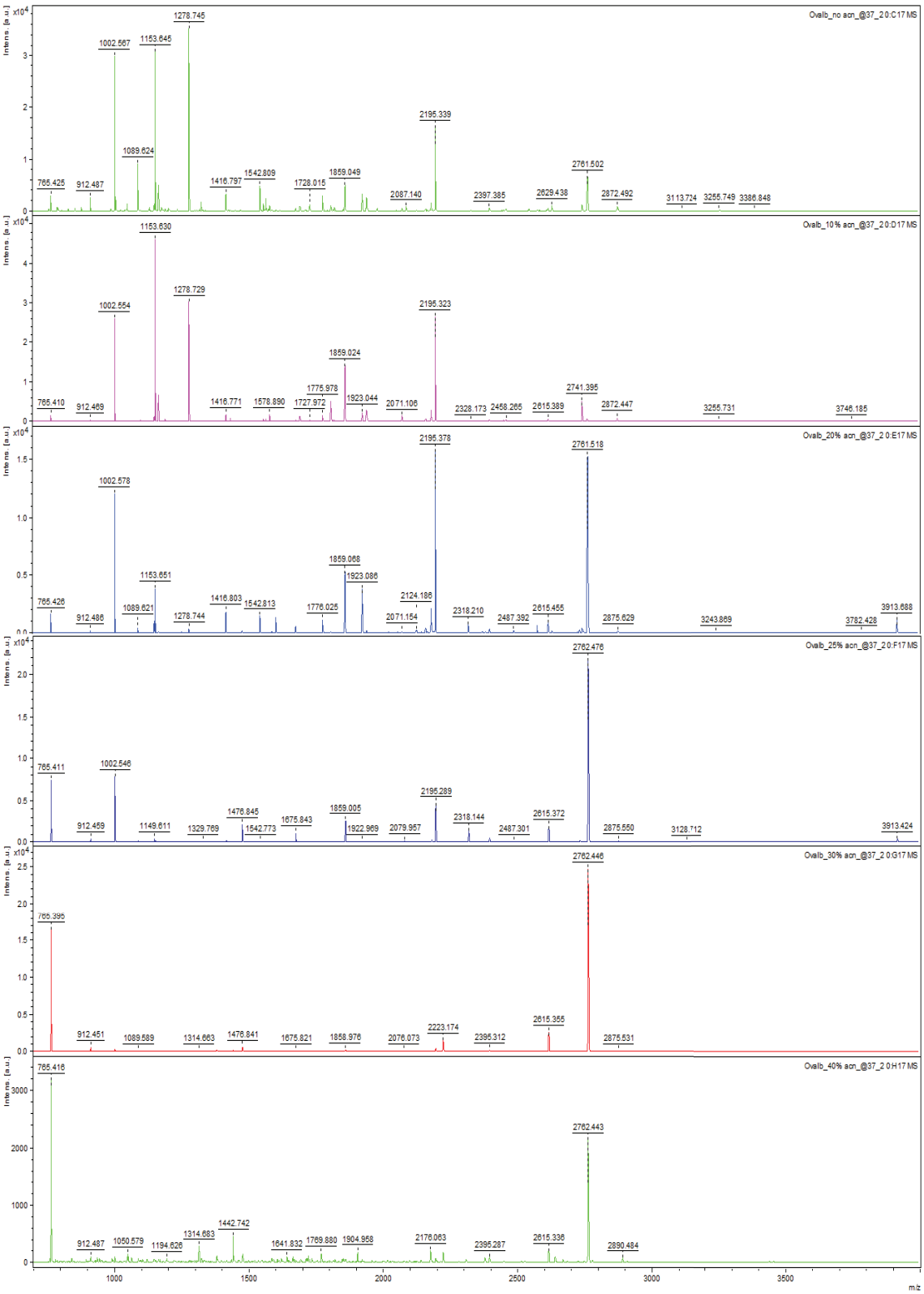
Ovalbumin protein digested at 37 °C and at room temperature (down)



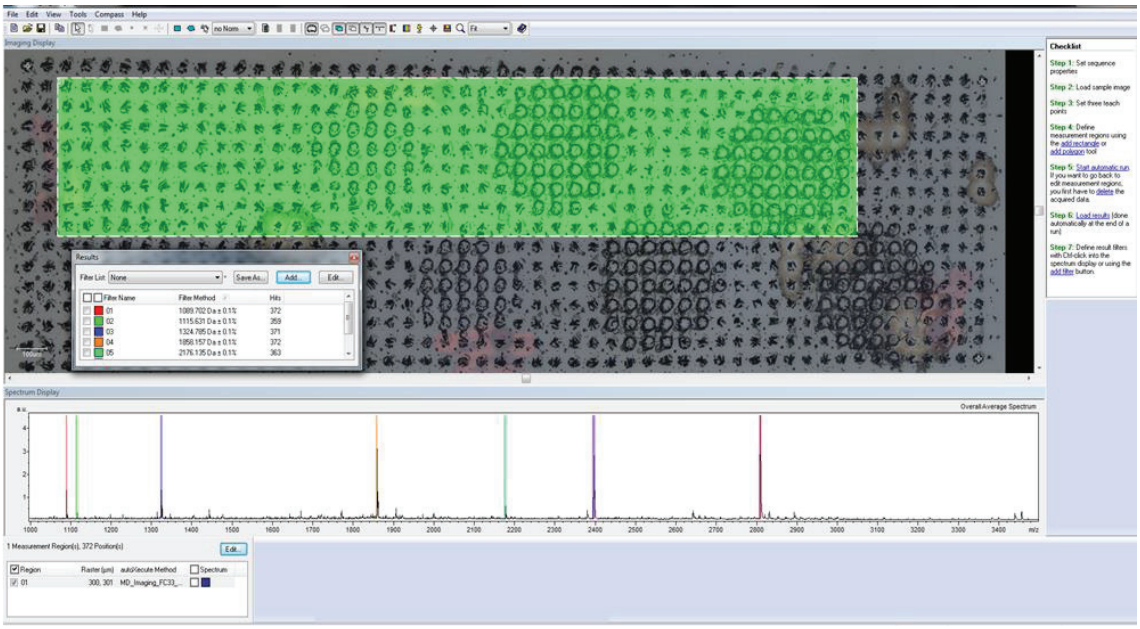
BSA protein digested by pepsin enzyme with (down) and without DHB matrix



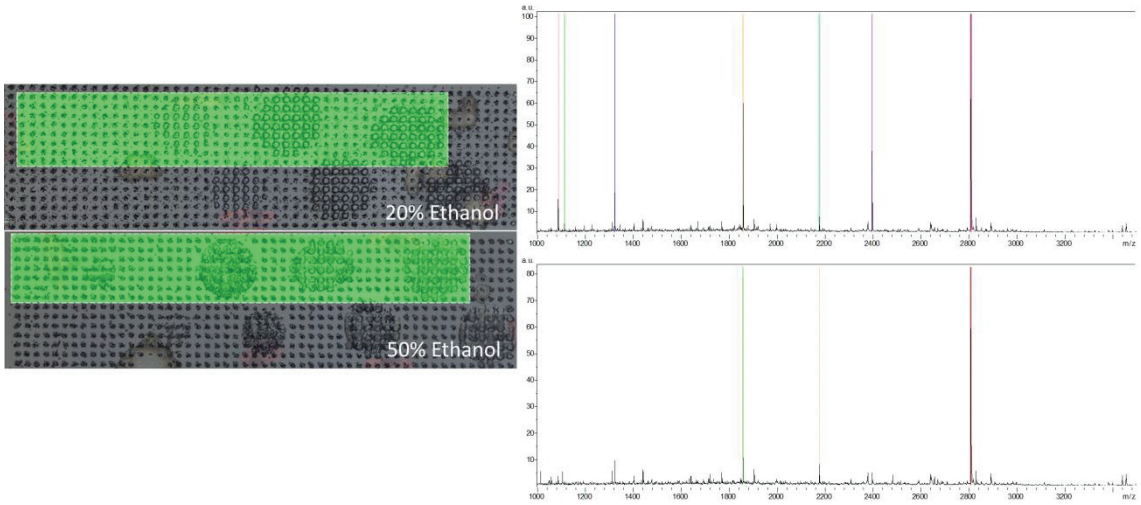
Cytochrome c protein digested by pepsin enzyme with (down) and without DHB matrix



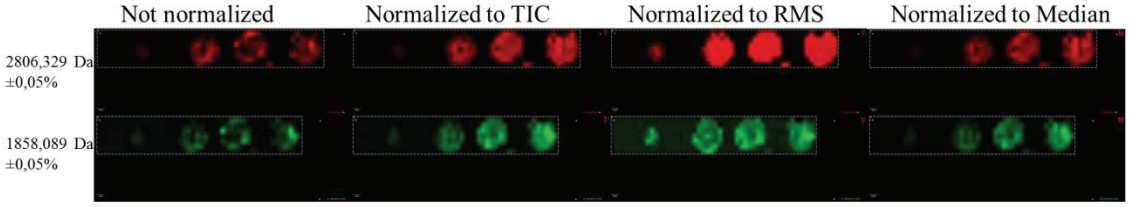
Acetonitrile content in the digestion medium from 0% to 40%



Sample screen shot of flexImaging graphical userface with acquired data



Spectral and optical image differences between matrix prepared in 20% and 50% EtOH



MSI image from BSA protein after matrix+pepsin application by printing →50% EtOH

VITA

This thesis was written and submitted by Melike DİNÇ who was born in 1986 in Bursa. She graduated from the Department of Biochemistry at the Faculty of Science in Ege University in June 2009. She earned her Master's degree in chemistry from Izmir Institue of Technology in December 2012. She has been studying mass spectrometry-based proteomics since 2010 thus attended several conferences and symposiums at home and abroad on mass spectrometry and protein analysis. In 2016, she received grant from student mobility program, Erasmus+ to do research about mass spectrometry imaging in Technische Universität Wien in Vienna, Austria. For the questions and remarks, contact e-mails are melike.dinc@hotmail.com and birmelikedinc@gmail.com