DIRECT DETERMINATION OF SURFACE PROTEINS OF LEISHMANIA PARASITE by PROTEOMIC APPROACH

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

MASTER OF SCIENCE

in Biotechnology

by
Ayşe Necla AKILLI

March 2023
İZMİR

ACKNOWLEDGEMENTS

I thank people who supported, guided, and encouraged me during my master's education.

First, I would like to thank Prof. Dr. Talat YALÇIN, who has always opened my horizons with his knowledge and experience and solved my questions and problems and I feel fortunate to be his student. Words cannot express my gratitude to Prof. Dr.Talat YALÇIN.

I want to express my deepest gratitude to Assoc. Dr. Alper ARSLANOĞLU for always spending time with me, answering my questions, and supporting me.

I want to thank my committee members Prof. Dr. Figen ZİHNİOĞLU and Prof. Dr. Çağlar KARAKAYA, for their suggestions and contributions.

I am deeply indebted to Dr. Melike DİNÇ and Dr. Çağdaş TAŞOĞLU from the National Mass Spectrometer Application and Research Center, who shared their knowledge and experience and guided me at every stage of my work.

I would like to thank Prof.Dr. Ahmet ÖZBİLGİN and his co-workers contributed to the isolation of *L.Tropica, L.Major, L.Infantum*, and *L.Donovani* that I used in my work. I also thank the parasites bank of Celal Bayar University for the Leishmania samples.

Special thanks to Merve BEYAZ, with whom I enjoy sharing the same laboratory ,who encourages me when I think I cannot do it, and who guides me on what to do.

I am very grateful to my dear friends Selin ÇOKDINLEYEN, Ceren H. BOZMAOĞLU, Irem AYERMOLALI, Aybike SARAÇ and Yasemin ŞAFAK for their help and moral support.

Finally, I would like to sincerely thank my dear father, my dear mother and my dear sister who were always there for me to complete my thesis.

ABSTRACT

DIRECT DETERMINATION of SURFACE PROTEINS of LEISHMANIA PARASITE by PROTEOMIC APPROACH

Leishmaniasis is a neglected tropical disease caused by the Leishmania parasite, primarily seen in developing and underdeveloped countries. Due to immigration to our country, the effective population of the disease has increased recently. The lesion, which has visceral and cutaneous forms, can be lethal when it acts on internal organs. Surface proteins are the most crucial part of the parasite and host interaction. The parasite attaches to the host cell through surface proteins, enters the cell, multiplies, suppresses the immune system, and allows many other biological functions. It is the most critical research part in vaccine and biomarker discovery. Generally, cell surface biotinylation and cationic colloidal silica beads with which the surface is coated are used to analyze surface proteins. These methods break down the cell and are more open to contaminant proteins from the cytoplasm and nucleus. In addition, since the surface proteins are rich in hydrophobic amino acids, they are difficult to dissolve in polar solvents.

The shaving method, uses a faster and minimal experimental workflow to cut only cell surface proteins without lysing the cell, has been tried in four Leishmania species (*L.Tropica, L.Infantum, L.Major, L.Donovani)*. The shaving method aims to digest cell surface proteins by treating the cell surface with a proteolytic enzyme for a short time. Thus, it is expected to contain fewer contaminants and unwanted proteins. As a result of shaving method analysis with the Fusion Orbitrap Mass Spectrometer, the rate of surface protein defined in 4 different species was 9.34% in *L.Tropica*, 7.55% in *L.Major*, 7.9% in *L.Infantum*, 7.52% in *L.Donovani*. Consistent with the literature and candidates for biomarker ISCL, KMP-11, Leishmanolysin, PSA-2, ABC transporter, and lanosterol 14α demethylase, proteins were identified. Familiar and different proteins were tabulated.

ÖZET

LEISHMANIA PARAZİT YÜZEY PROTEİNLERİNİN PROTEOMİK YAKLAŞIMLA DİREK TAYİNİ

Leşmanyöz, Leishmania parazitinin sebep olduğu, dünyada daha çok gelişmekte olan ve az gelişmiş ülkelerde görülen, ihmal edilmiş tropikal bir hastalıktır. Ülkemize yapılan göçler nedeniyle son zamanlarda hastalığın etki populasyonu artmıştır. Viseral ve kutanöz olarak iki formu bulunan leşmanyöz iç organlarda etki gösterdiğinde öldürücü olabilmektedir. Parazit ile konak etkileşiminde en önemli kısım yüzey proteinleridir. Parazit, konak hücreye yüzey proteinleri aracılığıyla bağlanır, hücre içine girer, çoğalır, immun sistemini baskılar ve daha birçok biyokimyasal fonksiyonun gerçekleşmesine izin verir. Bu nedenle aşı ve biyomarker keşiflerinin en önemli araştırma kısmıdır. Genellikle yüzey proteinlerinin analizinde hücre yüzeyi biyotinleme, yüzeyin kaplandığı katyonik colloidal silika boncukları gibi yöntemler kullanılmaktadır. Bu yöntemler hücreyi parçaladıkları için sitoplazma ve çekirdekten gelen kontaminant proteinlere daha açık olurlar. Ayrıca yüzey proteinleri hidrofobik aminoasitlerce zengin olduğundan polar çözücülerde çözünmesi zor olmaktadır. Bu çalışmada yaygın olarak kullanılan hücre yüzey proteini analiz yöntemleri yerine, dört leishmania türünde (*L.Tropica, L.Major, L.Infantum, L.Donovani*) hücreyi parçalamadan ilk olarak hücre yüzeyindeki proteinlerini kesmeyi amaçlayan daha hızlı ve minimal deneysel bir iş akışı olan tıraşlama yöntemi denenmiştir. Tıraş yöntemi, hücre yüzeyini kısa bir süre proteolitik bir enzimle işleyerek plazma zarı proteinlerini kesmeyi amaçlar. Böylece daha az kontaminant ve istenmeyen protein (analizlenmesi hedeflenen yüzey proteinleri dışındaki proteinler) içermesi beklenir. Fusion Orbitrap Mass Spectrometer ile analizlenen shaving metodunun sonucunda *L.Tropica, L.Infantum, L.Major, L.Donovani* den elde edilen örneklerden 4 farklı türde tanımlanan yüzey protein oranı sırasıyla *L.Tropicada* %9.34, *L.Major* de %7.55, *L. Infantum* da %7.9, *L.Donovani* de % 7.52 dir. Literatür ile uyumlu biyomerker adayı ISCL, KMP-11, leishmanolizin, PSA-2, lipoprotein, ABC transporter, lanosterol 14α demetilaz proteinleri tanımlandı. Ortak ve farklı proteinler tablolaştırıldı.

TABLE OF CONTENTS

LIST OF FIGURES

Figure **Page**

Figure **Page**

LIST OF TABLES

Table **Page**

CHAPTER 1

LEISHMANIA and LEISHMANIASIS

1.1. Leishmaniasis

Parasites are usually transferred to the host cell utilizing a vector. For example, they can be transferred from animals such as insects, mosquitoes, fleas, and ticks to the host. Leishmania spp is also transmitted by the bite of a female sandfly (Phlebotomine) to humans a vector-borne disease.[1] Leishmaniasis is the seventh most common tropical disease in the world. Despite the increasing lethal effect, the importance of researching diseases such as cancer and AIDS is still not shown for leishmania.[2] Leishmaniasis is in World Human Organisation (WHO) Neglected Tropical Disease (NTD).[3] That is, the disease is common in developing and under-developed countries. It spread more quickly in sub-tropical and tropical climates, where farming is common. Amphotericin B, pentavalent antimonials, and pentamidine are generally used in the treatment.[4-5]

1.2. Life Cycle

Leishmania has two form: Amastigote and promastigote. The promastigote contains flagella (the posterior) found in sandflies. Thanks to flagella, it can provide movement. The structure in a mammal's flagellar form is amastigote. The change in environmental conditions triggers this change. The amastigote forms an increase in temperature and a decrease in pH. After the sandfly bites the host (human), leishmania promastigotes are taken up by the host's phagocytic cells (macrophages, dendritic cells). Phagocytic cells protect the human body against foreign organisms such as bacteria, parasites, etc. This process occurs when leishmania is understood to be a pathogen and binds to the receptors on the macrophage surface. Thus, the phagocytosis process is initiated by the macrophages. Phagolysosomes contain lysosomal enzymes and

proteases.[6] Leishmania disrupts these mechanisms and multiplies by infecting the macrophage. In this case, leishmania transforms from the promastigote form with the flagellum. Then it allows it to act to protect itself from the host's immune system. The macrophages, from noticing it to the amastigote form without flagellum but with a greater replication ability. The environment of the parasitophorous vacuole lumen is acidic.[7] Therefore, the environment must be acidic for leishmania to stay in the host environment provide food and increase by binary fission.[8] Thus, the environment of amastigote is more acidic than promastigote. However, more than these adaptations are needed to stay in the host cell for a long time. As in many parasites, it is known that in leishmania, proteins are secreted to strengthen attachment to the host and facilitate more prolonged survival. Such as peroxiredoxin, activated C kinase proteins are found at the surface of the leishmania parasite and facilitate binding with the macrophage of the host.[9] Infected macrophages lyse, releasing amastigotes, and causing other macrophages to become infected. Bloodmeal is transferred to another sandfly and, amastigotes taken by other sandflies go to the midgut with increasing pH and decreasing temperature in the environment and transform into procyclic promastigotes.[10]

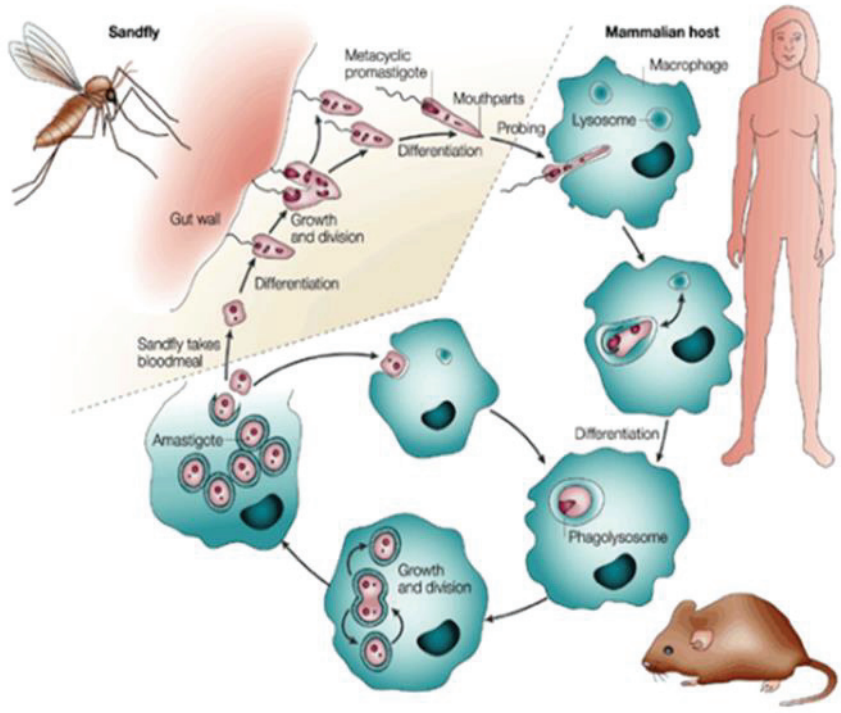


Figure 1.1. The Leishmania Life Cycle [10]

1.3. Type of Leishmaniasis

There are three types of leishmaniasis, name as visceral, cutaneous, and mucocutaneous. Cutaneous leishmaniasis is the most common form and usually causes skin lesions (caused by *L.Tropica, L.Major*). If the scars it creates do not fully heal, the patient can live with scars during their life. For this reason, people living in villages in ancient times tried to make natural vaccination by taking a swab from patients with a previous history of leishmania and applying it to their skin. Thus, they aimed to develop antibodies to the disease. It is thought that there are between 600.000 and 1 million new cases each year. Visceral leishmaniasis is 95% fatal if left untreated and targets the internal organs that cause a chronic disturbance.[11] It affects areas such as the liver, spleen, and bone marrow, and L.donovoni and L.infantum causes them. The Symptoms include bouts of fever, weight loss, abnormal enlargement of the spleen and liver, and anemia. rK39 serological test is currently used for the diagnosis of visceral leishmaniasis. According to the World Human Organization, 50,000 and 90,000 people are diagnosed with visceral leishmaniasis annually. If the parasite, which is expected to affect the internal organs, shows complications on the skin, it is known as Post-kala-azar dermal leishmaniasis (PKDL). Mucocutaneous leishmaniasis targets the nasal, mouth, and pharyngeal mucous membranes. It is mainly seen in Peru and Brazil.[11]

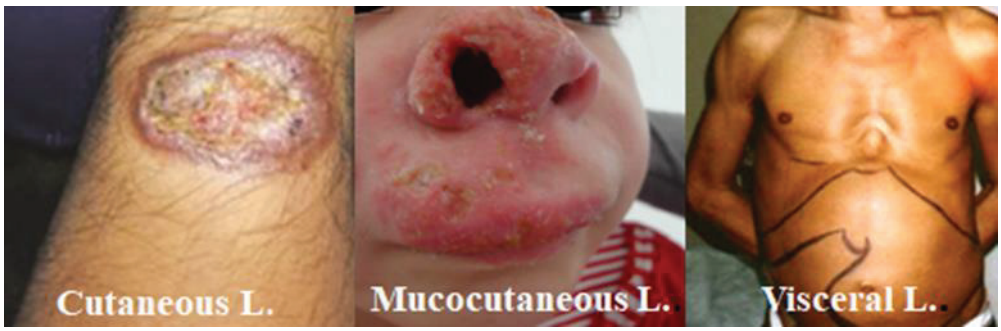


Figure 1.2. Type of Leishmaniasis [12-13-14]

1.4. Distribution of Leishmaniasis Around the World

It is mainly seen in rural areas. Leishmaniasis is found in 88 countries. (South and Central America, Africa, Asia, and Southern Europe) Visceral Leishmaniasis (VL) is generally found in Brazil, East Africa, and India. Cutaneous leishmaniasis(CL) cases are seen in the Middle East and Central Asia. 85% of the cases in 2022 occurred in Peru, Brazil, Iraq, Libya, Pakistan, Syria, Tunisia, and Afghanistan.[15]

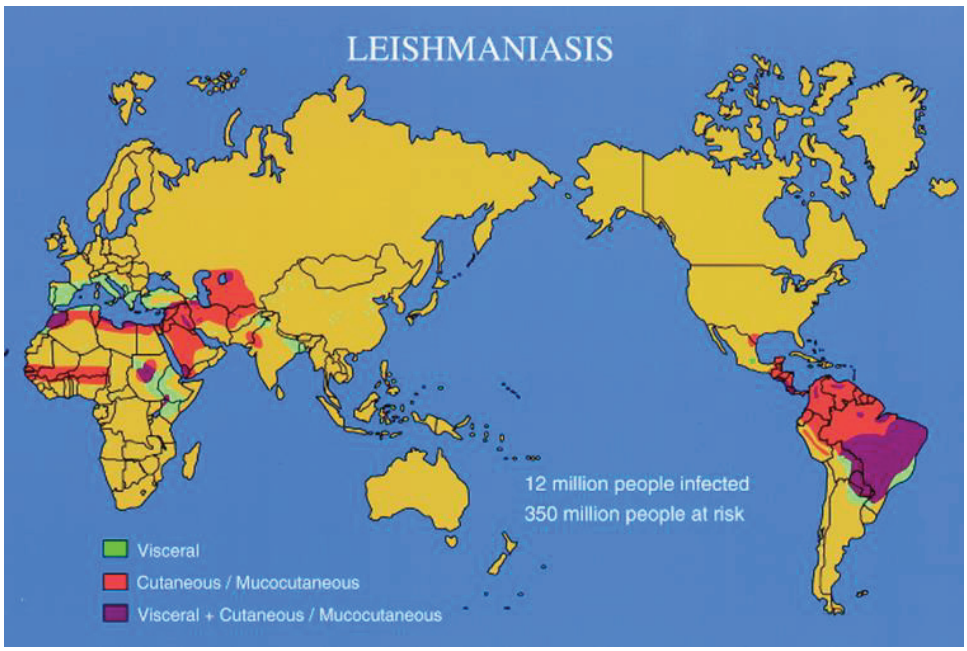


Figure 1.3. The Distribution of Leishmania in The World[16]

CHAPTER 2

MASS SPECTROMETRY and PROTEOMICS

2.1. Introduction of Mass Spectrometry

Mass Spectrometry is an analytical technique and, its general principle separates each ion according to its mass-to-charge ratio (m/z) by converting the sample into gas and ionizing it. It is an excellent system for analyzing samples whose contents are unknown. The use of mass spectrometry in many fields has made it popular for many departments which started with J.J Thomson's measures with the m/e cathode ray tube[17]. Some provide information about the century of historical buildings from carbon numbers, analysis of biomolecules, drug studies, determination of isotopic abundance, organic synthesis, and diagnosis of many diseases. Mass spectrometry is fast, reliable, and sensitive. The importance of mass spectrometry is increasing thanks to new devices with more sensitive and high resolution (e.g. orbitrap).

Mass spectrometry has many essential parts; ionization source, analyzer, detector, and high vacuum pump system. First of all, it is ensured that the sample becomes gaseous and ionized using ionization sources. Afterward, mass analyzers accelerate these ions in the electric or magnetic field and separate them according to their mass-to-charge ratios. Particles that do not become charged are deflected. Finally, charge particles are detected by a suitable detector.[18] The relative intensity of each sample can be found according to its mass-to-charge ratio. Mass spectrometry finds the exact mass of the molecules. These systems provide high vacuum and low pressure (10^{-5} to 10^{-7} torr). The reason for working under low pressure and high vacuum is to prevent unwanted collisions and to protect the low mean free path between charged particles. The mean free path between two molecules at atmospheric pressure is 6.5×10^{-8} meters. That is why impossible for two molecules not to collide. In mass spectrometry, the ions should not collide with each other and return to their neutral state. Thus, it is operated at low pressures such as 10^{-8} or 10^{-11} torr, in an ultra-high vacuum.

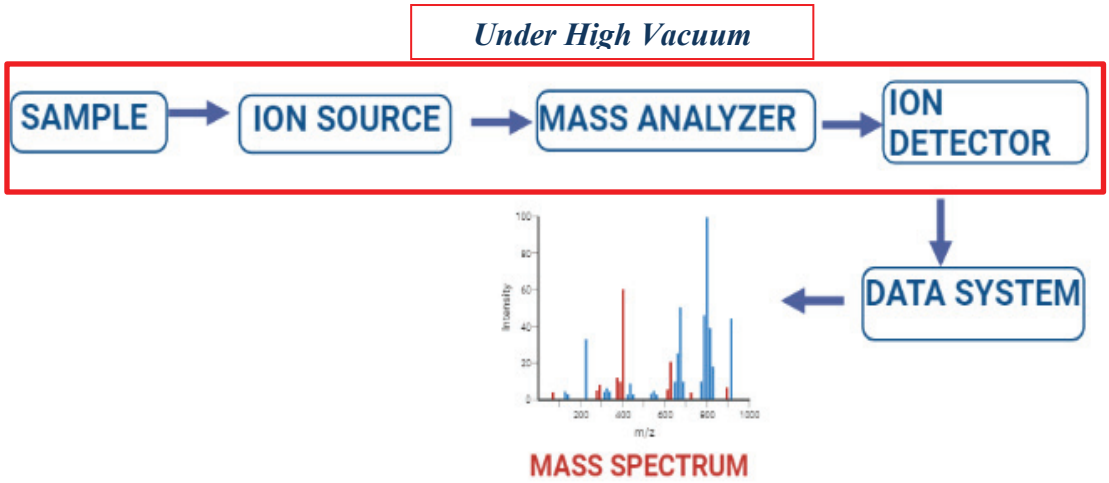


Figure 2.1. The Basic Component of Mass Spectrometry

2.2. Soft Ionization Techniques

Two 'soft' ionization methods have been developed to analyze samples between 400 and 3000 daltons, which are used to analyze biomolecules. These are Electrospray and Matrix Assisted Laser Desorption/Ionization. Usually, proteins are analyzed in positive ion mode. The presence of charged amino acids (Lys, His, Arg) in the peptide bond determines the maximum charge carried by the molecule.

2.2.1. Electrospray Ionization (ESI)

Electrospray from spray methods; It is the method in which the solution containing the analyte is sprayed into the mass spectrometry under atmospheric pressure. Electrospray is ideal for nonvolatile liquid samples such as protein. With nebulization, tiny droplets are created that are sprayed from atmospheric pressure to the mass spectrometer. A potential difference is usually applied in this spraying. This potential difference is vital for the movement of the solution from the ESI tip (microcapillary inlet) and into the mass spectrometry inlet. It evaporates with heated

nitrogen at this time, and charge droplets are formed. It is sprayed into the system as charge droplets from here. The solvent in the droplet heated with nitrogen gas evaporates, and collision with neutral gaseous (N2) and decrease in surface tension destabilize the droplets.

A coulomb explosion occurs when the droplets can no longer withstand the coulomb thrust against the surface tension. This process takes place repeatedly until the time when the charge particles are released. Solvents such as water, which reduce surface tension and make the analysis more effective, are used in the mobile phase to provide low surface tension quickly and to allow the droplets to burst and release the charge particles as soon as possible. Charged particles head toward the Mass analyzer. Electrospray and triple quadrupole or ion trap are the most commonly used mass analyzers. Since the Electrospray ionization method is a continuous system, it is necessary to use ion gates to be combined with pulse system mass analyzers such as time of flight.[19-20]

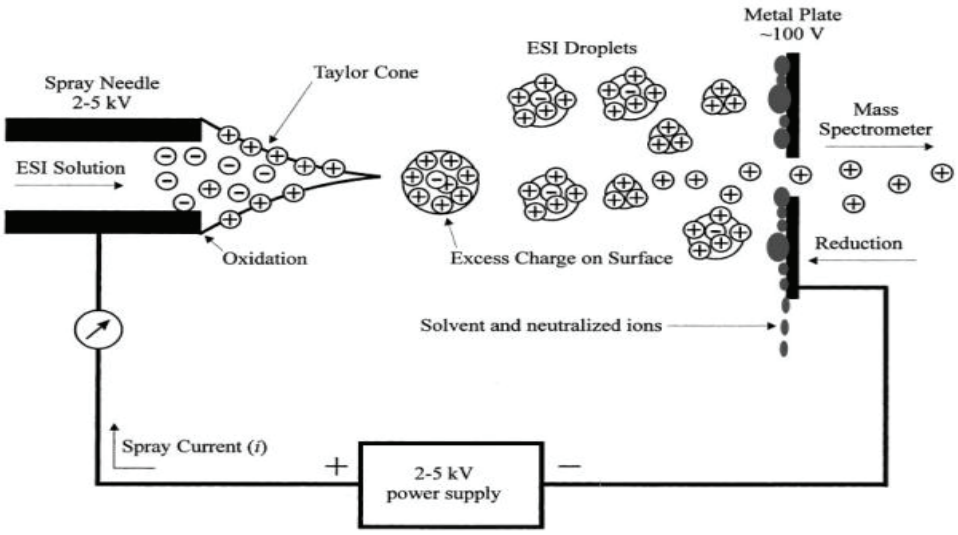


Figure 2.2. Principles of Electrospray (ESI)[21]

2.2.2. Matrix-Assisted Laser Desorption/Ionization (MALDI)

Matrix-Assisted Laser Desorption/Ionization (MALDI) allows the analysis of labile macromolecules and nonvolatile molecules.

The sample is trapped inside the matrix. Matrix is a weak organic acid. As an example of the matrices used, Sinapinic acid (3, 5-dimethoxy-4-hydroxycinnamic acid) (SA), 2,-(4-hydroxy-phenyl azo) - Benzoic acid (HABA), α-Cyano-4-hydroxycinnamic acid (CHCA) may be given. The amount of matrix is more than the sample (1: 1000 sample/matrix amount ratio). The matrix-sample mixture is placed on the MALDI target. The molecules of the analyte cocrystallize with those of the matrix when the solvent evaporates. The matrix is exposed to 337 nm nitrogen laser or laser light with neodymium/yttrium-aluminum-garnet (Nd-YAG) at 355 nm. The laser beam interacts only with the matrix. Thus, the sample can be ionized without being degraded. The interaction of the matrix with the laser beam causes the sample to pass into the gas phase and ionize. Then, the charged particles in the gas phase are accelerated and directed to the Mass analyzer[19].

The difference from fast atom bombardment (FAB) is that a solid matrix is used in MALDI. The second difference is that the sample collides with an atom or ion beam in FAB, while the laser beam is used in MALDI.

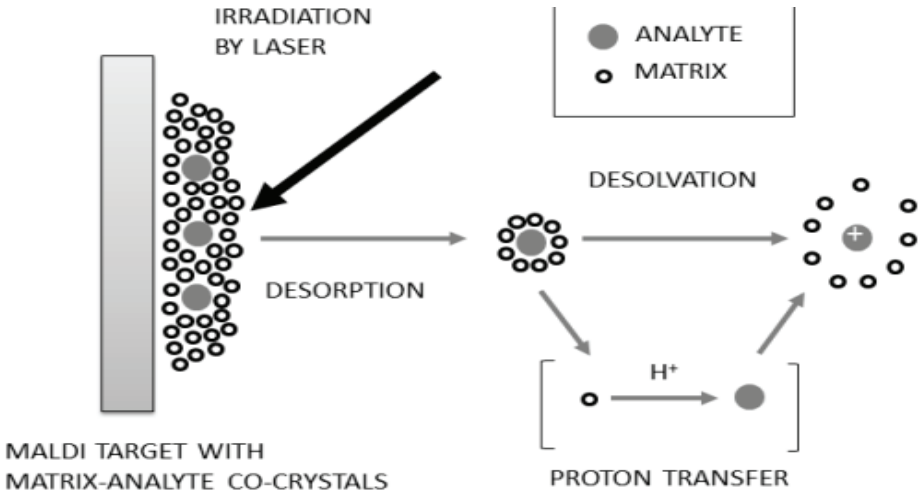


Figure 2.3. Principles of MALDI [22]

2.3. Mass Analyser

After the sample becomes gaseous and ionized, it is separated by each ion analyzer according to its mass-to-charge ratio. Generally, tandem MS is used in proteomics studies. This method, in which more than one mass analyzer is combined, provides a better separation. Combining multiple mass analyzers, such as Quadrupole-Time of Flight (Q-TOF) and Triple Quadrupole (QQQ), provides better separation. Since Electrospray is continuous, that is, it is sprayed continuously, it combines with Triple Quadrupole or Ion Trap suitable for continuous analysis. On the other hand, MALDI-TOF is a pulse system.[23]

2.3.1. Time of Flight (TOF)

It is a method that analyzes the mass-to-charge ratio according to the flight time of the ions to the detector. When ions come to the mass analyzer with the same energy with the effect of the electrostatic field and drift in the free-flight region, they reach the detector at different speeds due to distinct mass-to-charge ratios. The lighter ions hit the detector faster.[19]

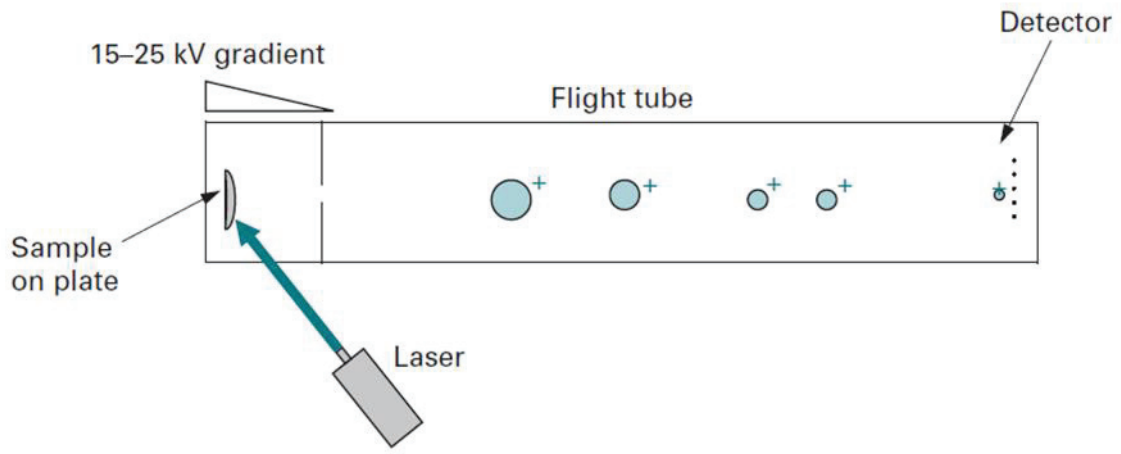


Figure 2.4. Time of flight (TOF) Mechanism[19]

2.3.2. Quadrupole

Quadrupole consists of 4 rods parallel to each other in cylindrical form. Radiofrequency (RF) and direct current (DC) are applied to each rod, applying a continuously varying electrical current along the length of the analyzer. The electric current is controlled throughout the analysis. Thus, only charged particles with a particular mass-to-charge ratio reach the detector, while those below the desired m/z ratio cannot proceed straight and are deflected. If RF and DC are applied proportionally, the charged particles or ions expected to be analyzed reach the detector. It has been observed that in the case of only RF or only DC applications, unwanted ones reach the detector. The same principle applies to the ion trap and ion cyclotron resonance analyzers. Quadrupoles can analyze samples with an m/z ratio of up to 3000.[19]

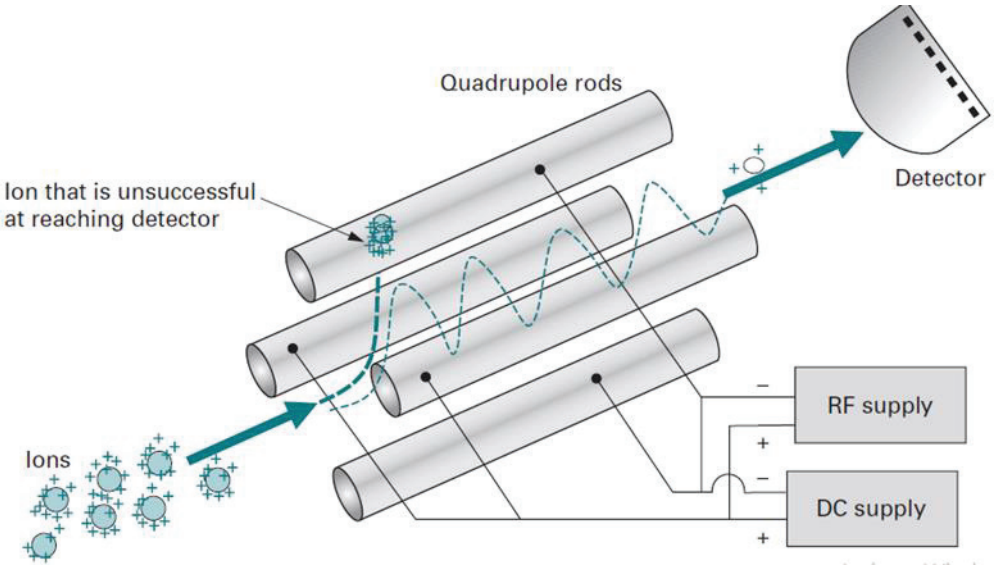


Figure 2.5. Quadrupole Mechanism[19]

2.3.3. Ion Trap

Wolfgang Paul invented the ion trap with an exact quadrupole mechanism in the 1980s. Ion trap aims to keep or trap ions and stores ions under magnetic and electric fields. After all, ions are imprisoned in the ion trap, and their mass separates them to charge ratios and reaches the detector with the change of the RF potential. Ions collide with helium gas in the ion trap. Thus, the ion's kinetic energy decreases. Finally, they can reach the detector. Reducing the pressure and working under ultra-high vacuum prevents the ions, namely the charged particles to be analyzed, from colliding with each other and becoming neutral.

It is aimed to analyze all trapped ions for the MS^n process. The resulting parent ions are collided with a low helium pressure, resulting in collision-induced dissociation of the ions. The fragment spectrum is defined from the obtained fragment ions. These fragment ions can be cleaved repeatedly if further structural information is desired. It provides accurate mass-to-charge determination of up to 100,000 daltons in tandem with MS analysis with an ion trap. Quadrupole ion trap (QIT) or 3D IT, consisting of a ring electrode and two endcap electrodes, are the standard ion trap analyzers.[19-24]

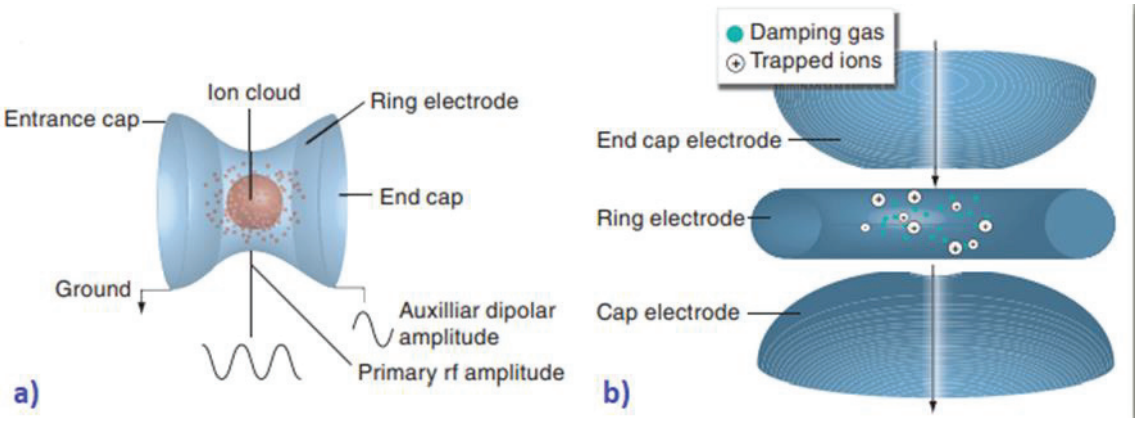


Figure 2.6. Ion-Trap Mechanism[25]

2.3.4. Orbitrap

Aleksandr MAKAROV invented the orbitrap principle; The central spindle is based on the trapping of ions at the electrode, and the mass-to-charge ratio is generated using the Fourier transform of the frequency of the oscillations of the orbitally trapped ions. For electrostatic field adaptation from TOF, ion trapping within the electrodes from ion trap analyzers, and current image measurement from FT-ICR, the Orbitrap mass analyzer has inherited important characteristics from earlier analyzers. In Orbitrap-based tandem MS systems, high resolution, high speed, mass accuracy, and pre-treatment stage before LC-MS will reduce the stage. LTQ-Orbitrap, Q-Exactive, and Orbitrap fusion mass spectrometry are used in tandem MS systems.

Ions are held captive in the orbitrap oscillating harmonically along an electric field's axis. Their m/z values are calculated using Fourier transforms to derive the mass spectrum from the ions' frequencies, which are detected without causing damage. The device has a wide m/z range (>6000), high mass resolution (up to 1.500.0 00), high mass accuracy (2-5 p.p.m.), and sub-femtomolar sensitivity.[26-27]

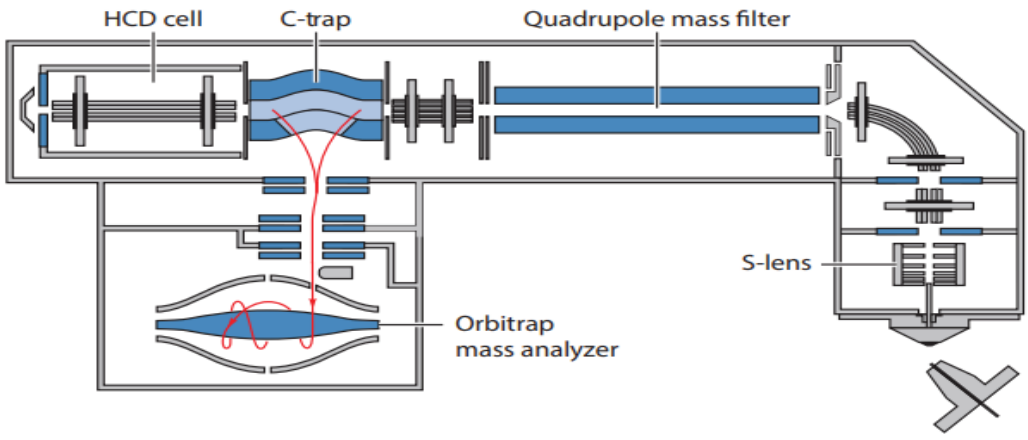


Figure 2.7. Q-Exactive Mass Spectrometer [27]

2.4. Ion Detectors

The detector detects the electrical signal of the ions that hit the detector. A mass spectrometry detector should be fast, accurate, long-lasting, noiseless, high amplitude, and have a dynamic range. Electron multipliers and multichannel plate detectors are generally used in proteomics studies.

2.4.1. Electron Multiplier (EM)

In Electron Multiplier, the principle is to detect the secondary electrons formed. In the positive ion mode, it is based on the secondary electron detachment from the surface when ions or charged particles hit the surface of the cathode and the detection of these detached electrons under a different voltage. For the secondary electron to be detached from the surface, the energy of the ion (charged particle) striking the surface must be sufficient to detach the electron. The electrodes accelerate and strike another dynode, reaching the anode due to differential potential voltages between the dynodes. Impacting this surface and breaking electrons is equal to the number of dynodes. Each impact breaks 2n electrodes, and when it gets closer to the last dynodes, it creates a high current. Electron Multiplier is generally preferred after mass analyzers such as Ion Traps and Quadrupoles.

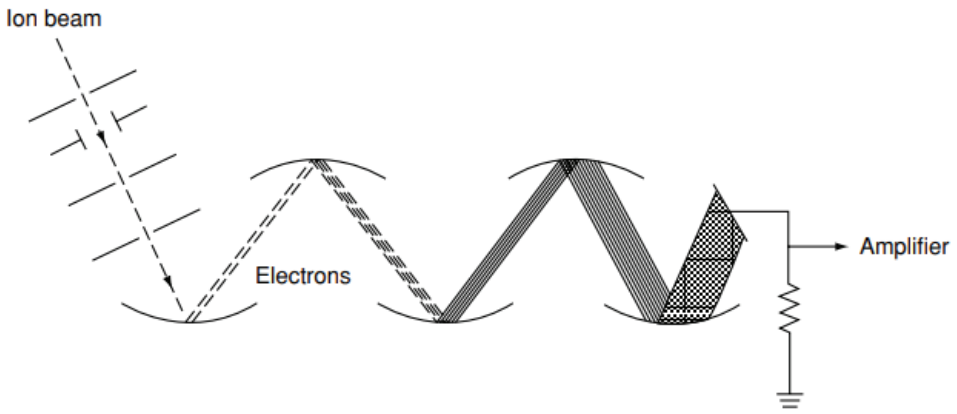


Figure 2.8. Electron Multiplier Mechanism[28]

2.4.2. Multichannel Plate

Here there is a principle similar to the principle in the electron multiplier. However, its ion capture capacity and electron removal rate are higher. The secondary electrode, formed by the impact of the ion on the cathode surface, moves in the electric field due to the applied potential difference and causes more electron detachment from the surface with each impact. Multichannel Plate is preferred for MALDI-TOF analysis.[29]

2.5. Proteomics

Marc WILKINS was the first person who used the word proteomics in 1994. Proteomics got its name from the combination of protein and genomics.[30] The primary purpose of the human genome project, which started in 1990 and was completed in 2003, is to map the human genome. The primary sequence of the genomic structure, which contains approximately 30,000 genes, has been resolved. This project sheds light on the treatment of many diseases. However, it is known that events in the human body are not only dependent on genome sequencing. For example, one gene is not equivalent to one protein. One of the shortcomings of the human genome project is not providing the interaction between genes and proteins. The best example is that although the caterpillar and the butterfly have the same genome, their functions differ because they contain different proteins. There are nearly 2 million proteins against 30,000 known genes. [31]Factors include;

- Any additions that may occur in the gene sequence,
- Post-translational modifications
  (such as methylation and phosphorylation),
- Differentiation of biological activity, and
- Protein degradation.

These changes show that the proteome is dynamic and the genome is static. Therefore, proteomic studies are essential for a complete understanding of biochemical events. For instance, Proteins change with the effect of diseases. The rate of methylation is seen more in cancer patients. In this way, proteomic analyses are essential for

convenience in diagnosing diseases and drug discovery. Thus, proteomics research; aims to describe the structure, function, and expression of all proteins-structural proteomics; models protein structures. Nuclear magnetic resonance spectrometry and X-ray crystallography are used to find the 3D structure of proteins. Functional proteomics focuses on the functions and interactions of proteins. It investigates how the interactions of protein-protein, DNA - protein, protein-lipid, protein - drug identify post-translational modification (PTM) and their effect on the protein's function. Expression proteomic studies focus on protein identification, expression, and differential expression depending on any biochemical or physical factor. It can identify disease-specific proteins. For example, it is known that after the parasite enters the host cell, it causes the expression of some different proteins in the host or increases the levels of some proteins. These studies are defined by 2-DE and mass spectrometry.[32-33-34]

Thanks to proteomic studies, protein sequences of many diseases can be found, while proteomic analyzes of parasites and their hosts have become more manageable. Furthermore, identify different biomarkers. Studies on proteomics continue in the fields of medicine, biochemistry, and chemistry

2.5.1. Proteomics Approach

There are two types of proteomic approaches, Top Down and Bottom Up. *Bottom-Up proteomics* is a widely used technique based on the characterization of peptides formed by the proteolysis of proteins. The characterization of peptides formed by proteolytic degradation of the protein mixture is called Shotgun proteomics. In shotgun proteomics, the protein mixture broken down by protease enzymes is fractionated and analyzed by LC/MS. Trypsin, one of the proteolytic enzymes, is often used in shotgun proteomics. Trypsin cleaves the carboxyl end of arginine and lysine. These amino acids are positively charged because they are basic amino acids. A positive ion at the carboxy-terminal end of trypsin treatment facilitates the peptide sequence's ionization in mass spectrometry. Analyzed peptides are identified by comparison with data from theoretical mass spectrometry. This technique; While it is advantageous for the identification of a large number of proteins; It has low-resolution protein identification capacity for PTM and alternative splicing studies. Another method is Top Down Proteomics, to directly analyze the protein (intact protein) without proteolytic

degradation and search the data. Generally separate with chromatically and ionize and analyze with MS. The disadvantage of this method is that fractionation, ionization, and protein characterization are more complicated than shotgun proteomics.[35-36]

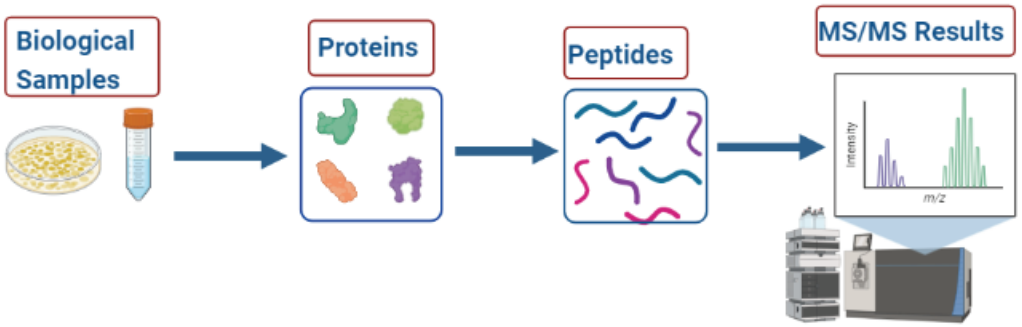


Figure 2.9. Shotgun Proteomics Workflow[37]

2.5.1.1. Mass Spectrometry-Based Proteomics

The invention of soft ionization techniques, such as Electrospray and MALDI, opens the possibility of nonvolatile molecules analysis via Mass Spectrometry. Today, ESI and MALDI have very important for proteomic studies. To summarize the steps in general MS-based- proteomics studies, the protein is extracted and purified, and DNA, cell components, and anything that may contaminate during the experiment are purified. Isolated proteins from cells or tissues are first separated by gel electrophoresis or HPLC, interfering with a proteolytic enzyme. Obtained mass Spectrometry analyzes proteolytic peptides. The results obtained from mass spectrometry are compared with the existing database using software such as Mascot, Sequest, and proteins are identified.

2.5.2. Separation Techniques before Mass Spectrometry

Samples can be prepared for mass spectrometry using gel-based, gel-free, and multidimensional separation methods.

2.5.2.1. Gel-Based

2D - Gel electrophoresis (2D-PAGE) separates proteins based on their isoelectric points (IEF) and molecular weights. Gel electrophoresis uses chaotropic and detergent to dissolve and precipitate the protein. Chaotropes such as urea weaken hydrophobic interactions and hydrogen bonds between proteins. However, SDS is a detergent that binds strongly to hydrophobic amino acids, that could prevent the interaction between amino acids and ensures denaturing of the protein. After 2D gel electrophoresis, protein bands can be detected by coomassie or silver staining. The results' bands of interest (containing unknown protein) are excised and cut using a proteolytic enzyme. Trypsin enzyme is frequently used because it is more compatible with mass spectrometry and cuts at specific points in the protein.[36]

Some of the disadvantages of this method are;

- Gel preparation, staining, and destaining steps are time-consuming.
- Proteins with less solubility are more challenging to see in the gel.
- For proteins with a pH range of less than 4 and more than 9, proteins with a molecular weight of fewer than 10 kDa and more than 200 kDa cannot be detected.

2.5.2.2. Gel-Free Based

The interaction of the sample between the mobile phase and the stationary phase is vital in all chromatographic techniques. While the mobile phase enables the sample to move along the column, the stationary phase may contain different properties depending on the separation desired. The sample may adhere to the column or elute by interacting

with the stationary phase. Chromatograms are created according to the time it takes for the sample to be injected and detected.[40]

HPLC is frequently used with ESI-QqQ mass spectrometry systems. The samples separated in HPLC are transmitted directly into the electrospray to form droplets. HPLC has been an excellent approach for analyzing rare peptides, removing salts that may cause interference, and reducing the ion suppression effect.

The high-performance liquid chromatography technique (HPLC), invented in the early 1970s, is essential for proteomics studies due to its high resolution, high sensitivity, reproducibility, and small sample size sufficient for analysis. The difference from gas chromatography is that it is compatible with nonvolatile compounds. HPLC includes various separation methods such as hydrophobic, hydrophilic, size, affinity, and reversed-phase (RP-HPLC). For example, in the reverse phase; On the other hand, long hydrocarbon-structured nonpolar compounds such as C18 and C8 are trapped in the stationary phase, that is, the column. Ultrapure water and organic solvent are preferred as mobile phase. Acetonitrile and trifluoroacetic acid (TFA) is preferred in organic solvents. When the sample is given to HPLC, the nonpolar components in the sample will be adsorbed in the column while the other compounds will be eluted. Then, the sample retained in the column is eluted with organic solvents such as acetonitrile. Isocratic or gradient flow can be provided. Gradient flow (i.e. increasing acetonitrile ratio) is generally preferred for the complete separation of peptides and the best separation from impurities. Therefore, the compound with the longest retention time will be the most nonpolar.[40]

HPLC is frequently used with ESI-QqQ mass spectrometry systems. The samples separated in HPLC are transmitted directly into the Electrospray to form droplets. Combining HPLC with Electrospray; has been very good in terms of better analysis of rare peptides, removal of salts that may cause interference, and reduction of ion suppression effect.

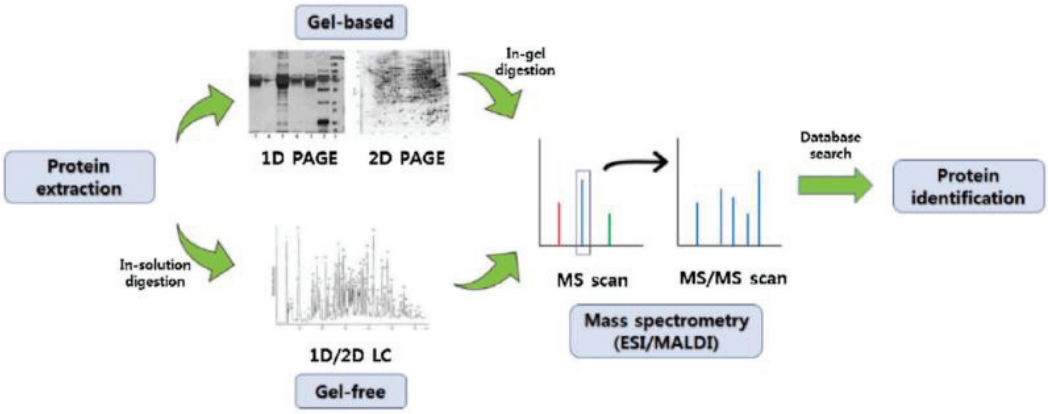


Figure 2.10. Gel-Free and Gel-Based Proteomic Workflow[41]

2.5.2.3. Multi-Dimensional Separation (MudPIT)

Due to the complex structure of proteins, more than one-dimensional HPLC is needed. A multi-dimensional system has been developed to achieve better and clearer separation. Thus, a distinction is made according to multiple separation options, such as using a strong cation exchange column and reverse phase. This system first separates the sample according to its isoelectric point with a strong cation exchange column (SCX). Then it separates the sample according to its non-polarity in a reverse phase column with apolar character. Thus, a better resolution is provided and, the peptides are enriched. There are separations made using two reverse-phase columns.[39]

2.5.3. Shotgun Proteomic Approach

In the general flow of Shotgun proteomics, proteins are first broken down with the help of a proteolytic enzyme (such as trypsin, proteinase K), followed by separating the peptides formed by liquid chromatography and then their analysis with tandem MS. The data obtained after tandem MS analyses are compared with the reference protein databases with the help of software and show which protein or proteins can match the data obtained. Examples of software used are Sequest, X! Tandem, and Mascot.

2D gel electrophoresis (or Multi-dimensional chromatographic separation systems) primarily aims to reduce the sample's complexity. Multi-dimensional chromatographic separation systems are preferred in Shotgun proteomics. Because 2D – PAGE has some problems, such as incomplete imaging of membrane proteins, each process makes the sample more complex. Sample losses may be experienced, and time loss in preparing gel electrophoresis solvents. Fractionation of samples before tandem MS provides protein enrichment, increased analytical dynamic range, and better peak resolution. Two different Multidimensional separations can be used in gel-free methods such as shotgun proteomics. These are on-line and off-line.

In on-line separation, peptides are directly analyzed by 2D- HPLC. It is a method in which more than one column is used together for chromatographic separation and, peptides are separated according to different properties. For example, in SCX-RP HPLC ESI-MS-MS systems, by using SCX and RP together, firstly, peptides are separated according to their isoelectric point by a strong cation exchange column (SCX column). At the same time, in the reverse phase (RP), polar compounds are aimed to be eluted first by attaching apolar peptides to the column.

Strong cation exchange (SCX) may be disadvantageous for tryptic peptides because peptides formed by cleavage with trypsin have 2+, 3+, or 4+ charges. Therefore, peptides can form groups and, a clear distinction cannot be made. Because of this, reverse phase - reverse phase HPLC (RP-RP-HPLC) produces a more precise and non-overlapping peak. Although RP-RP HPLC can distinguish better, it prevents orthogonality formation. To mimic SCX-RP and provide orthogonality, two very different pH ranges, such as the 4-10 pH range, are required.

In offline separation, there is fractionation. Tryptic peptides are first separated by a high pH RP column and collected and fractionated on the microplate. The aim is to provide a better separation by combining peptide mixtures with two different properties. Thus, the analysis result in MS-MS is improved. It is then separated by a low pH RP column and analyzed with Electrospray - tandem MS systems.[24-41-42-43]

2.5.4. Peptide MS/MS Spectrum and Fragment Ions

In mass spectrometry, amino acid sequences of peptides can be found from ions obtained from tandem mass spectrometry by collision-induced dissociation (CID) of peptides in positive mode. Tandem mass spectrometry is generally used in proteomic studies. In these systems where more than one analyzer is combined (Tandem MS or MS/MS) with collision cell dissociation(CID). In the first mass analyzer, the ion of interest is selected and collides with inert gases (nitrogen) to fragment into daughter ions. The second mass analyzer helps to obtain the CID spectrum of the ion of interest as well as the amino acid sequence of the precursor ions.

The mobile proton model is the most common theory explaining these 'protonated peptides' fragmentation state. Fragmentation can occur from any region, and the charge can be seen at fragments N or C ends. Therefore, in the MS/MS spectra, ions at the N terminal are named b, while ions at the C terminal are named y.

In Figure 2.11. breaks starting from the N-terminal are denoted by b, and the cleavage of a single amino acid from the peptide is called b1. This nomenclature continues incrementally for each amino acid cleavage up to the C terminal.[44-45]

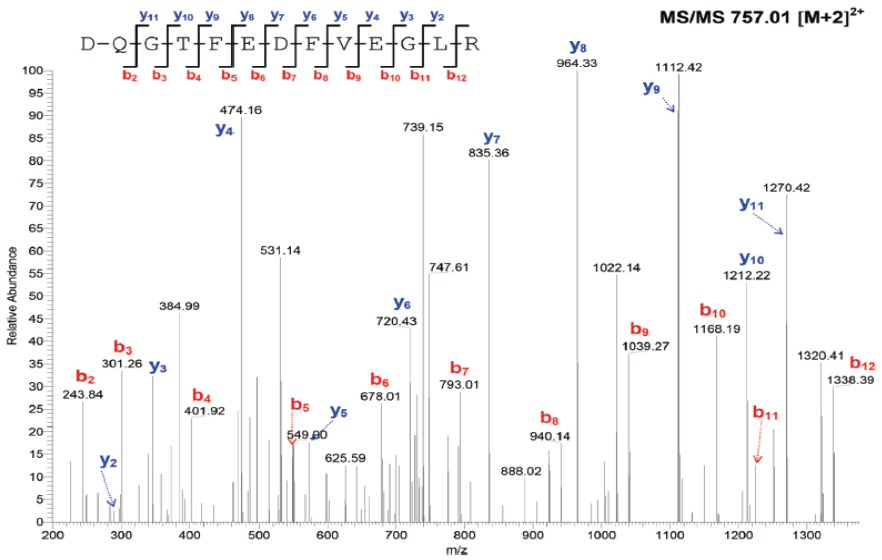


Figure 2.11. MS/MS spectrum[45]

2.5.5. Peptide Sequencing and Database Search

In biochemical analysis, Peptide sequencing identifies one amino acid at a time from the N-terminal end of the peptide. In this method, phenyl isothiocyanate (PITC) is added to the N-terminal end of the peptide. Specifically, for hydrolysis to occur, it breaks the nearest amide bond and forms a cyclic structure containing the amino acid called phenylthiohydantoin, which is a phenyl carbamoyl derivative. The obtained phenylthiohydantoin (which contains amino acid and phenyl carbamoyl derivative compound) is analyzed by LC. While Edman degradation is faster and more effective than the Sanger method, separating and sequencing individual amino acids poses a problem in sequencing large proteins. Interactions in the protein, such as hydrophobic interactions, are unpredictable, although they do not cause any damage to the protein.

Because of these problems, software was developed to search for potential proteins that overlap with m/z obtained from MS by mass fingerprinting (PMF) and MS/MS ion search rather than peptide determination by sequencing. For instance, automatic software such as Mascot, Sequest, and X! Tandem can be given. In these software, easy and fast analyzes are performed.[24-46-48]

Raw data from MS/MS and existing database information are loaded into this software. Softwares such as Sequest, and Mascot scan the raw data information against the existing database (such as uniparc databases). Some database addresses are given in the table 1.1.

Tablo 1.1. Several Protein Sequence Databases Addresses[46]

Protein Sequence Databases	Web Adress
Reference Sequence (RefSeq)[a]	http://www.ncbi.nlm.nih.gov/RefSeq/
UniProt - UniParc	http://www.uniprot.org/
International Protein Index (IPI)[c]	http://www.ebi.ac.uk/IPI/IPIhelp.html

Suppose trypsin, creatine, and MALDI/MS, which can create contaminants in the results, are performed in the software settings. In that case, if it is known in which interval the interference peaks may come from the matrix, and salts give peaks, this information can be entered to prevent the results from having contaminant values.[43]

The mascot settings shown in Figure 2.12. with which enzyme the degradation is made, peptide tolerance range, load range of peptides, the maximum desired missed cleavage range, and expected or possible modifications should be indicated.

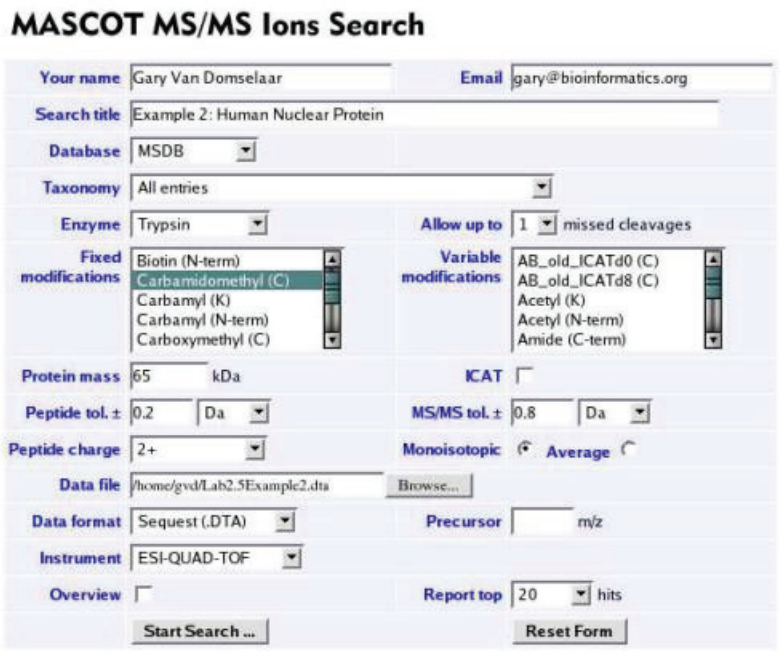


Figure 2.12. Mascot Interference Setting[47]

When the scan is completed, a sequence of proteins with the score is obtained. The importance of scoring indicates the match's success, and the error rate must be known. For this, the degree of misidentified (false) proteins can be found by searching the data in the fake database.[43]

2.5.6. Plasma Membrane Perspective and Surface Proteins

The plasma membrane (PM) contains lipids, carbohydrates, and proteins. The phospholipid bilayer, which is the basic structure of the cell membrane, is amphipathic. Structurally, the outer part is hydrophilic, interacting with the cytosol and the extracellular membrane, while the inner part is hydrophobic and consists of fatty acids. Among the essential components of the membrane are proteins. Plasma membrane (PM) protein is on the cell surface and can also be called surface protein. Cell and plasma membrane proteins surround the cell and define the boundaries between the cytoplasm and the extracellular. PM proteins, comprising 50% by weight of PM, include integral proteins embedded in the lipid bilayer, membrane-associated peripheral proteins, and proteins anchored to the lipid bilayer (such as GPI-anchor proteins). It is crucial for the cell due to its functions, such as protein-protein interaction, enzymatic activities, receptor, signal, and ion transport. *Peripheral proteins* are proteins associated with the phospholipid membrane. These proteins, more associated with the hydrophilic part, are attached to the membrane by ionic or hydrogen bonds. Certain conditions should be considered to extract the membrane proteins, such as pH increase and high salt concentrations can be applied to break the bonds. Integral proteins, conversely, are embedded in fatty acids in the phospholipid bilayer (type of integral protein; alpha-helix, β-barrel, etc.)[49]. These proteins, rich in hydrophobic amino acids, are more challenging to extract. Amphipathic detergents are preferred (such as SDS and Triton X-100). Finally, there are lipid-anchored membrane proteins.

Transmembrane proteins are a type of integral membrane; Plasma membrane proteins are proteins that stay in the hydrophilic part and bind signal molecules that cannot enter the lipid bilayer and act as signal transduction. The most critical example is G-proteins.[50]

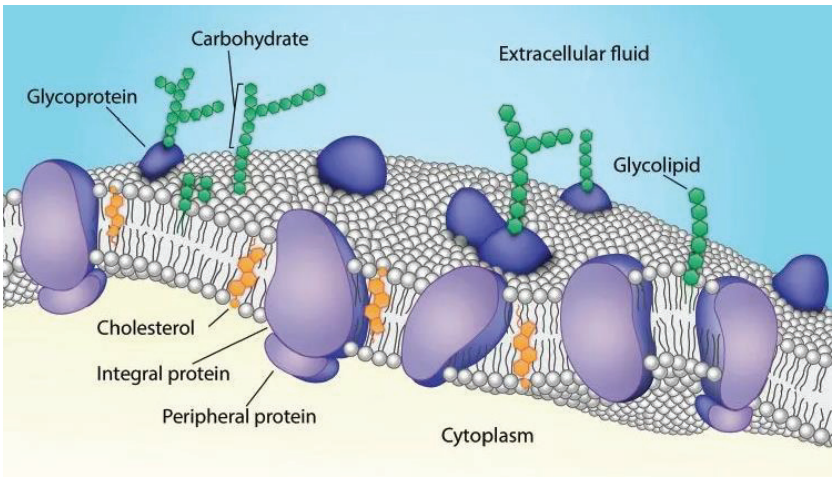


Figure 2.13. Plasma Membrane Component[51]

It is known that plasma membrane proteins move by rotational and lateral diffusion. The movement of plasma membrane proteins was first proven by forming a heterokaryon cell. According to this experiment, human and mouse plasma membrane proteins were artificially monitored. It is labeled with two antibodies, fluorescein, and rhodamine, for easy monitoring by electron microscopy. Thus, antibodies have two colors, green and red, which were created and easily followed. As a result of the experiment, It has been observed that human and mouse surface proteins act with each other.[50]

If the cell encounters any invading species(such as parasites or viruses), the expression levels of proteins in the plasma membrane may change. The effect of surface proteins greatly accelerates the diagnosis of parasites, viruses, or cancer cells. Post-translational modifications, such as glycosylation, generally occur on the cell surface. Sometimes are triggered by the immune system. The expression levels of some proteins on the cell surface increase significantly, and their detection is crucial for detecting the disease.

2.5.7. Plasma Membrane Proteins of Leishmania

The plasma membrane (PM) carries out important functions. Some of these are membrane processes, transports, cell-to-cell and protein-to-protein interactions, and the directing of environmental stimuli.

A few crucial aspects of the studies with cell surface proteins are that they shed light on how cells establish a connection and the pathways that the changes that occur with cell interactions create or trigger in the cell or organism. Surface proteins for pathogens play critical key roles in host-parasite interactions, such as adhesion to the host organ or tissue, invasion, increased toxicity, and inhibition of the immune system response. Leishmania metacyclic promastigotes cause suppression of interleukin and interferons. Amastigotes form proteins in mammalian macrophages, causing the phosphorylation of JAK-STAT pathways and hiding from the cell's immune system protein can bind strongly to the host membrane by maintaining the stability of lipophosphoglycan (LPG) and glycoprotein 63 (GP63). It is an essential determinant for the start of phagocytosis and survival in the host membrane. Other parasite surface proteins are peroxiredoxin, cytochrome C oxidase, and inositol phosphosphingolipid phospholipase C-like. Leishmania provides both survival and spread of the parasite by renewing itself in energy metabolism, protein interactions, and many biochemical processes throughout its life cycle. After the research, some important virulence surface proteins were found in leishmania. Some of these are proteinases, lipophosphoglycan (LPG), proteophosphoglycan (PPG), and kinetoplastid membrane protein-11 (KMP-11) proteins. Proteinases (such as cysteine protease) reduce the host's immune response. It hydrolyzes peptide bonds and disrupts the structure of proteins. It affects the host immune system pathways, such as JAK/STAT, by phosphorylating and dephosphorylating the immune system with serine-threonine phosphatases. It reduces NO production. Thus, Leishmania species increase the possibility of escaping from the immune system in the cell and continuing their lives. It is known that PPGs are an essential protein in infecting the host cell when the parasite is in the metacyclic promastigote state. Therefore, biomarker targeting is an important research topic in vaccine and drug development studies, as surface components and surface proteins are the first steps in host-parasite interactions.[52-53]

Cell-cell interactions and protein binding have been engaging in proteomics for many years. However, the fact that membrane proteins have an amphiphilic structure creates difficulties in many steps (such as dissolving proteins) during the experiment, and generally, abundance is low.[53] Also, while it has a static structure, proteomics is dynamic. Although PTM, missense mutation, and biological activities change, mass Spectrometry still prevents complete proteomics from being developed. Increase solubility and separation with gel electrophoresis for membrane proteins and gel-free method only digest with proteolytic enzyme and analyze them. As technology advances, the analysis of membrane proteins becomes more manageable.

The most commonly used methods for extracting surface proteins from cells; are chemical capture of cell surface glycoproteins, coating the cell surface with cationic colloidal silica beads, biotinylation, and membrane shaving.[54-55]

The general principle in biotinylation is; the intact cells are biotinylated using appropriate reagents. It is aimed at the lysis of the cell and purification of biotinylated proteins by affinity chromatography. Obtained proteins are digested with proteolytic enzymes and analyzed by mass Spectrometry. An example of protein biotinylation is iodoacetyl-biotin, which binds to free sulfhydryl groups on cysteine residues. This binding results in a reduction of disulfide bonds.[55]

The principle in cationic colloidal beads is based on the electrostatic interaction of proteins on the cell surface with polyacrylic acid. Intracellular proteins are removed by centrifugation. The raw membrane pellet is purified by ultracentrifugation. Membrane proteins obtained after these processes are cut with proteolytic enzymes, and mass spectrometry protocols are applied.[56]

These methods lyse the cell and involve the risk of further analysis of the protein from within the cell. Direct determination methods are being developed to minimize the contamination steps that may come from inside the cell or the experimental environment. In studies where surface proteins are directly identified, ionic solutions (LiCl), protein denaturants (urea buffer), and shaving methods have been tried.[56]

2.5.8. Shaving Methods

Bottom-up technology is a general approach in proteomics. However, when the sample is treated with an enzyme, it must dissolve for the enzyme to break down the protein completely. The fact that the membrane proteins are insoluble prevents the enzyme from fully reaching the protein and breaking it down at specific cut points. Recently, solvents such as detergents (SDS), organic compounds, and chaotropic agents (urea) have been used instead of water. Thanks to organic and chaotropic agents used instead of water, the precipitation of hydrophobic proteins is prevented. However, membrane proteins are a complex area to study because they are both small and can be found embedded in the membrane. However, gel-based MS studies are not very suitable for surface proteomics because of the limitations in the analysis, such as the precipitation of hydrophobic amino acid-rich transmembrane proteins during ısoelectric focusing (IEF).

Rodr´ıguez-Ortega et al. used the 'shaving' method for surface protein analysis, which is simples, more reliable, direct, and faster. It is also an advantageous method due to the including small peptides of membrane proteins. The shaving method focuses on a gel-free MS experimental setup, where the cell is not broken down but digested with a proteolytic enzyme for only a short time.[57]

Principle of the method: cell was grown, harvested, and washed with buffer (such as phosphate saline buffer). Buffers keep the pH in the environment constant and do not have any denaturing or precipitating effect on the protein. The cell is not lysed in any way. Cells washed with buffer are kept briefly by adding proteolytic enzymes (trypsin, proteinase K, pepsin). This time can vary from 30 minutes to 1 hour. This process only aims to shave the proteins on the cell's surface.[56-57]

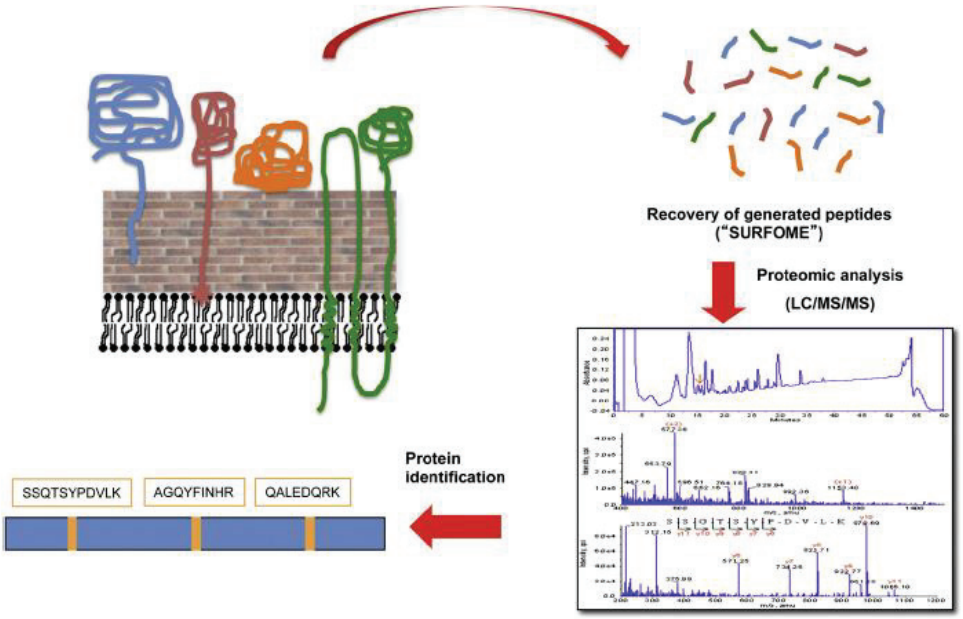


Figure 2.14. 'Shaving ' Method Workflow'[57]

The image above briefly explains the shaving method. LC/MS analyses are performed directly after shaving the surface proteins. Finally, the raw data from mass spectrometry are loaded into software tools such as the Mascot and scanned by comparing them with the existing database.

Trypsin, which performs specific digestion, and Proteinase K, which performs non-specific digestion, are generally used as proteolytic enzymes. Trypsin (EC:3.4.21.4) is a serine protease enzyme. The most preferred enzyme because it always breaks down lysine and arginine from the C terminal end. Because it has a specific digestion site, works well in-gel digestion and in-solution digestion procedures, and facilitates ionization in MS thanks to the charged peptides it creates. However, the trypsin enzyme may not be able to identify too many peptides in the preferred shaving method. Therefore, non-specific enzymes such as proteinase K and chymotrypsin are also studied. While proteinase K can work between pH 2-12, its optimum pH is 8. Ca +2 must be present in the environment for maximum activity from the enzyme. The method (e.g., pH=6 and no Ca+2 in the buffer) was optimized, and peptides at a rate of 500-3000 m/z could be produced in MS analyses. The peak intensity in the MS/MS spectra is also lower than that of tryptic peptides. Since proteinase K cut points are not specifically from the basic amino acid regions of the peptide, the peptide mixture is

more challenging to ionize by the ionization source. Enzymes are not very suitable for fragment formation with CID. Therefore, MS/MS spectra, which cannot be easily explained, can be obtained. As a result, according to the results obtained in the experiments, the number of peptides obtained from proteinase K is less than the number of peptides obtained with the Trypsin enzyme. However, in research with S. pyogenes, transmembrane peptides that could not be 'shaved' by trypsin were identified thanks to the proteinase K enzyme.[57]

Since the enzyme is desired to cut only the proteins on the cell surface, it is treated with the enzyme briefly. In addition, since cell lysis is prevented, the amount of proteins that can come from the cytoplasm or nucleus is minimized. However, in the data obtained from MS/MS studies, proteins from the cytoplasm or nucleus are encountered. Trypsin and proteinase K are enzymes with a weight of approximately 23 kDa. Since they are small, they can pass through the plasma membrane by passive diffusion and cut unwanted proteins. (such as proteins to be cleaved from nuclei or cytoplasmic regions). To overcome this, Tjalsma and coworkers immobilized trypsin to agarose beads. When the results with immobilized trypsin and free trypsin are compared, it is seen that cytoplasmic protein is obtained in both. Thus, it has been understood that proteins obtained from the cytoplasm or nucleus have other factors besides trypsin. These proteins burst in case of any shock for the cell's self-defense. Alternatively, they should be found in the cytoplasmic part of the phospholipid layer, but they usually are encountered on the surface by membrane phospholipid bilayer movements.

This method is used in gram-positive bacteria such as S. pyogenes, Streptococcus agalactiae, Streptococcus pneumoniae, Bacillus subtilis, Staphylococcus aureus, in gram-negative bacteria such as Chlamydia trachomatis, Escherichia coli, Salmonella enterica and in eukaryotes such as Trypanosoma cruzi has been tried.[57]

Leishmania is among the trypanosomatid parasites, and this family includes Trypanosoma cruzi and Trypanosoma brucei. The study of Rayner M. L. Queiroz and colleagues with Trypanosoma cruzi was the first known study for this family. In this study, biotinylation and shaving of surface proteins were compared. The results determined that both methods were suitable for surface proteins.[58] However, the ratio of the number of identified proteins in the cell was also found to be high.

Topological analysis of 37 transmembrane proteins of M1_SF370 group A Streptococcus (GAS) has been demonstrated using the shaving method with trypsin and

proteinase K. Regions of peptides obtained after proteolytic degradation by topological prediction with PSORT are colored in red. These red areas are peptide regions cut by proteolytic enzymes and identified as a result of MS/MS analysis. The identified proteins are vaccine-candidate proteins.[59]

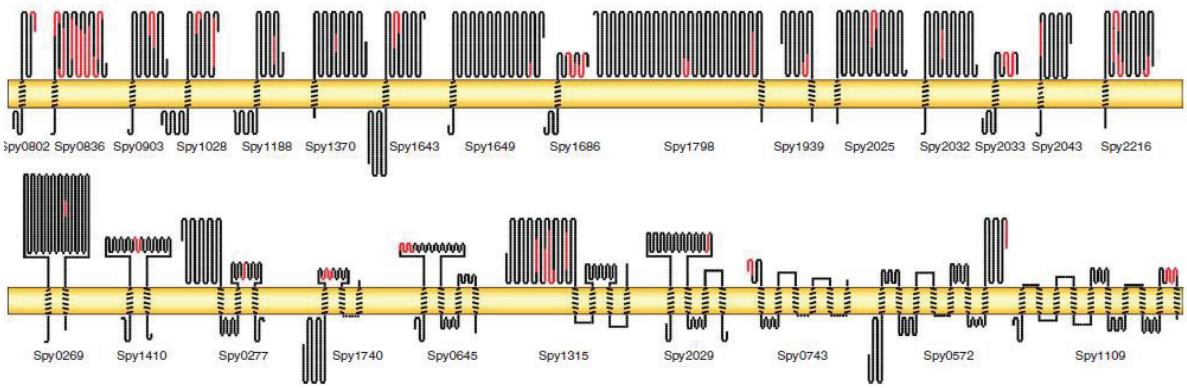


Figure 2.15. Prediction of Transmembrane Proteins Topology[59]

2.6. Aim of the study

The leishmania parasite is in the neglected tropical disease group. This disease is generally common in developing or underdeveloped countries. Therefore, most countries need to do more work on the treatment process. The treatment methods found as a result of the studies are costly. In addition to the expensive treatment process, there need to be more serological tests to diagnose the disease. The widely used rK39 40 is a practical test on visceral leishmaniasis. However, in addition to the difficulties experienced in distinguishing leishmania from other parasites in recent years, visceral and cutaneous leishmaniasis cannot be distinguished due to genetic variation between species. Recently, the incidence of leishmaniasis has increased with increasing immigration to Turkey. In this increasing disease, studies on the diagnosis and treatment process should be increased and the disease should be detected more quickly.

Because of these problems, we focused on leishmania's surface proteins to diagnose quickly and directly. As can be understood from the 'shaving' method, it is

aimed to shave the cell surface only with trypsin and thus digest only the proteins on the surface. Thus, it aims to understand better the interactions between the surface of the parasite cell and the host cell.

First, the common proteins between the *L.Tropica, L.Major, L.Infantum, L.Donovani* species will be determined. Thus, it was aimed to find a biomarker to distinguish leishmania from other parasites. Secondly, according to the results obtained, it aimed to find the proteins that distinguish them from other parasites and separate leishmania species from each other. It is expected to shed light on future research. Thus, if a protein can be a potential biomarker, it will be more obvious.

CHAPTER 3

MATERIAL and METHODS

3.1. Growth Leishmania Isolates and Shaving Method Protocol

Four leishmania species samples *(L.Donavani, L.Infantum, L.Tropica, L.Major)* were cultivated, isolated, and extracted using Prof. Dr. Ahmet Özbilgin's group in Celal Bayar University, Faculty of Medicine, Department of Parasitology.

Frozen leishmania promastigote cells were first grown in NNN (Novy-MCNeal-Nicolle) medium at 25 ^{o}C. After the cells were grown in the NNN medium, they were transferred to the RPMI-1640 medium to produce large numbers. RPMI-1640 medium containing 10% FCS, 200 U of penicillin/mL, and 0.2 mg of streptomycin/mL. When the cells reached the $10^{\wedge 7}$ logarithmic phase, they were centrifuged for 10 minutes at 4400 rpm in a sterile 15 mL tube at 4 ^{o}C. At the end of centrifugation, the pellet is incubated overnight in FCS-free RPMI 1640 medium.[24]

The medium was removed and washed three times with Tris-buffered saline (TBS). For 50mL of 10x Tris base, 1.2g Tris and 4.4g NaCl are prepared and adjusted to pH 7.4 with HCl with a pH meter. PH 7.4 is the optimum pH at which cell integrity is ensured. It was then centrifuged at 1500 rpm for 5 minutes. 8mL of TBS was added to the pellet. It was gently shaken until homogeneity was achieved and incubated overnight. It was divided into two tubes of 4 mL each. 20µg/mL trypsin was added to each tube. It was incubated in a water bath at 28°C for 30 minutes. After incubation, it was centrifuged at 3000g for 10 minutes to remove cells. The supernatant was taken into a different tube. It was centrifuged at 6000g for 5 minutes. The supernatant was stored at -20^{o} C.[58]

3.2. Experimental Workflow

Gel electrophoresis was applied as the first step of the experiment to see if the shaving method was applied successfully and if trypsin digest the plasma membrane proteins. In the second step, the shotgun proteomic workflow was applied.

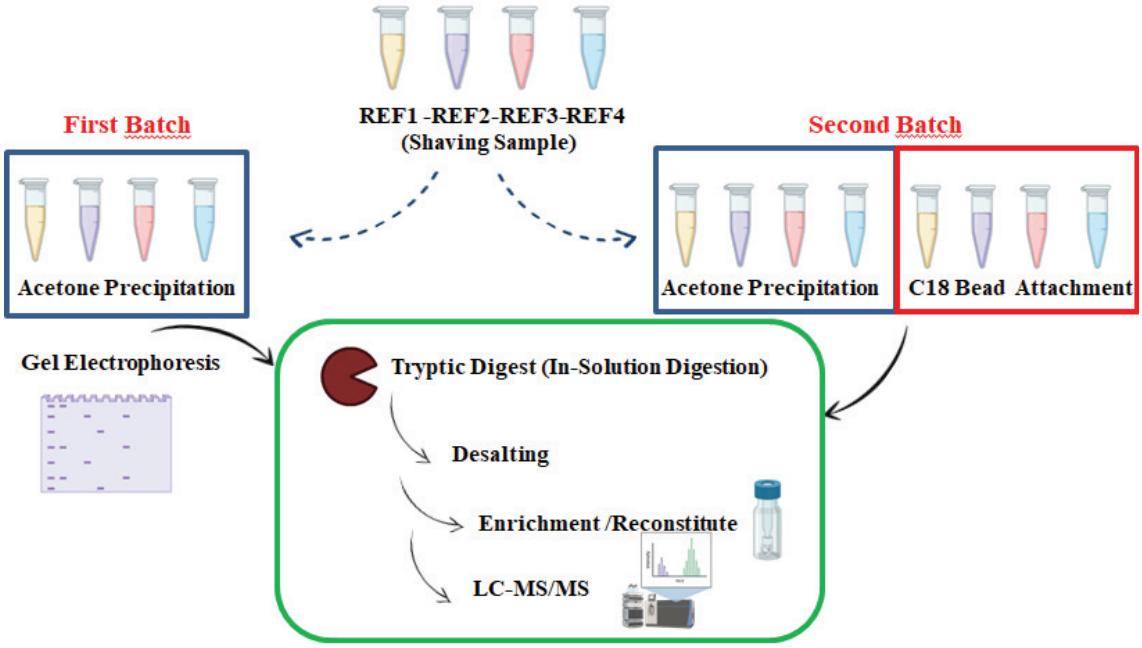


Figure 3.1. Experimental Workflow

3.3. First Batch

3.3.1. Acetone Precipitation, Gel Electrophoresis, and Pre-treatment Steps

First, gel electrophoresis was applied to see if the shaving method worked. In the first run the initial sample volume is set as 300 µL. Acetone precipitation was applied and acetone-evaporated samples were dissolved with 4µL 6M urea and 0.1M tris-HCI.

Control groups are included in the first applied Shaving samples. Both control and trypsin-treated samples (shave samples) were added to the gel. Four times the added sample buffer (16µL) was added to the samples to which 4 µL. Vortex and short-term centrifugation were done to ensure complete mixing.

Resolving gel and stacking gel were prepared for gel electrophoresis. % 12 10 mL resolving gel and %5 5 mL stacking gel were prepared.

Tablo 3.1. Requirements for the Preparation of 12% Resolving Gel and 5% Stacking Gel

12 % 10 mL Resolving Gel	5% 5 mL Stacking Gel
3.3 mL H_2 O	3.4 mL H_2 O
4.0 mL 30% Acrylamide mix	0.83mL 30% Acrylamide mix
2.5 mL 1.5M Tris (pH 8.8)	0.63mL 0.5M Tris (pH 6.8)
0.1 mL 10% SDS	0.05mL 10% SDS
0.1 mL 10% Ammonium persulphate	0.05mL 10% Ammonium persulphate
0.004 mL TEMED	0.005mL TEMED

Ammonium persulfate weighed 35 grams and made up to 350 µL with distilled water. (APS) needs to be prepared freshly. APS is important for bisacrylamide and acrylamide to interact to initiate polymerization. Tetramethyl ethylenediamine (TEMED) accelerates this process.

After preparing the resolving gel, the wells were filled. While pouring the resolving gel, it was covered with distilled water to eliminate the air bubbles formed and cut off its contact with the air. It waited for 30 minutes for the resolving gel to freeze. During the waiting period, a 1X running buffer was prepared, adding 3 grams of Tris base, 14.4 grams of glycine, and 1 gram of SDS. SDS is a detergent and causes foaming. That is why SDS was added last. Since the running buffer contains glycine, the sample can gain density.

1 µL of cytochrome C (12.000 Da) and 2 µL of human serum albumin (66.500 Da) and 2 µL of carbonic anhydrase (30.000 Da), and 20 µL of sample buffer were added and centrifuged, and this mix was added to the first well as a marker. In addition, a second mix was created with 1µL of cytochrome C and 2µL of human serum albumin, and 20µL of sample buffer, and centrifuged was added to the last well. It was pipetted next to the markers in trypsin-treated (trypsin-treated samples) and trypsin-untreated (control samples).

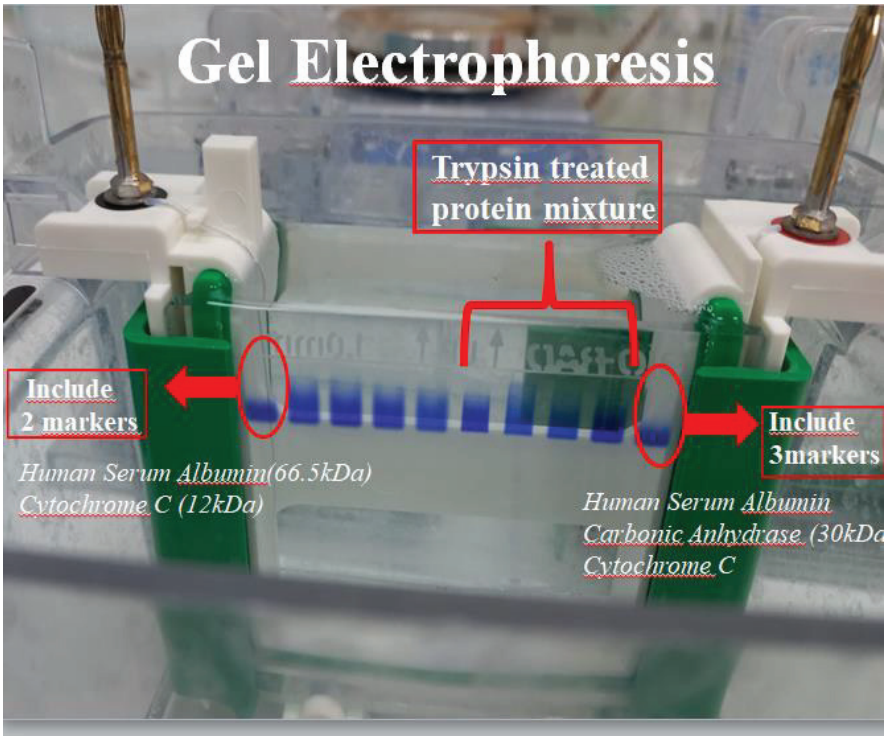


Figure 3.2. Gel Electrophoresis Design

Then it was added to the running buffer gel. 60 volts were applied for 1 hour. After this volt was applied to pass the stacking gel part, 100 V was applied for 2 hours to wait for the separation in the resolving gel. After the process, the gel was washed with distilled water and shaken for 10 minutes. Finally, Coomassie Brillant Blue G250 (CBB-G250) dye was spilled, and shaking was done overnight at 37 ^{o}C.[60]

After that, it was washed with distilled water. The gel stained with CBB-G250 dye was destained, and the dye was removed. For the first destaining, 5% (v/v) acetic

acid, 10% (v/v) methanol, and distilled water were added to the gel and shaken for 15 minutes. This process first served to remove the dyes in the background of the gel. A second destaining was done again. 5% (v/v) acetic acid, 50% (v/v) methanol, and distilled water were used. It aimed to increase the methanol rate and destroy the dyes on the bands. It can be used in ethanol instead of methanol.

The gel stained with CBB-G250 is destained, and the dye is removed. Then the silver staining protocol is applied. It is a sensitive technique that can show up to 1ng of protein. Silver staining was performed to see more precise results from the gel stained with Coomassie.

Tablo 3.2. Silver Staining Solutions Chemical

Solutions	Quantity of chemicals
Fixation Solution	%30(v/v) Methanol + %10 (v/v) acetic acid + %60 distilled water (d.water)
Rinsing Solution	%20(v/v) ethanol + %80 d.water
Sensitization Solution	75mg Sodium thiosulphate pentahydrate in 250 mL d.water
Silver Solution	0.5 g Silver nitrate in 250 mL d.water
Development Solution	3.75mg Sodium thiosulphate pentahydrate + 7.5g sodium carbonate in 25 mL. Add 175 µL formaldehyde.
Stop Solution	6.25mL acetic acid + add water + 12.5g Tris base dissolve with a magnetic stirrer

A fixation solution was added to the gel and shaken for 1 hour, and this solution was poured, and the rinsing solution was added and shaken for 15 minutes. And then it was shaken with water for 15 minutes. After pouring the sensitized solution, a shake was applied for precisely one and a half minutes. The solution was washed with distilled water twice for 20 seconds after pouring. The silver solution was added and shaken for

30 minutes. Since silver is a heavy metal, the waste solution was poured into the bin. It was rinsed again with distilled water for 20 seconds. A development solution was added and slightly shaken for 2 to 5 minutes. Then the reaction was stopped by adding a stop solution. It is a reaction that terminates the reaction in the silver staining method, depending on the pH change. After Tryptic digestion, desalting and LC-MS/MS were performed.

3.4. Second Batch

3.4.1. The Samples Using The Acetone Precipitation Method

3.4.1.1. Acetone Precipitation

The purpose of acetone precipitation is, for example, to remove contaminants and to allow the protein to precipitate. Loss may occur for small peptides depending on the peptide size it contains during acetone precipitation.

We have four trypsin-treated samples. These samples were obtained from 4 different leishmania parasites *(L.Donavani, L. Infantum, L.Tropica, L. Major)*. These proteins obtained from reference leishmania cells were named REF. It was named REF1 obtained from *L.Tropica*, REF2 obtained from *L.Major*, REF3 obtained from *L.Infantum*, and REF4 obtained from *L.Donovani*. Samples stored at -20 were thawed. Vortex was used to ensure complete homogenization. 10 mL of acetone was added to 2 mL of sample (1: 5 ratio) and left at -20 for 4 hours. It was centrifuged for +4 ^{o}C, 15 minutes at 14000 rpm. Then acetone (supernatant) was decanted. After pouring acetone, the remaining small amount was expected to evaporate in 3 minutes. 100 µL of 6M urea and 0.1M TrisHCl were added. Short-term centrifugation and vortexing were applied again.

3.4.1.2. In Solution Digestion

Samples will be digested with trypsin enzyme during in-solution digestion. The optimum pH of the trypsin enzyme is 8-8.5.

0.1 M DTT (154.25g/mol), and 0.1 M IAA (184.96g/mol) stock solution were prepared. It was covered with aluminum foil as it was affected by light. 10mM DTT and 20mM IAA were prepared from 0.1M stock solutions, respectively.

Tablo 3.3. Steps in the Experiment

1) 11 µL DTT
2) 25 µL IAA
3) 30 µL DTT (DTT was pipetted at 2 times the volume of the first added DTT.)
4) pH adjusting with 0.1M tris-HCI (600 µL)
5) 0.5 µg trypsin containing 5 µl 0.1 M tris-HCl was added to each sample.
The sample was kept at 37 ^{o}C overnight

3.4.1.3. Desalting - C18 SCE Procedure (Stage-Tip)

Solid phase extraction methods are an essential step before HPLC as it cleans the sample, remove impurities and substances that may cause interference, and ensures that the flow in the column can be better.

It was passed with C18 containing 50% MeOH and 50% d.water in 1000 µL pipette tips. Pipette tips containing C18 weighed 15mg. Pass through with 0.1% TFA to remove methanol from the medium. If the methanol is not removed, the samples will not adhere. 100 µL of 5% TFA was added to increase the medium's acidity. It was then passed with 50µL of 50% ACN-water, 50µL of 50% ACN-water, 80% ACN-water, and

100% ACN, respectively. The samples were evaporated with SpeedVac until a small portion remained in the Eppendorf.

Stage-tip pipettes (homemade zip-tip) were prepared as C18 columns.[61] The purpose of solid phase extraction (SPE) aims to enrich the peptides before LC/MS and purify them from impurities. The C18 column is hydrophobic. Then, the sample is eluted with a solution with a higher acetonitrile (ACN) ratio. The sample remaining in the Eppendorf is dried with a Speed-Vac. This drying process is crucial for removing acetonitrile from the sample environment. Because only solvents in mobile phase A (100% d.water+0.1% FA) should be present in the sample before LC/MS, Acetonitrile remaining in the sample can disrupt the gradient flow in HPLC. Gradient formation can be seen earlier than expected. However, a little solution should remain in the sample during the removal of acetonitrile with the Speed-Vac process because the hydrophobic peptides will be challenging to dissolve since mobile phase A contains 100% d.H2O and 0.1%FA.

The samples were reconstituted to 15 µL. After the nanodrop measurement, it was added to the vial, and LC/MS was made.

3.4.2. The Samples Using The C18 Bead Attachment

C18 bead attachment was tried instead of acetone precipitation to see if sample loss occurred with acetone precipitation. The other method's flow aim is the same.

Homemade zip-tip making (Stage-Tip); 4 different 1000µL pipette tips were filled with C18 (Solid phase exchange cart material) material containing 50% (v/v) methanol + 50% water. After that, the tips contain 15mg of C18. Then, C18 loaded pipette tips were passed through 200 and 500 µL d.water to remove methanol. After washing, it was checked whether there was flow from the pipette tips without pipetting. 2mL samples were passed through the tip in a C18-loaded pipette. It was rewashed with 100 µL of water. Then 30 µL of DTT was passed and left for 30 minutes. 30 µL of IAA was passed and left in the dark for 30 minutes. 0.5µg of trypsin in 0.1M tris-HCl was added to the C18 micropipette tips and overnight incubated. The samples retained in the tip were passed through 50%-50%-80% ACN-water and taken to Eppendorf. It was evaporated with a speed vac. It was then made up to 15µL with 0.1%TFA. Nanodrop measurements were made. LC-MS/MS analysis was performed.

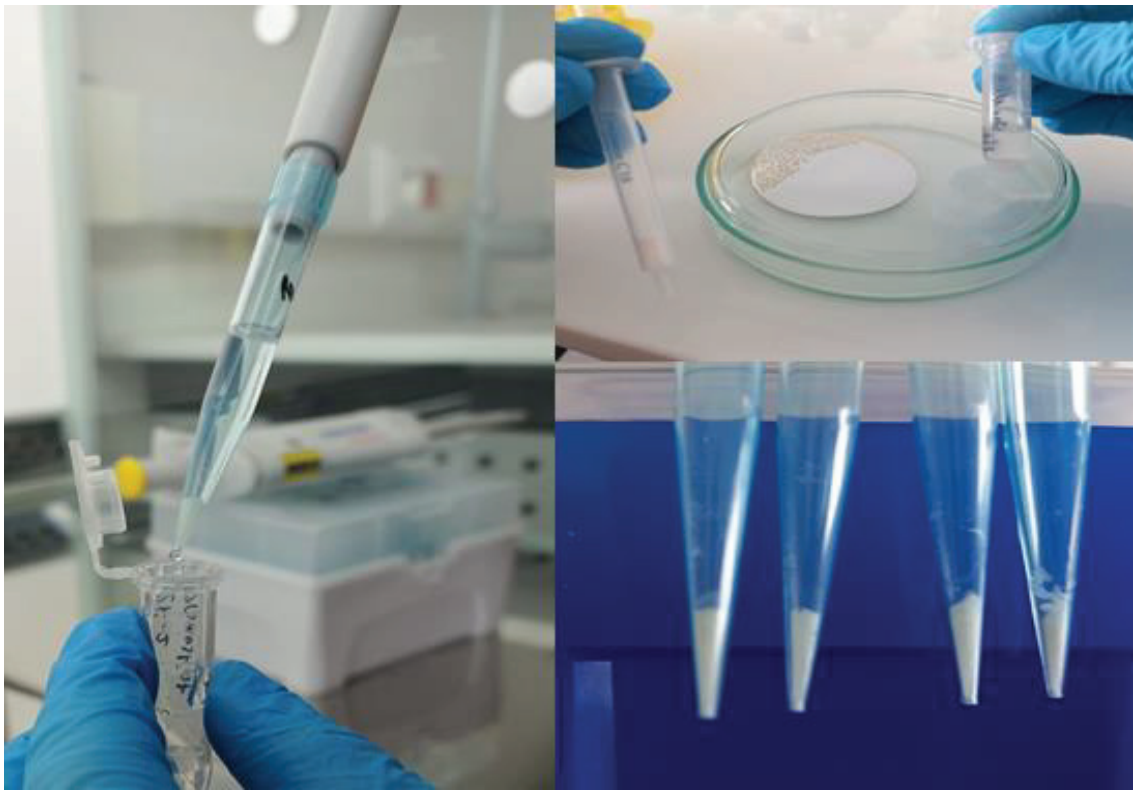


Figure 3.3. Homemade C18- Stage Tip Preparation

3.5. MS/MS Analysis and Database Search

After evaporation, the peptides were diluted with a 2% ACN, 0.1% TFA solution identical to the mobile phase of the loading pump. 5 µL of the sample was loaded onto the trap column (5µm; 0.3x5mm) for 4.5 minutes. The flow rate was kept at 10ul/min. Then, the trapped peptides were directed to the separation column (75µmx25cm C18; 3µm 100A) with gradient flow at 0.3µl/min for 60 minutes. The mobile phases were 0.1% formic acid in water for solvent A and 0.1% FA in acetonitrile for solvent B.

Tablo 3.4. HPLC Gradient Flows

Time (min)	**%B**
0	2
5	2
6	5
36	25
40	35
42	90
45	90
50	4
60	4

Data were collected according to data-dependent acquisition mode. The Nanospray voltage was 2000 V at positive mode, and the ion transfer tube temperature was 275 C. MS data were collected with an orbitrap analyzer with 120000 resolution. The scan range was 375-1500 m/z. The intensity threshold for the MS/MS analyses was 5000. The peptide signals whose charge state was between 2-7 exceeded the threshold; They were ranked in the fragmentation list for 60 minutes (dynamic exclusion time is 60 minutes) and then transferred to the ion trap analyzer during 3 seconds cycle. The isolation width of the quadrupole analyzer was 1.6 m/z, and the activation type was HCD fixed at 35%. The data type is centroid.

Results from LC-MS/MS analysis were combined for each sample. The results were searched in the UniParc database with Sequest. UniParc database was installed from the UniProt website (www.uniprot.org) in FASTA format. The Uniparc, Leishmania species 58,985 proteins have been identified in this database. UniParc database was established from the UniProt website (www.uniprot.org) in FASTA format. The search parameters set in Sequest are shown in Table 3.4.

Tablo 3.5. Sequest Parameter

Input Data	
Protein Database	Uniparc_Leish38568
Enzyme Name	Trypsin
Static Modification	Carbamidomethyl (+57.021 Da)
Dynamic Modification	Oxidation (+15.995 Da), Acetyl (+42.011 Da)
Max Missed Cleavage Sites	2
Min. Peptide length	6
Max peptide length	144
Peptide Charge	2+ and 3+
Tolerances	
Precursor Mass Tolerance	10 ppm
Fragment Mass Tolerance	0.6 Da
General Setting	
Precursor Selection	Use MS1 precursor
Scan Event Filters	
Polarity Mode	+
Other parameters are neglected	

As a result of database scanning with Sequest, common and different proteins obtained from 4 different Leishmania species of were classified and determined in an Excel sheet.

CHAPTER 4

RESULTS and DISCUSSION

Enzyme administration time was kept short so the cell would not be lysed. Keeping the treatment with the trypsin enzyme short is important so that the cell does not feel under stress and does not explode. Trypsin breaks down proteins on the cell surface. It cuts the cells from the amino acids arginine and lysine at the C-terminus of the proteins in the parts that adhere to the flask.

This procedure was applied for four Leishmania species *(L.Donavani, L.Infantum, L.Major, L.Tropica).*

Tablo 4.1. Sample Names and Species of Parasite

REF1	**REF2**	**REF3**	**REF4**
L.Tropica	L.Major	L.Infantum	L.Donavani

4.1. First Batch Results

The first study included a control group. The results were compared with the control group. The purpose of performing gel electrophoresis in this study was to understand whether trypsin could cut through proteins around the cell. It was performed to estimate the average mass of the proteins in the sample and to make a comment about whether the trypsin enzyme breaks down the proteins on the cell surface. Coommassive staining and silver staining were done. There was no clear result in staining with Coomassie Brillant Blue G250. Silver staining, which is approximately 100 times more sensitive, was performed. The molecular weights of three different markers loaded on

the gel are as follows: Cytochrome C (12.000 Da) and human serum albumin (66.500 Da), and carbonic anhydrase (30.000 Da). Trypsin is a 25 kDa enzyme. The most prominent spots come from the trypsin. The spots in the area marked with red in figure 4.1 belong to trypsin-treated samples.

The reasons for not showing results other than trypsin with gel electrophoresis may be;

- Although gel electrophoresis in effectivelty dissolve membrane proteins with zwitter ionic detergents such as SDS, it may be insufficient for the visualization of small hydrophobic peptides. Samples were treated with trypsin twice. (Shaving method and tryptic digest before LC-MS) Therefore, medium has plenty of trypsin compared with trypsin-treated samples.
- The abundance of arginine and lysine amino acids on the plasma membrane enhances the cut by trypsin. Thus, the sample was thought to contain small molecular weight peptides.

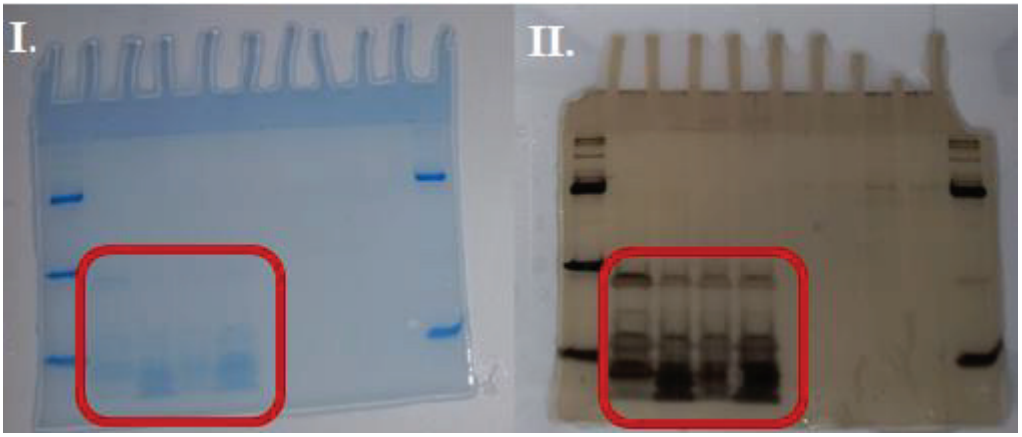


Figure 4.1. I.Result of Coomassie Staining, II. Result of Silver Staining

The number of proteins identified especially in *L. tropica* and *L. Infantum* was higher than in the control groups showing that the shaving method worked. (Shown in the Figure 4.2) In the second batch, the sample amount was increased and incubated with TBS overnight, and it was aimed to remove the dead cells.

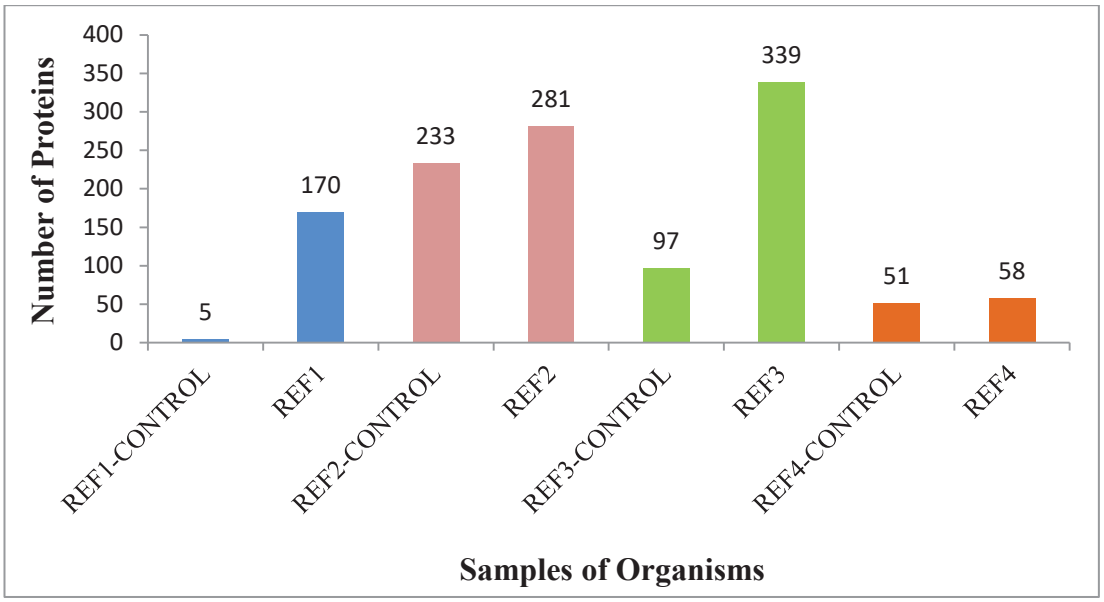


Figure 4.2. Protein Counts Defined in *L.Tropica, L.Major, L.Infantum, L.Donavani*

4.2. Second Batch Results

Uniparc database was used as the database. In this database, the Uniparc Leishmania species 58,985 proteins have been identified. The UniParc database was preferred over UniProt to reduce the redundancy problem in the results. Uniparc and LC-MS data results were matched in the Proteome Discoverer 2.5 (Thermo Scientific) using the Sequest software. Result tables were arranged by expressing the obtained UniParc ID codes as UniProt. UniParc code can correspond to more than one UniProt ID code. Therefore, peptide domains defined in raw data for UniProt expression from Uniparc code were blasted.

4.2.1. Results of Samples Using Acetone Precipitation Method

The work with the Fusion Orbitrap mass spectrometer with the identified proteins is shown in Figures 4.3 and Figure 4.4. Samples from all species of Leishmania were analyzed in two technical runs by HPLC-MS are shown in Figure 4.3. Proteins identified in both technical runs are shown in figure 4.4. The number of identified common proteins is REF1: 717, REF2: 278, REF3: 619, and REF4: 718.

The similar and different protein numbers of the proteins identified in *L.Tropica, L.Major, L.Infantum, and L.Donovani* in both methods are shown in the Venn diagram in Figure 4.5.

For the shaving method to work effectively;

- The amount of enzyme added to the Leishmania promastigote cells,
- The incubation time,
- The temperature of the medium,
- The pH of the buffer,
- The effective growth of the cells during cell culture is important.[58]

In addition, using nano LC and Orbitrap fusion mass spectrometer with 120.000 resolution affects the number of proteins identified in the results.

REF1.1 and REF1.2 are samples obtained from *L.Tropica*. It means that two technical runs are made from each vial. The reason for the difference in the number of identified proteins between REF1.1 and REF1.2; depends on the instrument. Influencing factors;

- Reasons such as relative intensity difference in peaks,
- Voltage variation during analysis (Electrospray voltage)
- Data-dependent acquisition randomness
- Signal deterioration due to voltage variation
- Peaks with low intensity may not be matched in the Sequest scan

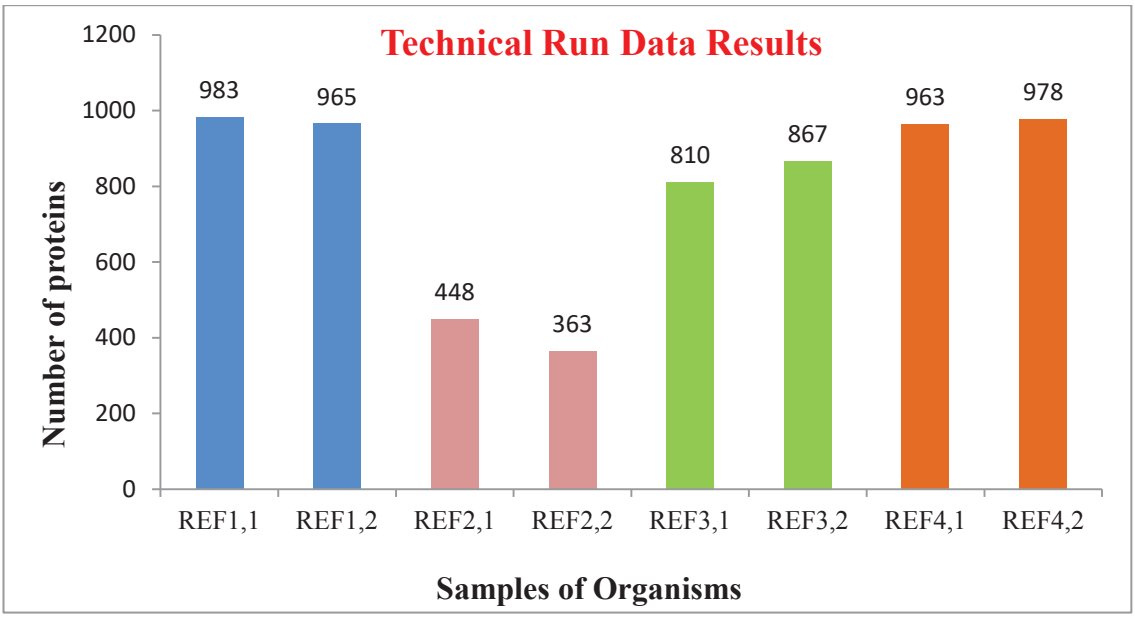


Figure 4.3. Protein Counts Defined in *L.Tropica, L.Major, L.Infantum, L.Donavani*

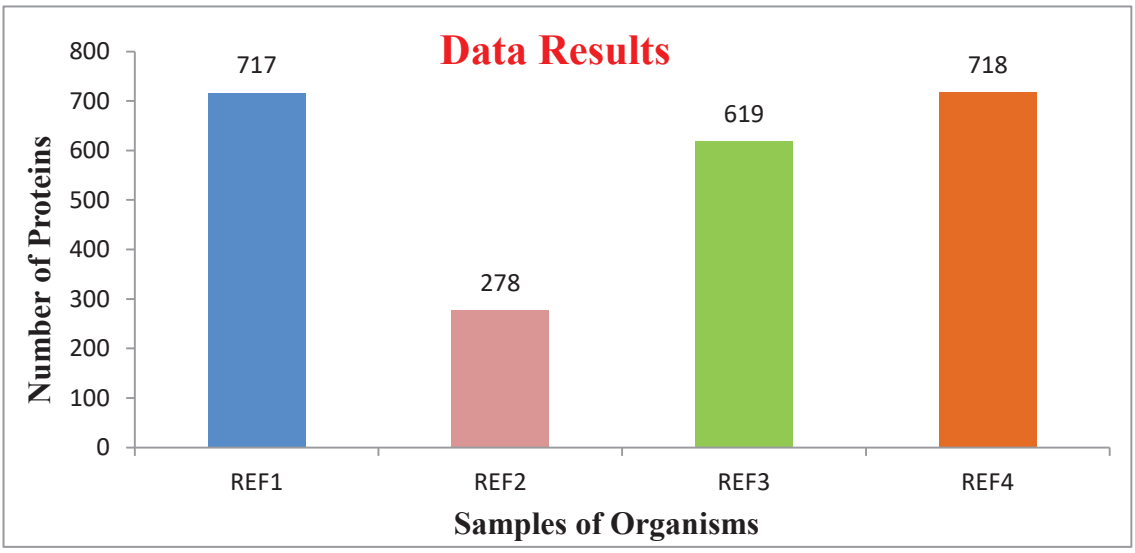


Figure 4.4. Number of Proteins Identified as Common in Both Technical Run

The protein numbers identified from *L.Tropica, L.Infantum, and L.Donovani* are close; however; the number of proteins identified in *L. Major* is low. This difference is because *L. Major* did not grow adequately during cell culture studies.

In Figure22. Identified proteins in 4 species were compared and schematized in a Venn diagram. From this, the results were concluded by first starting with the proteins seen in only one species and then examining the commonly identified proteins.

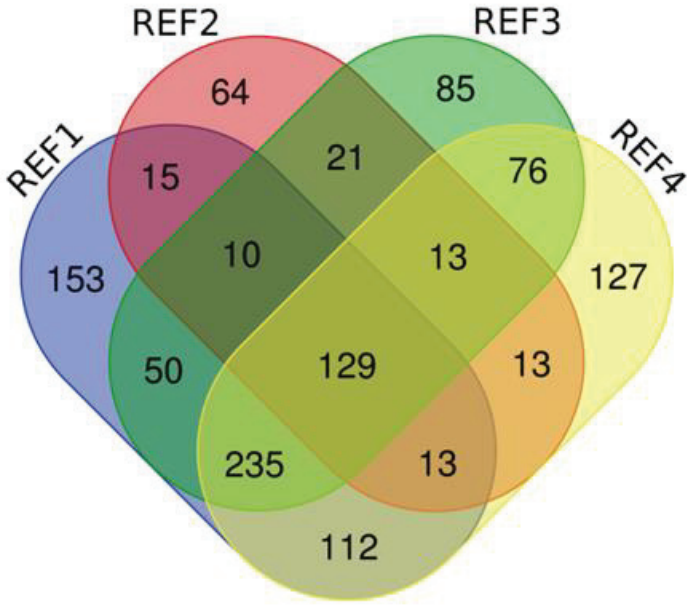


Figure 4.5. Venn diagram (REF1: *L.Tropica,* REF2: *L.Major,* REF3: *L.Infantum,* REF4: *L.Donavani*)

Tablo 4.2. Numbers of Commonly Identified Proteins in 4 Different Samples

Samples	Surface Proteins / Total Identified Proteins
REF1-REF2-REF3-REF4	10/129
REF1-REF2-REF3	2/10
REF1-REF2-REF4	1/13
REF2-REF3-REF4	0/13
REF1-REF3-REF4	17/235
REF1-REF2	2/15
REF3-REF4	5/76
REF1-REF3	4/10
REF1-REF4	9/112
REF2-REF3	3/21
REF1	18/153
REF2	4/64
REF3	7/85
REF4	8/127

The subcellular localization of the identified proteins is shown in the pie chart in the Figure 4.6. When compared with the literature, it is seen that cytosolic proteins are very abundant in the sample. Some of the reasons; may include the long incubation period during cell culture, the presence of exoproteome proteins, high resolution, and nano-level analysis. It is possible to 'exoproteome' proteins on the cell surface. The definition of exoproteome includes the cell's lysis, proteins found on the cell surface, and secreted proteins. The intracellular proteins seen in the results are predicted to be triggered by unexpected stress, shock, or burst of cell or environmental factors and released to the cell surface.[62]

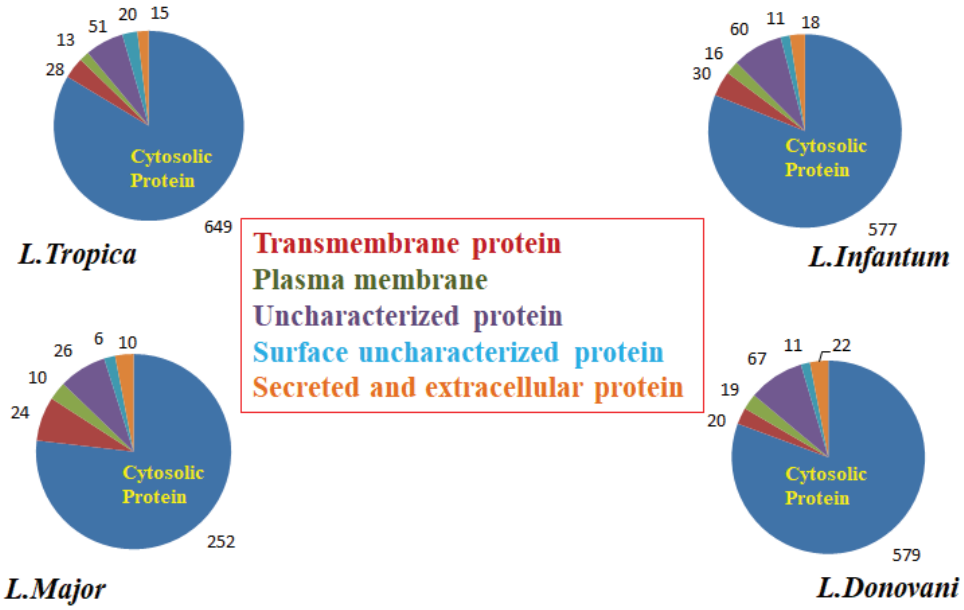


Figure 4.6. Protein Distribution

The proteins in the tables are associated with the cell membrane and the extracellular region. The proteins are likely to be found on the surface of that species. All identified proteins are given in Appendix B. The tables give protein length, peptide spectrum match (PSM), Xcorr score, and antigenic property score.

The one number of the PSM is also included in the table. Proteins with a PSM of one may also have been randomly matched, However, they could be adressed because of the important plasma membrane proteins identified in each biological repeat.

The number of samples obtained from *Leishmania Tropica* samples (REF1) is 717. The similar protein of REF1 with REF2 is 15, the similar protein with S3T is 50, and the similar protein with REF4 is 112. Among the proteins found only surface in REF1, those found on the surface are given in Table 4.3. Eighteen protein out of 153 protein as surface proteins were identified.

Tablo 4.3. Proteins Identified only in *Leishmania Tropica* (REF1)

UniProtKB Code	**Identified Protein**	**Length**	**Peptides**	**PSM**	**Theoretical MH+ (Da)**	**X Corr**	**Probably Antigenic Score (x>0.5)**	**Location of Protein**
A4HYA0	ATP-binding cassette protein subfamily F, member 2 (ABC transporter)	612	1	1	1260.607	3.13	0.49 no antigenicity	Plasma membrane
E9B1J0	Putative calpain-like cysteine peptidase	784	5	6	1618.793	4.29	0.54	Extracellular
Q4QEA3	META domain-containing protein	444	5	6	1216.61	3.87	0.49 no antigenicity	Cytoplasmic, cilium, extracellular, plasma membrane
P14700	Membrane antigen containing repeating peptides (Clone 39)	53	2	5	1102.512	2.68	0.94	Membrane, secreted
E9ACH4	FYVE-type domain-containing protein	2354	1	1	1498.761	3.8	0.74	Extracellular
E9AFQ3	FYVE-type domain-containing protein	377	1	1	-		0.74	Extracellular
E9B2S6	Putative ATP-dependent zinc metallopeptidase (EC 3.4.24.-)	719	1	1	1615.89	3.8	0.66	Membrane
A0A504X297	Nucleoside transporter family protein	513	1	1	1184.579	2.2	0.53	Membrane
A4I4T9	Uncharacterized protein	1237	1	1	2101.975	2.95	0.60	Secreted,membrane
E9BJ67	Snf7, putative	215	1	1	1680.719	3.34	0.46	Secreted, vacuolar transport
A0A6J8FD18	Glycerophosphoryl diester phosphodiesterase family putative	353	1	1	1461.733	3.88	0.67	Extracellular,Plasma membrane
A0A504Y4S8	Leucine-rich repeat-containing protein 51	875	1	1	1691.773	2.93	0.51	Cilium

(Cont. on next page)

Table 4.3. (cont.)

UniProtKB Code	Identified Protein	Length	Peptides	PSM	Theoretical MH+ (Da)	X Corr	Probably Antigenic Score (x>0.5)	Location of Protein
A0A504Y3V5	Uncharacterized protein	150	1	1	2385.055	3.66	0.8	Membrane
Q4QCJ6	Uncharacterized protein	724	1	1	1354.667	2.5	0.51	Integral membrane-phagocytic vesicle
E9AHI1	Uncharacterized protein	123	1	1	1418.662	3.51	0.95	Mitochondrial, plasma membrane
A4HU23	Uncharacterized protein	615	3	3	1428.733	4.71	0.38 no antigenicity	Extracellular, secreted
A0A504Y3V5	Uncharacterized protein	150	1	1	2385.055	3.66	0.80	Membrane
E9AHI1	Uncharacterized protein	123	1	1			0.95	Membrane

Transport proteins maintain the homeostasis of the cell with the metabolites they carry between the inside and outside of the cell. Some transport proteins identified by the Shaving method; are PPM (PP2C), Ca+2 ATPases, Vacuolar H+-ATPase (V-ATPase), and ABC transporters. ATP-binding cassette protein (ABC) is the most prominent transporter family in that proteins significantly regulate the import and export of substances across plasma membranes. Driven by ATP, ABC transporters can carry a variety of substances across intracellular and extracellular membranes, such as small inorganic or organic molecules, sterols, metal ions, lipids, sterols, and polypeptides. ABC proteins, also known as ATP binding cassette proteins, are found in various locations in leishmania.[63-64] These proteins throw thiol-antimony compounds used in the treating leishmania out of the cell and create drug resistance. Antigenic estimates of the identified ABC proteins were analyzed with VaxiJen 2.0 [65] as threshold 0.5. In this case, a 0.49 antigenic ratio was found in the ABC transporter protein. The defined peptide area (N' LDELDADTAAAR 'C) does not contain antigenic or epitope properties.

In the results, the antigenic properties of FYVE-type domain-containing protein (E9ACH4), Uncharacterized protein (E9AHI1), Uncharacterized protein (A0A504Y3V5), SAM domain (A0A6L0XLC2) proteins identified only in *L.Tropica* are above the threshold (0.5). The antigenicity of the N'QQLAANAEELQQR'C peptide range found in the extracellular region defined in the FYVE-type domain-containing protein (E9ACH4) is 0.72. This peptide region with high antigenicity may be important

in peptide-based diagnosis kits. Each amino acid in the N'EAPAISMWEADAK'C peptide domain identified in the uncharacterized protein (E9AHI1) shows an average epitope probability of 0.6. The epitope domains in proteins showing antigen characteristics are important because they are antibody-binding sites. The identified peptide domain is located in the cytoplasmic region. It may have been seen in the plasma membrane or the extracellular region, possibly by binding to a ligand.

The number of proteins obtained from *Leishmania Major* samples (REF2) is 278. Its similar protein with REF3 is 21, and its similar protein with REF4 is 13. Among the proteins found only in REF2. The proteins found on the surface are given in Table 4.4. Four proteins out of 64 proteins as surface proteins were identified.

Tablo 4.4. Proteins Identified only in *Leishmania Major* (REF2)

UniProtKB Code	**Identified Protein**	**Length**	**Peptide**	**PSM**	**Theoretical MH+ (Da)**	**X Corr**	**Probably Antigenic Score (x>0.5)**	**Location of Protein**
P08148	Leishmanolysin	602	3	3	2057.981	3	0.58	Plasma Membrane
Q1X7K5	Excreted/secreted protein 27	104	1	1	810.5448	2.12	0.78	Secreted
Q4Q843	Putative glycoprotein 96-92	716	1	1	1304.596	3.52	0.91	Cytoplasm, secreted
Q4QIM8	Calcium transporting ATPase	1104	2	2	1389.613	3.14	0.50	Transmembrane

Although the function of the excreted-secreted 27 protein (Q1X7K5) is not known precisely, it contains a signal sequence. (1-44 aa position) The secreted pathway increases the probability of the protein being found in the extracellular region or plasma membrane. This protein was identified by the N'IPVLQLK'C peptide sequence. This

peptide region is located in the signal sequence and the average probability of being the epitope region is around 0.5 per amino acid. The antigenicity of the signal region is strong. It can be an important peptide domain for epitope and antibody interaction. Excreted secreted 27 protein was identified in all repeats despite processing with a single peptide and peptide spectrum matching (PSM).

The number of proteins obtained from *Leishmania Infantum* samples (REF3) is 619. Its similar protein with REF4 is 76. The number of REF1, REF2, and REF3 similar proteins is 10. The similar protein identification number in REF2-REF3-REF4 is 13. Among the proteins found only in REF3, the proteins found on the surface are given in Table 4.5. Seven proteins out of 85 proteins as surface proteins were identified.

Tablo 4.5. Proteins Identified only in *Leishmania infantum* (REF3)

UniProtKB Code	**Identified Protein**	**Length**	**Peptide**	**PSM**	**Theoretical MH+ (Da)**	**Xcorr**	**Probably Antigenic Score (x>0.5)**	**Location of Protein**
A0A6J8FFA7	Hydrophilic acylated surface protein b	134	1	2	2425.936	4.63	1.2	Secreted, Golgi, endoplasmic reticulum
A4HVP9	Nucleobase transporter	501	3	3	1666.744	3.02	0.66	The plasma membrane, membrane
A0A504WXZ6	Calpain cysteine protease family protein	724	1	1	1685.816	3.11	0.48	Cell membrane, cytoplasm
E9APR0	Elongation of fatty acids protein (EC 2.3.1.199	285	1	1	1003.557	2.2	0.43	Membrane
A4HVP9	Nucleobase transporter 3	501	1	1	1666.744	2.88	0.66	Plasma Membrane
A4HYI8	Putative ATP-dependent zinc metallopeptidase	598	1	1	1174.596	2.89	0.37	Membrane
E9BTP4	Uncharacterized protein	106	1	1	1171.574	2.03	0.37	Secreted, plasma membrane

Hydrophilic Acylated Surface Protein (HASP)(A0A6J8FFA7) proteins are heterogeneous acidic surface proteins expressed in infective metacyclic promastigote and amastigote species of Leishmania and are important in metacyclogenesis.

HASP proteins are 'non-classically' secreted proteins. In other words, it does not contain the signal area, and instead of being released from the Golgi and endoplasmic reticulum, it follows a different path in secreting. The research shows that the 18 amino acid domain (MGSSCTKDSAKEPQKSAD) in the N-terminal domain of the HASPB protein is necessary for the surface to be secreted. Include dual acylation at the N terminal domains such as myristoylation, and palmitoylation. Lipid modification increases its hydrophobicity. It is an adequate protein for transmitting the parasite to the host and its survival within the macrophage. The peptide domain analyzed as our experiment results are N' terminal NEDGHNVGDGANDNEDGNDDQPK' C terminal. When this peptide domain is released to the surface of the metacyclic promastigote, it is cleaved by trypsin from the peptide domain closer to the C-terminal part.[66-67-68] The HASPB protein is highly antigenic. The defined peptide domain is located at positions 107-129. Its epitope probability is 0.7. (Threshold 0.5)

It was determined that the nucleobase transporter protein has 11 transmembrane domains with the DeepTMHMM - Predictions bioinformatic tool. The co-identified N'GNEIYGDLVDEEADK'C peptide domain is at amino acid position 274-288. This area is located inside the cell. A tryptic digest may have been made in the intracellular part due to the movement of membrane proteins. This protein has many amino acid positions with high antigenicity and high epitope probability. Its significane for leishmania is unknown, but the lack of expression of this protein affects the survival of Leishmania.

The number of identified proteins from *Leishmania Donavani* samples (REF4) is 718. The similar protein number obtained in all species is 129. Among the proteins found only in REF4, the proteins found on the surface are given in Table 4.6. Eight proteins out of 127 proteins as surface proteins were identified.

Tablo 4.6. Proteins Identified only in *Leishmania Donavani* (REF4)

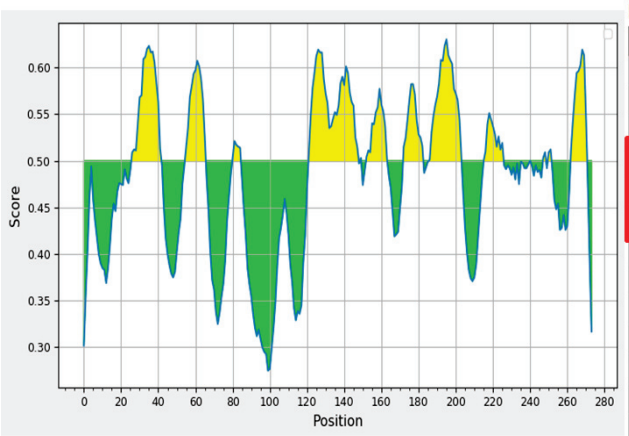
UniProt KB Code	Identified Protein	Length	Peptides	PSM	Theoretical MH+ (Da)	X Corr	Probably Antigenity Score (x>0.5)	Location of Protein
A4I7B5	Phosphoglycan beta 1,3 galactosyltransferase 5	1130	1	2	1115.599	2.79	0.48	Extracellular, mitochondrial
Q4QIZ9	Proteophosphoglycan ppg4	2783	1	1	784.46756	2.23	0.70	Plasma membrane, nucleus
A4HWN7	Phosphodiesterase (EC 3.1.4.-)	940	5	5	1115.599	3.49	0.51	Secreted
A0A3Q8IUV8	H(+)-transporting two-sector ATPase (EC 7.1.2.2)	610	5	6	1308.665	2.66	0.41	Membrane- cilium –variable location
A4HXI2	Hydrolase-like protein	390	2	2	1198.679	2.67	0.58	Plasma membrane
A4I4S8	Uncharacterized protein	1436	1	1	1504.885	2.95	0.60	Plasma membrane
E9BC31	Uncharacterized protein	148	1	1	1266.662	1.92	0.63	Mitochondrial and plasma membrane
A4IAI3	Uncharacterized protein	274	1	1	1660.865	2.5	0.53	Mitochondrial, cytoplasm, plasma membrane

Phosphoglycan beta 1.3 galactosyltransferase 5 (A4I7B5) protein identified in *L.Donovani* plays a role in lipophosphoglycan (LPG) synthesis. LPG is an essential immunological molecule for Leishmania. LPG in the parasite membrane; helps to hide the parasite from the immune system of the host cell by inhibiting NADPH oxidase activity.[68] But the antigenicity score is below the threshold.[69]

Phosphodiesterase (A4HWN7) enzyme was found in the cytoplasm using the Deeploc1 bioinformatic tool, but localized in the plasma membrane in the analysis performed with the Wolf PSORT II.[71] Phosphodiesterase regulates cytosolic cyclic AMP.[70]

The distribution plot of the antigenic property of the uncharacterized protein (A4IAI3) and the candidate epitope peptide domain is given in Figure 4.7. A threshold of 0.5 was accepted. While the yellow parts indicate the amino acid position with the antigenic feature; The green-colored regions indicate amino acid positions below the threshold. The analysis performed with UniProt, showed that the homologous protein

with 100% amino acid similarity to this protein is the alpha/beta hydrolase family protein. Peptide region identified from raw data 140N' APANATGLAELNSYL R' C155. This peptide domain is a region of high antigenicity. However, the protein having a PSM ratio of one requires repeating the results.



Predicted peptides:

No.	Start	End	Peptide	Length
1	27	42	NESIADLPLSRLTPTE	16
2	56	66	FSEAAPMTSDS	11
3	81	85	EDILR	5
4	122	148	AQLQRCLPSSYPVGELLRRAPANATGL	27
5	150	150	E	1
6	153	164	SYLRDLIPDPAE	12
7	173	183	EVRRGLARFRF	11
8	187	204	LLQYTEPFVLNSPAGSTF	18
9	217	226	GAISPEEEAR	10
10	241	241	T	1
11	248	249	SP	2
12	251	252	VR	2
13	263	271	MLGEINEAN	9

Figure 4.7. B Cell Epitope Predicted Peptides Areas

The number of identified proteins in *L. Tropica* (REF1) and *L. Major* (REF2) is 15. Among these proteins, the protein on the surface or is thought to be on the surface is also given in Table 4.7. Two proteins out of 15 proteins as surface proteins were identified.

Tablo 4.7. Common Proteins Identified in *L.Tropica* and *L.Major*

UniProtKB Code	**Identified Protein**	**Length**	**Peptides**	**PSM**	**Theoretical MH+ (Da)**	**X Corr**	**Probably Antigenic Score (x>0.5)**	**Location of Protein**
A0A3Q8I931	Tb-292 membrane-associated protein-like protein	14173	4	4	2059.941	3.67	0.58	Secreted
Q4QB61	Uncharacterized protein	150	1	1	1211.455	2.23	0.98	Secreted

L.Tropica (REF1) and *L. Infantum* (REF3) similar identified protein number is 50. Among these proteins, the protein on the surface or is thought to be on the surface is also given in Table 4.8. Four proteins out of 50 proteins as surface proteins were identified.

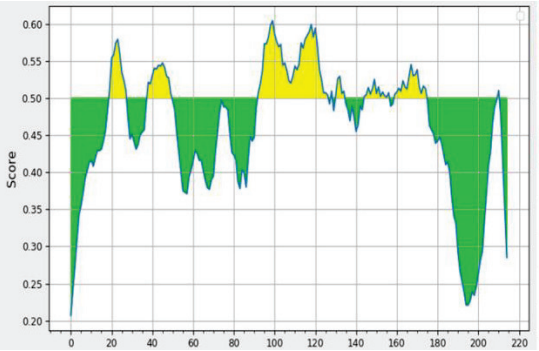
Tablo 4.8. Common Proteins Identified in *L.Tropica* and *L.Infantum*

UniProt KB Code	**Identified Protein**	**Length**	**Peptides**	**PSM**	**Theoretical MH+(Da)**	**Xcorr**	**Probably Antigenic Score (x>0.5)**	**Location of Protein**
P36982	Protein phosphatase 2C (PP2C) (EC 3.1.3.16)	563	7	9	1664.7	3.51	0.57	Flagellar protein, cytoplasm, secreted
Q4QIG6	Putative Qc-SNARE protein	215	4	4	1252.602	2.49	0.63	Vesicular membrane transport, secreted (lysosome, vacuole)
A0A504XAQ2	Eukaryotic translation initiation factor 3 subunit 7 (eIF-3) family protein	984	1	1	1130.569	2.56	0.55	Plasma Membrane
A0A6J8FRU2	Clathrin heavy chain	1693	10	12	2513.318	3.09	0.46	Cytoplasmic vesicle membrane, secreted

It is known that serine protein phosphatases (EC: 3.1.3.16) are excreted-secreted (ES) proteins secreted during the interaction of host and parasite. Although the exact biochemical function of the metal-dependent protein phosphatase (PPM-PP2C), in this protein class, is unknown, it is known to trigger a decrease in nitric oxide production and down-regulation of cytokines (TNF alpha-IL-6) by macrophages in the host.[72]

The SNARE (soluble N-ethylmaleimide-sensitive factor adapter proteins receptors) protein is also a fusion protein. SNARE protein is an intracellular vesicle-mediated membrane-bound protein. These proteins are classified as Q or R according to the residue in their center. Depending on the position of glutamine in the protein, it is

classified as Qa, Qb, and Qc. The Qa class SNARE (syntaxins) protein is a membrane-integrated protein participating in exocytosis.[73] SNARE protein complexes are important fusion proteins for membrane fusion. SNARE proteins bind to other proteins and cause structural changes in the cell membrane. This structural change causes the cell membranes to be pulled together. SNARE can affect the host-parasite interaction by influencing the molecule, and protein expression level on the cell surface. Putative Qc-SNARE protein was identified in *L.Tropica* and *L.Infantum* by peptides N' IVACVTSDMR'C, N' GSVGPQGAELR'C. Both peptides are located in the intracellular portion of the membrane. However, N' GSVGPQGAELR' C peptide is at position 92-103 and has high and linear antigenicity. There are some important steps for selecting the epitope region. One of them is that it is antigenic and it is advantageous that it has a large epitope peptide domain.



Predicted peptides:

No.	Start	End	Peptide	Length
1	20	28	AGEQPKLAE	9
2	39	50	PKEYRHKTIEDK	12
3	93	127	GSVGPQGAELRNGKKLLQQKMEFYNNPQNDKITAL	35
4	129	129	D	1
5	131	135	INQVV	5
6	145	157	VLARGDRIDTLHE	13
7	160	175	STLADQAQQFQQRSTE	16
8	211	211	F	1

Figure 4.8. B Cell Epitope Predicted Peptides Areas[74]

L.Tropica (REF1) and *L.Donavani* (REF4) similar identified protein number 112. Among these proteins, the protein on the surface or is thought to be on the surface is also given in Table 4.9. Ten proteins out of 112 proteins as surface proteins were identified.

Tablo 4. 9. Common Surface Proteins Identified in *L.Tropica* and *L.Donavani*

UniProt KB Code	Identified Protein	Length	Peptide	PSM	Theoretical MH+ (Da)	X Corr	Probably Antigenic Score (x>0.5)	Location of Protein
A4HTH1	Inositol phosphosphingolipid phospholipase C-Like	653	1	1	1112.54	2.28	0.54	Plasma membrane
Q4Q8L3	Leishmanolysin (EC 3.4.24.36)	566	1	1	1206.615	2.37	0.58	Plasma membrane
E9ATU8	VHS domain-containing protein	508	1	1	1625.886	3.04	0.67	Secreted
Q4QCE4	Putative ATP-binding cassette protein subfamily E, member 1	659	1	1	1247.587	2.71	0.47 No antigenicity	Plasma membrane, variable locations
E9AM02	Putative lipophosphoglycan biosynthetic protein	771	1	1	-	-	0.65	Extracellular – endoplasmic reticulum
Q4QI04	VTC domain-containing protein	813	1	1	1394.782	1.95	0.56	Vacuolar transport, plasma membrane
A4I752	Putative lipase	367	1	1	1224.568	2.35	0.35	membrane
Q5EXB3	Phosphotransferase (EC 2.7.1.-)	471	8	9	1272.691	3.78	0.45	Membrane, secreted
Q4QAH5	Uncharacterized Protein	600	1	2	2191.122	3.22	0.49	Cilium
Q306G5	Prohibitin	268	1	1	1115.5	1.8	0.50	Secreted,mitochondrial inner membrane

The putative lipophosphoglycan biosynthetic protein has been identified with different peptides for *L. Tropica* and *L.Donovani*. The PSM ratio of all identified peptides is one.

Lipases (EC 3.1.1.-); They are enzymes that catalyze the hydrolysis of ester bonds of lipids. Leishmania contains abundant sphingolipids and phospholipids on its surface. These lipids play a role in the host-pathogen relationship, in the function of proteins and receptors necessary for cell entry. In addition, leishmania is found in vacuoles within the host where lipases mediate lipid exchange across the vacuole membrane. However, there is more than one lipase enzyme. In the search for potential biomarker proteins that can be used in the diagnostic kit, the first target-specific, rapid and virulence-like proteins are expected to be identified.

Sphingomyelin is a vital membrane component for cell-to-cell recognition and signaling. The Inositol phosphosphingolipid phospholipase C-Like protein (A4HTH1) is a member of the neutral sphingomyelinase enzyme. This protein, expressed in some yeast and parasite species, is not expressed in humans. Leishmania is responsible for the degradation of host sphingomyelin and inositol phosphoryl ceramide. Inositol phosphoryl ceramide, is expressed by leishmania instead of sphingomyelin. In the absence of inositol phosphosphingolipid phospholipase C-like (ISCL), the parasite has become hypersensitive to the acidic environment in the host. ISCL is a plasma membrane protein identified by shaving; It is a candidate biomarker in that it has a significant difference in the amino acid sequence and breaks down the sphingomyelins in the cell membrane and has a virulence effect.[75-76-77]

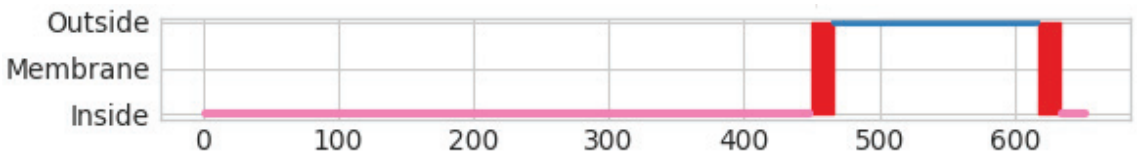


Figure 4.9. ISCL protein

This protein has been identified in *L. Tropica* and *L.Donovani*. In other words, both visceral and cutaneous agents have been identified in the parasite. In Figure 4.9, the membrane localization of the ISCL protein is shown. The amino acid sequence obtained in *L. Tropica* is N terminal IYDMAAATTR' C terminal and the amino acid sequence obtained in *L.Donovani* is 'N terminal IYDMAAATTR'C terminal. This peptide sequence is in the 484-493 amino acid range of the ISCL protein. This area is located in the extracellular region.

ISCL protein, belongs to the neutral sphingomyelinase protein family, and is not expressed in humans. This protein is expressed in some yeasts, parasites, and bacteria. As a result of the alignment of ISCL protein with Homo Sapiens, Schizosaccharomyces Pombe, Leishmania, and Leptomonas, 21.45%, 25.07%, 100%, and 58.58% similarity was found, respectively. Although ISCL and nSMase are homologous proteins, it is understood that they have few similarities in terms of the amino acid sequence. Although its Xcorr value is below three and its PSM is one, it is

suggested that ISCL protein may be a false discovery protein. However, during the experiment, the ISCL protein was defined as a similar protein for at least two species in two different biological Leishmania samples.

The ISCL protein contains the exonuclease-endonuclease-phosphatase domain. This domain contains these conserved domains of the ISCL protein containing the amino acid sequence 14-191.Evolutionarily conserved amino acid domains are effective in protein recognition and structural properties. The conserved domain of the ISCL protein is located in the cytoplasmic region of the plasma membrane.

Although the antigenicity of the VHS protein (E9ATU8) is high, it has not been evaluated as a b-cell epitope peptide because it does not have a transmembrane domain. Prohibitin (Q306G5) is a mitochondrial inner membrane protein. It is described in the literature as secreted on the cell surface. The fact that it is not associated with the plasma membrane has a PSM value of one, and has a low Xcorr score suggests that this protein is a random match. Since it contains a putative lipophosphoglycan biosynthetic protein (E9AM02) signal domain, will likely to be released into the extracellular region. Despite its high antigenicity, it has not been examined in detail because of the absence of the transmembrane domain and the fact that the peptide and PSM matching is one of the results.

L. Major (REF2) and *L. Infantum* (REF3) similar identified protein number 21. Among these proteins, the protein on the surface or is thought to be on the surface is also given in Table 4.10. Three proteins out of 21 proteins as surface proteins were identified.

Tablo 4.10. Common Proteins Identified in *L. Major* and *L . Infantum*

UniProt KB Code	**Identified Protein**	**Length**	**Peptide**	**PSM**	**Theoretical MH+ (Da)**	**Xcorr**	**Probably Antigenic Score (x>0.5)**	**Location of Protein**
A0A504X8M0	Cation diffusion facilitator transporter family protein	1183	1	1	1259.659	2.84	0.57	Membrane
A4HZB2	Phosphotransferase (EC 2.7.1.-)	471	6	6	1912.968	3.37	0.46	Membrane bound organelle
E9B1S0	H(+)-exporting diphosphatase (EC 7.1.3.1)	802	1	1	1731.876	4.11	0.46	Membrane

The similar protein number identified in *L.Major* (REF2) and *L.Donovani* (REF4) is 13. Among these proteins, the protein on the surface or is thought to be on the surface is also given in the Table 4.11. One protein out of 13 proteins as surface protein was identified.

Tablo 4.11. Common Proteins Identified in *L. Major* and *L.Donovani*

UniProtKB Code	**Identified Protein**	**Length**	**Peptide**	**PSM**	**Theoretical MH+(Da)**	**X Corr**	**Probably Antigenic Score (x>0.5)**	**Location of Protein**
E9AD26	Putative calpain-like cysteine peptidase (EC 3.4.22.-)	4553	3	3	1080.551	2.71	0.63	Cytoplasm, Plasma membrane

L.Infantum (REF3) and *L. Donovani* (REF4) similarly identified protein number 76. Among these proteins, the protein on the surface or is thought to be on the surface is also given in Table 4.12. Five proteins out of 76 proteins as surface proteins were identified.

Tablo 4.12. Common Proteins Identified in *L. Infantum* and *L. donovani*

UniProt KB Code	**Identified Protein**	**Length**	**Peptide**	**PSM**	**Theoretical MH+(Da)**	**X Corr**	**Probably Antigenic Score (x>0.5)**	**Location of Protein**
A4HRH2	Long-chain-fatty-acid-CoA ligase, putative	698	5	2	1250.591	2.7	0.42 No antigenicity	Plasma membrane
E9AGG4	Soluble promastigote surface antigen PSA-50S /PSA-2	463	1	1	1291.7	2.48	0.59	Plasma membrane
A0A3S7X1J0	Conserved plasma membrane protein	138	1	1	1098.569	2.56	0.38	Plasma membrane
E9BIZ3	Stress-response A/B barrel domain-containing protein	110	1	1	1690.828	2.48	0.79	Extracellular, cytoplasm
A4IAI3	Uncharacterized protein	274	1	1	1660.865	2.86	0.53	ER, plasma membrane

Promastigote surface protein-2 (PSA-2), is one of the proteins identified in *L.Infantum* and *L. Donovani*. (PSA-2); moreover it is a membrane-associated, glycolipid-bearing secreted protein. PSA-2; It contains leucine repeated region. The leucine-rich region effectively recognizes the parasite, signal transduction, and protein-protein interactions. The leucine-rich repeat protein (LRRP) domain was in four Leishmania species. However, due to the alignment of the same protein LRRP with two different UniProt entry names (A0A504XHY0, Q4QH36) with the amino acid sequence, they were found to have 21.58% similarities. From this, it is understood that even if the catalytic activities or functions of the proteins are similar, the same protein obtained from different species is not in the same amino acid sequence. Proteins can undergo mutations according to their host or parasite. The PSA-2 protein is a specific protein to the leishmania family.[78-79] It is a transmembrane protein containing a signal domain in the 1-32 amino acid range. The peptide range defined for both *L. Infantum* and *L. Donovani* is N'TNTLAVLEAFGR'C. Epitope regions in the protein were predicted with ABCpred prediction. The epitope probability of the N'LAVLEAFGRAIPELGK'C peptide, which also includes the identified peptide domain, is 0.68. No potential allergy was found with AllerCatPro 2.0.

Although the conserved plasma membrane protein (A0A3S7X1J0) is crucial because in that it contains conserved areas, it has yet to be examined because its antigenicity is below the threshold. While Wolf Psort II for Uncharacterized protein (A4IAI3) was found in plasma membrane localization in the scan with animals and fungi, DeepTMHMM was not found to be associated with the membrane in the transmembrane region scan.

L.Tropica (REF1), *L.Major* (REF2), and *L. Infantum* (REF3) similarly identified protein number is 10. Among these proteins, the protein on the surface or is thought to be on the surface is also given in Table 4.13. Two proteins out of 10 proteins as surface proteins were identified.

Tablo 4.13. Common Proteins Identified in *L.Tropica*, *L.Major*, and *L. Infantum*

UniProtKB Code	Identified Protein	Length	Peptide	PSM	Theoretical MH+(Da)	X Corr	Probably Antigenic Score (x>0.5)	Location of Protein
E9ASW8	Putative nucleoside transporter 1	499	1	1	1694.743	3.15	0.50	Plasma membrane
A0A504Y8T0	Peptidase M3 family protein	1196	1	1	1843.94	3.63	0.42	The plasma membrane, cytoplasm

Like many parasites, transport proteins are of great importance in Leishmania. It has essential roles in transport by transmembrane proteins, controlled uptake and release of lipophilic components, the release of signal molecules, and nutrient exchange. To survive Leishmania needs at least one purine pyrimidine transporter, such as the nucleoside transporter protein. It is not known precisely how the nucleoside transporter protein affects Leishmania.[80]

The number of *L.Tropica* (REF1), *L.Major* (REF2), and *L.Donovani* (REF4) common proteins is 13. No surface proteins were identified.

The number of *L.Tropica* (REF1), *L.Infantum* (REF3), and *L.Donovani* (REF4) common proteins is 235. Among these proteins, the protein on the surface or is thought to be on the surface is also given in Table 4.14. Seventeen proteins out of 235 proteins as surface proteins were identified.

Tablo 4.14. Common Proteins Identified in *L.Tropica, L.Infantum,* and *L. donovani*

UniProtKB Code	Identified Protein	Length	Peptide	PSM	Theoretical MH+(Da)	X Corr	Probably Antigenic Score (x>0.5)	Location of Protein
E9B9H0	Amastin surface glycofamily protein	300	1	1	2244.135	3.22	0.60	Plasma membrane
A0A6L0XSK0	Kinetoplastid membrane protein-11	92	2	4	1105.589	2.52	0.56	Plasma membrane - extracellular

(Cont. on next page)

Table 4.14. (cont)

UniProtKB Code	Identified Protein	Length	Peptide	PSM	Theoretical MH+(Da)	X Corr	Probably Antigenic Score (x>0.5)	Location of Protein
A0A6L0WJR8	Leishmanolysin (EC 3.4.24.36)	151	2	2	1110.536	2.28	0.80	Plasma membrane-secreted
Q94593	Leishmanolysin (EC 3.4.24.36)	617	1	1	2869.294	5.83	0.80	Plasma membrane-secreted
E9AUL9	Vesicle-fusing ATPase (EC 3.6.4.6)	738	3	3	1512.827	3.04	0.31	Cytoplasm-secreted
A0A3S7X9N0	Cys/Met metabolism PLP-dependent enzyme family protein	409	3	4	2264.311	3.52	0.40	Secreted
E9ARJ9	Putative ATP-dependent zinc metallopeptidase	598	1	2	1842.996	3.39	0.59	Plasma membrane
Q9NDD1	Surface antigen P2 (PSA-2)	626	2	2	1762.742	4.07	0.58	Extracellular /plasma membrane
A0A3Q8IHK1	AP-1 complex subunit gamma	831	2	2	2844.459	3.77	0.53	Cytoplasmic vesicle, membrane
A4HTF0	Putative vacuolar-type Ca2+-ATPase (EC 3.6.3.-)	929	3	3	1377.701	2.61	0.53	Plasma membrane
Q71S90	Iron Superoxide dismutase (EC 1.15.1.1)	208	1	1	1333.696	3.29	0.53	Secreted
A2TEF2	Putative lanosterol 14-alpha-demethylase (EC 1.14.13.70)	480	1	1	1583.9	3.17	0.66	The plasma membrane, endoplasmic reticulum, mitochondrial
A0A6L0XPB6	Transmembrane 9 superfamily member	637	1	1	1273.581	2.61	0.58	Membrane
A0A6L0XHZ8 */L.Tropica, L.Infantum*	TPR-repeat protein putative	488	1	1	1240.726	2.1	0.53	Variable locations

(Cont. on next page)

Table 4.14. (cont)

UniProtKB Code	Identified Protein	Length	Peptide	PSM	Theoretical MH+(Da)	X Corr	Probably Antigenic Score (x>0.5)	Location of Protein
E9AK23/ *L.Donovani*	TPR_REGION domain-containing protein	361	2	2	1855.976	3.38	0.59	Variable locations
A4I0B5	Uncharacterized protein	556	3	3	1352.615	2.59	0.64	Membrane
Q4Q3G7	Uncharacterized protein	420	1	1	1752.814	3.87	0.62	Secreted
A0A504Y241	Putative integral membrane protein	1601		2	2362.094	6.15	0.68	Plasma membrane

In the cell's normal functioning, some proteins (gp63, PSA-2) are secreted out of the cell, that is, into the extracellular region. Secreted proteins out of the cell; It is carried out in two ways, classical and non-classical. ER-Golgi organelles are carried out classically, and proteins are released from the cell with the Golgi vesicles. In its classically secreted pathway, the protein contains a signal peptide domain. This domain is rich in hydrophobic amino acids (like Lysine-R). In the non-classical secretory pathway, proteins are secreted without the ER-Golgi complex and the signal peptide domain. Examples are the proteins secreted from the lysosome and those emitted from the exosome. Secreted proteins are important biomarker proteins in host and parasite interaction in terms of their antigenic properties.[81]

Leishmanolysin (A0A6L0WJR8) is seen in *L.Tropica*, *L.Infantum*, and *L.Donovani*. Leishmanolysin can be found in both forms. It may also be secreted outside the cell during interaction with the host or a protein bound to the extracellular region. This protein is found in visceral and cutaneous parasites; Promastigote is a zinc-dependent metalloproteinase expressed in Leishmania. This protein, which molecularly weighs 63 kilodaltons, plays an vital role in the interaction between the parasite and the host. For example; the interaction of leishmanolysin with fibronectin receptors, facilitates the adhesion of the parasite to macrophages. It helps the parasite survive by acting on the host macrophage. Leishmanolysin is a candidate biomarker protein

because it is a specific protein for the leishmania family and also has high antigenicity.[82-83]

Kinetoplastid membrane protein -11(A0A6L0XSK0) is a surface glycoprotein unique to the Trypasonoma family. It is crucial for the stability of lipophosphoglycan.[85]The expression level of kinetoplastid membrane protein is increased in metacyclic promastigote and amastigote forms. It is essential as a surface protein for the parasite to enter the cell and internalize in macrophages. While increasing the production of interleukin-1O and arginase in the immune system; suppresses nitric oxide. In this way, the parasite survives in the host macrophage. Both its high antigenicity and its specificity to the leishmania family make the peptides of this protein essential for B Cell epitope peptide regions.[84]

The exact function of amastin surface glycoprotein protein (E9B9H0 has yet to be clarified, and it is assumed to be involved in signal transmission between the host and the parasite. It is known that amastin surface glycofamily proteins may play a role in ion transfer in the cell membrane and are effective in host-parasite interactions.[84]In previous studies of this protein, N-terminal peptides' antigenicity was higher than many glycoproteins. It is generally used in the diagnosis of visceral leishmaniasis patients. In contrast, the C terminal peptide domain was analyzed in our results. It was determined that the prediction peptide ratio for the epitope of this peptide domain defined in N'RSSEAGLGESAVALLQGEDERLL'C was 0.72. The threshold is antigenically restricted for above 0.5. However, predictions of epitope peptides that will provide specific binding with antibodies do not only pass antigenicity test; this is just the first step and requires repetition.

Putative vacuolar-type Ca+2 ATPases (A4HTF0) levels provide calcium homeostasis in the plasma membrane and provide signal transmission. Some proteins in the plasma membrane (such as phosphoinositide phospholipase C) use calcium as a substrate. In the interaction of parasite and host, the activity of calcium ATPases increases, and calcium is overexpressed in the parasite environment. It has been observed that this facilitates entry into tissues.[87]

The lanosterol 14α-demethylase enzyme (EC 1.14.13.70)(A2TEF2) is the animal isoenzyme of cytochrome p450. This enzyme catalyzes the removal of the C-14alpha group, which is vital in sterol biosynthesis, from lanosterol. This demethylation step is the first checkpoint in converting of lanosterol to other essential sterols within the cell. According to studies, leishmania needs to make its own sterol biosynthesis,

since it cannot use the cholesterol of the host cell.[88-89] It has high antigenic properties as a candidate biomarker in diagnosis kit development. Still, since it is analyzed with a single peptide and a single peptide spectrum match score, the results must be repeated. Whether the Transmembrane 9 superfamily member (A0A6L0XPB6) identified in the samples obtained from *L.Tropica, L.Infantum,* and *L. donovani* contains an alpha helix transmembrane domain was checked with the transmembrane domain determination page of Denmark Technical University[91].

However, the fact that it contains only a transmembrane domain does not prove that the protein is on the surface. This protein can also be found in the membranes of intracellular organelles such as mitochondria and endoplasmic reticulum. Therefore, the locations of uncharacterized proteins, and transmembrane proteins were additionally found with Wolf Psort II and Gene Ontology (GO).

Although the exact function of the Uncharacterized protein (A4I0B5) identified in *L.Tropica, L. infantum, and L.Donovani* is unknown, it is likely to be an immunogenic protein for future studies. Because; It is a transmembrane protein and indicates that the shaving method is working, it contains a signal domain, so signal transduction is likely to be involved in the interaction of the host with the parasite. In addition, the protein comprises highly antigenic peptide domains. It includes a signal domain in the amino acid region at positions 1-32, and amino acids at amino acid positions 5-32 have highly antigenic properties. The conserved domain has yet to be defined yet. Any allergenicity that the uncharacterized protein may cause in the body could not be detected.

Protein kinase-like proteins are enzymes that change their function or stop by phosphorylating the protein. During this process, 1 ATP enters, and 1 ADP exits. Many protein kinases are involved in the biological cycle of Leishmania. The activity of the protein is excellent in differentiation, replication, signal transduction, and host survival. For example, protein kinase A dephosphorylated some proteins in the phosphorylated state when differentiation from the promastigote form to the amastigote form begins. Therefore, many therapeutic drugs are being studied to inactivate protein kinases. The protein kinase was found to consist of 9 alpha-helical transmembrane domains using the TMHMM 2.0 server.[91] The same peptide sequence was analyzed for four species due to LC-MS/MS. This sequence (N terminal-LDEAVDLSTRGDEDAAK-C terminal) is the protein's 764 to 781 amino acid range with 1899 amino acids. Figure 4.10 shows this

area (764-781 amino acid sequence) in the cytoplasm inside the cell. Its appearance in the results is its mobility in the plasma membrane.

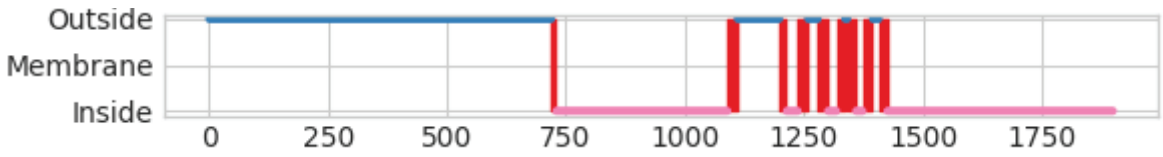


Figure 4.10. Protein Kinase-Like Proteins

The number of *L.Major* (REF2), *L.Infantum* (REF3), and *L.Donovani* (REF4) similar proteins are 21. No surface protein was detected.

The number of *L.Tropica* (REF1), *L.Major* (REF2), *L.Infantum* (REF3), and *L.Donovani* (REF4) similar proteins are 129. Among these proteins, the protein on the surface or is thought to be on the surface is also given in Table 4.15. Ten protein out of 129 proteins as surface proteins were identified.

Tablo 4.15. Common Proteins Identified in *L.Tropica, L.Major, L.Infantum,* and *L. donovani*

UniProt KB Code	**Identified Protein**	**Length**	**Peptide**	**PSM**	**Theoratical MH+(Da)**	**X corr**	**Probably Antigenic Score (x>0.5)**	**Location of Protein**
A0A381MFJ0	Plasma membrane ATPase (EC 7.1.2.1)	974	2	4	1423.729	3.95	0.50	Plasma membrane
A0A504X036	H(+)-exporting diphosphatase (EC 7.1.3.1)	1649	6	7	1491.729	3.55	0.54	Cilium, membrane
E8NHQ7	Putative ATPase alpha subunit	574	5	7	-	-	0.45	Variable locations
A0A504XVT2	Vacuolar (H+)-ATPase G subunit family protein	320	1	1	1614.747	3.9	0.51	Plasma membrane

(Cont.on next page)

Table 4.15. (cont.)

UniProtKB Code	Identified Protein	Length	Peptide	PSM	Theoratical MH+(Da)	Xcorr	Probably Antigenic Score (x>0.5)	Location of Protein
E9AD27 *(L.Tropica, L.Major, L.Infantum, L.Donovani*	Putative calpain-like cysteine peptidase	821	3	3	2246.16	4.07	0.48	Cytoplasmic-secreted
A4HS39 (*L.Tropica, L.Infantum, L.Donovani)*	Putative calpain-like cysteine peptidase (EC 3.4.22.-)	855	4	4	1696.798	3.36	0.48	Cytoplasmic-secreted
A0A504XEX2 *(L.Major, L.Infantum, L.Donovani)*		6170	40	47	2748.447	4.75	0.48	
E9ADF0	BAR domain-containing protein	340	3	4	1484.712	2.41	0.60	Ciliary membrane, plasma membrane
E9AKI4	Uncharacterized protein	147	2	2	1269.624	2.41	0.60	Membrane
A0A3S7X7R9	Hypothetical protein conserved	352	1	1	1781.903	4.46	0.47	Membrane

Calpain-like cysteine peptidase (EC 3.4.22.52) protein identified in *L.Tropica, L.Major, L.Infantum,* and *L. Donovani*, is a serine peptidase. This protein, frequently encountered in many organisms, is responsible for cytoskeletal re-organization, signal transduction, cell cycle control, differentiation, and controlled and uncontrolled cell death.[69] It has been shown that calpain-like cysteine peptidase inhibitors in Leishmania and Trypanosomatids cause a significant decrease in amastigote levels in the host macrophage. Peptidases want to transport themselves out of the cell. It is likely encountered in the samples because it is found in excess and wants to rise to the cell surface.

Tablo 4.16. Homolog Proteins Identified in *L.Tropica, L.Major*, L.Infantum, and *L. Donovani*

UniProt KB Code	**Identified Protein**	**Length**	**Peptide**	**PSM**	**Theoretical MH+(Da)**	**X corr**	**Probably Antigenic Score (x>0.5)**	**Location of Protein**	**Leishmania Species**
E9AKM8	Surface antigen-like protein	284	2	3	1679.723	3.01 2.58	0.60	Cilium	*L.Tropica, L.Donovani*
E9ACQ0		284	1	1	1328.572	2.99	0.61		*L.Major*
Q9XZY0		284	1	1	1679.723	3.01	0.72		*L.Infantum*
E9AKJ6		286	1	1			0.54		*L.Major, L.Donovani*

As a result of surface antigen-like protein alignment defined in *L.Tropica, L.Major, L.Infantum, and L.Donovani* with different UniProt codes, it was found that E9AKM8 and E9ACQ0 were 91.71% similar, with a high probability of being homologous proteins, but 19.71% similarity with Q9XZY0. Sometimes the same proteins can have different amino acid sequences and perform the same function. Mutations occur naturally and adjust the protein to host requirements. The peptide domain N' AMDAGTVSNCFR'C identified in *L.Tropica, L.Donovani*, and *L.Major* is in the conserved region. Conserved regions often have functional features. Therefore, it may be significant that the epitope regions are in these parts. Neverthless, it is a peptide domain with low antigenicity. Although their exact biochemical function has not been defined, their containing signal domains, signal transduction, and presence on the surface may be necessary for antigen-antibody research in future studies.

Tablo 4. 17. Homolog V –Type Proton ATPase

UniProt KB Code	**Identified Protein**	**Length**	**Peptide**	**PSM**	**Theoretical MH+(Da)**	**X corr**	**Probably Antigenic Score (x>0.5)**	**Location of Protein**
E9AWI3 *(L.Tropica L.Donovani)*	V-type proton ATPase subunit A	775	1	1	1211.674	2.27	0.44- No antigenicity	Lysosome/ vacuole/ Plasma membrane
Q4Q821 *(L.Tropica, L.Donovani)*	Vacuolar proton pump subunit B (V-ATPase subunit B)	495	7	7	1435.692	3.85	0.28-No antigenicity	Membrane
E9ARJ3 (*L.Tropica, L.Infantum)*	V-type proton ATPase subunit C	380	1	1	1520.707	3.83	0.50	Membrane
A4I3X4 *(L.Major)*	Vacuolar proton pump subunit B	497	7	7				Acidocalcisome, membrane
Q4Q82 *L.Donovani*	Vacuolar proton pump subunit B	495	7	7	1610.926	2.08	0.28	Acidocalcisome, membrane

Vacuolar H+-ATPase (V-ATPase) is a membrane-associated protein that pumps the proton from the cytoplasm toward the organelles. The parasitophorous vacuole lumen of the parasite must remain acidic for Leishmania to survive in the macrophage. During phagolysosomal internalization, its pH decreases compared to its form outside the host cell (the promastigote form of Leishmania has higher pH and lower temperature) while its temperature increases.[90] V-ATPase activity is important for providing and maintaining this acidity. However, its antigenicity is below the determined threshold.

Proteins such as V-type ATPases, PSA-2, and surface antigen-like proteins were identified with different UniProt codes. Homologous proteins can undergo some mutations according to the parasite's need and create genetic and amino acid-based changes such as single nucleotide polymorphism. When these proteins were aligned within themselves, it was understood that the amino acid sequences were different.

4.2.2.Results of Samples Using C18 Beads Attachment Method

C18 bead attachment (C18-Solid phase extraction) was also applied to determine whether there was sample loss with acetone while acetone precipitation was used. With the C18 bead attachment samples, the number of proteins identified in *L.Tropica, L.Infantum, and L.donovani* decreased while *L.major* increased (Figure 4.11). The insufficient growth of *L. Major* and the identification of many proteins compared to the acetone-precipitated samples could not be explained. Perhaps there was too much hydrophobic peptide in the throw-away acetone due to acetone precipitation, and these peptides allowed the identification of samples attached to C18 beads. However, the situation is also strange for *L.Tropica, L.Infantum, and L.Donavani*. While the number of proteins identified in the acetone-precipitated samples is around 700, it is around 300 in the C18 bead attachment samples. Since the results were not straightforward, acetone precipitation was thought to be more reliable. The analysis with homemade C18 beads attached in micropipettes should be repeated and optimized.

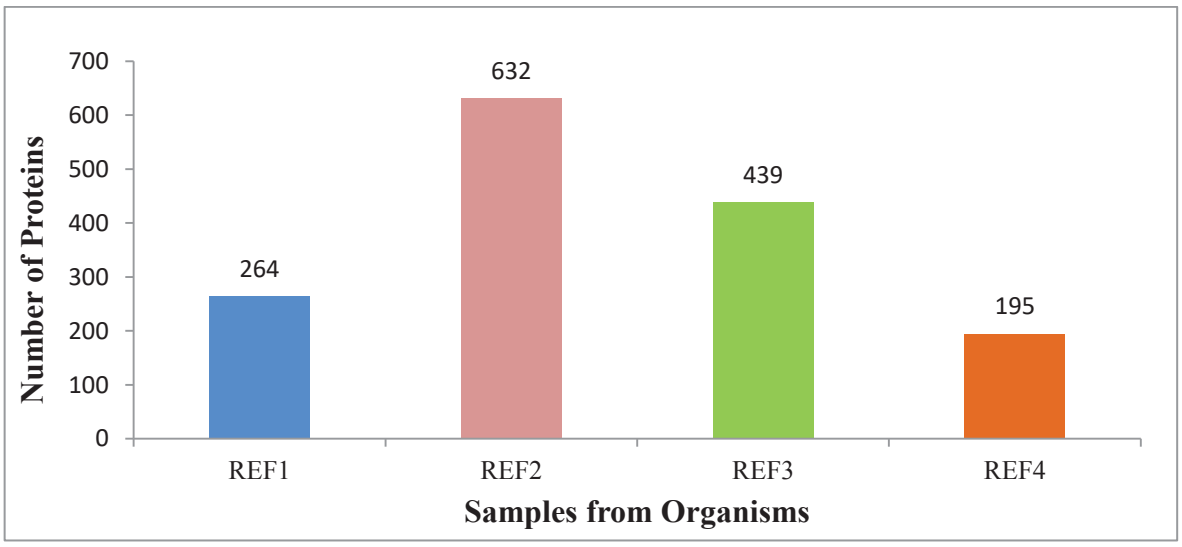


Figure 4.11. Protein Counts Defined in *L.Tropica, L.Major, L.Infantum, L.Donavani*

CHAPTER 5

CONCLUSION

Leishmaniasis is a neglected tropical disease that has become more and more deadly with recent migrations to Türkiye. There is no rapid and effective diagnostic kit for this disease caused by the leishmania parasite. Surface proteins are the target of biomarker and diagnostic kit studies, as they are directly effective in the first interaction of the parasite and the host cell, facilitating penetration and invasion within the host. Although surface proteins are very vital in many biomarker detection studies, they are challenging to study and identify because they are rich in hydrophobic amino acids and embedded in the phospholipid bilayer. This study aimed to challenge the difficulty of identifying plasma membrane-associated proteins by applying the shaving method in Leishmania. With a short 30-minute trypsin treatment, proteins on the cell surface were cut and enzyme activity was stopped. The shaving method was applied to *L.Tropica, L.Major, L.Infantum, and L.Donovani* species. Different and similar proteins were identified between species.

In the LC-MS/MS analysis results; 717 proteins from *L.Tropica*, 278 from *L.Major*, 619 from *L.Infantum*, and 718 from *L.Donovani* were identified. Among the 717 proteins identified in *L.Tropica*, 67 proteins associated with the plasma membrane and extracellular were found. It was found that 21 of the 278 proteins identified in *L. Major* were associated with the membrane. 49 of the 619 proteins identified in L. Infantum are surface-associated or secreted proteins. It was found that 54 of the 718 proteins identified in L.Donovani were associated with the surface. The defined surface protein ratio was 9.34%, 7.55%, 7.9%, and 7.52% for *L.Ttropica, L.Major, L.Infantum*, and *L.Donovani*, respectively.

Transport proteins such as ABC transporter, V-type proton ATPase, Leishmanolysin, calpain cysteine protease, KMP-11, ISCL, PSA-2, lanosterol 14α demethylase, surface antigen-like protein in cell surface and parasite-host interaction important literature compatible proteins were detected. Among the proteins identified in *L. Major and L. Infantum*, inositol phosphosphingolipid phospholipase C-Like protein

(ISCL) has a neutral sphingomyelinase protein, which is homologous to humans. Although the catalytic activity of the enzymes is similar, the similarity in terms of an amino acid sequence is 21.45%. This difference may be an essential convenience for diagnosis kits in future studies. It is crucial for a potential biomarker that it is associated with the plasma membrane, has a virulence effect, and has an antigenic score above 0.5. In addition, kinetoplastid membrane protein-11, leishmanolysin *L.Tropica, L.Infantum*, and *L.Donovani*, have high antigenicity and leishmania family-specific proteins, were identified. The reason *why L. Major* was not identified because *L. Major* did not grow sufficiently during cell culture studies. The predicted B cell epitope peptide regions of these proteins may be preferred in future studies in diagnosis kits.

Many uncharacterized proteins were identified. Subcellular localization of these proteins was found bioinformatically. Among these proteins, the structure of the uncharacterized protein (A4I0B5) is unknown; However, it it was thought to be effective in signal transduction between the host and the parasite since it is located in the plasma membrane and contains a signal domain. This protein with high antigenic properties was identified by at least 3 peptides in *L.Tropica, L.Infantum, and L.Donovani*. The results found proteins generally found in the nucleus and ribosome such as histone H2A protein and elongation factor 1-alpha. As it can be understood, the cell initiates apoptosis and performs controlled cell death, or there is a vesicular transfer from the cell to the outside of the cell.

Despite the difficulties in analyzing surface proteins, it has been shown that surface proteins are identified quickly and practically with the shaving method. A few steps can be followed in future studies to increase the amount of defined surface protein and to make more different protein biomarker targets. For example, although trypsin is the most compatible enzyme with MS and proteome discoverer, it can be aimed to identify uncharacterized proteins by trying other enzymes by cutting the protein from different regions (such as pepsin). In addition, since small molecular weight peptides are lost with acetone precipitation, either the acetone precipitation method should be improved. Some methods that can be tried in future studies are; Evaporation of acetone after acetone precipitation, mixing acetone with salts to increase the salting-out effect[92] (to increase the precipitation ratio of the peptide), or other separation methods that may be preferred.

REFERENCES

1) Spickler, A. R. Leishmaniasis (Cutaneous and Visceral). Iowa State University **2017**, 1–11**.**

2)Arenas, R.; Torres-Guerrero, E.; Quintanilla-Cedillo, M. R.; Ruiz-Esmenjaud, J. Leishmaniasis: A Review. F1000Res **2017**, 6 (May), 1–15. https://doi.org/10.12688/f1000research.11120.1.

3) Hotez, P. J.; Aksoy, S.; Brindley, P. J.; Kamhawi, S. World Neglected Tropical Diseases Day. PLoS Negl Trop Dis **2020**, 14 (1), 1–4. https://doi.org/10.1371/JOURNAL.PNTD.0007999.

4) Özsoylu, Ş. Treatment of Visceral Leishmaniasis. Turkish Journal of Pediatrics. July **2003**, p 280. https://doi.org/10.4103/0974-777x.62883.

5) Reimão, J. Q.; Coser, E. M.; Lee, M. R.; Coelho, A. C. Laboratory Diagnosis of Cutaneous and Visceral Leishmaniasis: Current and Future Methods. Microorganisms **2020**, 8 (11), 1–30. https://doi.org/10.3390/microorganisms8111632.

6) Tasew Guma, Geremew. The Contribution of Epidermal Apoptosis to the Clinical Outcome of Leishmania Aethiopica Induced Cutaneous Leishmaniasis and Its Laboratory Diagnosis Using Low Cost Culture Media; Karolinska Institutet, **2010**.

7) Jean Hamilton, H. The Leish Niche: The Secretome of Leishmania and Its Role in Parasite Virulence. **2019**

8) Lypaczewski, P. Virulence and Adaptation of Leishmania Donovani Parasites; **2020**.

9) Verma, S.; Deep, D. K.; Gautam, P.; Singh, R.; Salotra, P. Proteomic Analysis of Leishmania Donovani Membrane Components Reveals the Role of Activated Protein C Kinase in Host-Parasite Interaction. Pathogens **2021**, 10 (9). https://doi.org/10.3390/pathogens10091194.

10) McKenna, S. E. Functional and Structural Studies of the Hydrophilic Acylated Surface Proteins from Leishmania Donovani. University of York **2015**, No. August..

11) World Health Organizastion ,Leishmaniasis. https://www.who.int/news-room/fact-sheets/detail/leishmaniasis.

12) Pan American Health Organisation, Visceral Leishmaniasis. https://www.paho.org/en/topics/leishmaniasis/visceral-leishmaniasis

13) Siriwardana, Y.; Deepachandi, B.; Gunasekara, C.; Warnasooriya, W.; Karunaweera, N. D. Leishmania Donovani Induced Cutaneous Leishmaniasis: An Insight into Atypical Clinical Variants in Sri Lanka. J Trop Med 2019, **2019**. https://doi.org/10.1155/2019/4538597.

14) Elfaituri, S. S.; Matoug, I.; Elsalheen, H.; Belrasali, Y.; Emaetig, F. Mucocutaneous Leishmaniasis in an 11-Year-Old Girl with Ataxia Telangectasia - Case Report. Libyan Journal of Medicine **2015**, 10. https://doi.org/10.3402/ljm.v10.26432.

15) Asmaa, Q.; AL-Shamerii, S.; Al-Tag, M.; AL-Shamerii, A.; Li, Y.; Osman, B. H. Parasitological and Biochemical Studies on Cutaneous Leishmaniasis in Shara'b District, Taiz, Yemen. Ann Clin Microbiol Antimicrob **2017**, 16 (1), 1–15. https://doi.org/10.1186/s12941-017-0224-y.

16) Handman, E. Leishmaniasis: Current Status of Vaccine Development. Clin Microbiol Rev **2001**, 14 (2), 229–243. https://doi.org/10.1128/CMR.14.2.229-243.2001.

17) Rajawat, J.; Jhingan, G. Mass Spectroscopy; Elsevier Inc., 2019. https://doi.org/10.1016/B978-0-12-816548-5.00001-0.

18) El-Aneed, A.; Cohen, A.; Banoub, J. Mass Spectrometry, Review of the Basics: Electrospray, MALDI, and Commonly Used Mass Analyzers. Appl Spectrosc Rev **2009**, 44 (3), 210–230. https://doi.org/10.1080/05704920902717872.

19) Maity, A.; Kesh, S. S.; Palai, S.; Egbuna, C. Electrophoretic Techniques; **2022**. https://doi.org/10.1016/B978-0-12-822654-4.00007-5.

20)Ribeiro, D. A Short Overview of the Components in Mass Spectrometry Instrumentation for Proteomics Analyses. Tandem Mass Spectrometry - Molecular Characterization **2013**, 1. https://doi.org/10.5772/54484.

21) Banerjee, S.; Mazumdar, S. Electrospray Ionization Mass Spectrometry: A Technique to Access the Information beyond the Molecular Weight of the Analyte. Int J Anal Chem **2012**, 2012 (January), 1–40. https://doi.org/10.1155/2012/282574.

22) Šebela, M. Solid Mixed Matrices and Their Advantages in Matrix-Assisted Laser Desorption/ Ionisation Time-of-Flight Mass Spectrometry; 2016; Vol. 28. www.spectroscopyeurope.com.

23)Karpievitch, Y. V.; Polpitiya, A. D.; Anderson, G. A.; Smith, R. D.; Dabney, A. R. LIQUID CHROMATOGRAPHY MASS SPECTROMETRY-BASED PROTEOMICS: BIOLOGICAL AND TECHNOLOGICAL ASPECTS. Annals of Applied Statistics **2010**, 4 (4), 1797–1823. https://doi.org/10.1214/10-AOAS341.

24) Submitted, T. Leishmania Tropica CAUSING VISCEROTROPISM IN HUMANS USING MASS SPECTROMETRY A Thesis Submitted To. **2022**, No. July.

25) Ahmed, F. E. Utility of Mass Spectrometry for Proteome Analysis: Part I. Conceptual and Experimental Approaches. Expert Rev Proteomics **2008**, 5 (6), 841–864. https://doi.org/10.1586/14789450.5.6.841.

26) Hu, Q.; Noll, R. J.; Li, H.; Makarov, A.; Hardman, M.; Cooks, R. G. The Orbitrap: A New Mass Spectrometer. Journal of Mass Spectrometry **2005**, 40 (4), 430–443. https://doi.org/10.1002/jms.856.

27) Eliuk, S.; Makarov, A. Evolution of Orbitrap Mass Spectrometry Instrumentation. Annual Review of Analytical Chemistry **2015**, 8, 61–80. https://doi.org/10.1146/annurev-anchem-071114-040325.

28) Kinter, M. Mass Spectrometry. Anal Chem **1995**, 67 (12), 493–497. https://doi.org/10.1021/ac00108a034.

29) Ladislas Wiza, J. Microchannel Plate Detectors. Nuclear Instruments and Methods **1979**, 162 (1–3), 587–601. https://doi.org/10.1016/0029-554X(79)90734-1

30) Parker CE, Warren MR, Mocanu V. Mass Spectrometry for Proteomics. In: Alzate O, editor. Neuroproteomics. Boca Raton (FL): CRC Press/Taylor & Francis; **2010**. Chapter 5. Available from: https://www.ncbi.nlm.nih.gov/books/NBK56011/

31) Pace, B. S. The Human Genome Project. Renaissance of Sickle Cell Disease Research in the Genome Era **2007**, 27–44. https://doi.org/10.1142/9781860947964_0003.

32) Capelli-Peixoto, J.; Mule, S. N.; Tano, F. T.; Palmisano, G.; Stolf, B. S. Proteomics and Leishmaniasis: Potential Clinical Applications. Proteomics Clin Appl **2019**, 13 (6), 1–16. https://doi.org/10.1002/prca.201800136.

33) Al-Amrani, S.; Al-Jabri, Z.; Al-Zaabi, A.; Alshekaili, J.; Al-Khabori, M. Proteomics: Concepts and Applications in Human Medicine. World J Biol Chem **2021**, 12 (5), 57–69. https://doi.org/10.4331/wjbc.v12.i5.57.

34)Valcu, C. M.; Kempenaers, B. Proteomics in Behavioral Ecology',. Behavioral Ecology **2015**, 26 (1), 1–15. https://doi.org/10.1093/beheco/aru096.

35) de Castro, M. D. L.; Fernández-Peralbo, M. A. The Role of Microwaves in Omics Disciplines. Food Engineering Series 2013, No. April **2013**, 127–180. https://doi.org/10.1007/978-1-4614-4830-3_6.

36) Gao, Y.; Yates, J. R. Proteomic Analysis by Shotgun Proteomic. Mass Spectrometry-Based Chemical Proteomics **2019**, 1–38. https://doi.org/10.1002/9781118970195.ch1.

37) Biorender application (https://www.biorender.com/)

38)Abdallah, C.; Dumas-Gaudot, E.; Renaut, J.; Sergeant, K. Gel-Based and Gel-Free Quantitative Proteomics Approaches at a Glance. Int J Plant Genomics 2012, **2012**. https://doi.org/10.1155/2012/494572.

39) Delahunty, C. M.; Yates, J. R. MudPIT: Multidimensional Protein Identification Technology. Biotechniques **2007**, 43 (5), 563–569. https://doi.org/10.2144/000112604.

40) Gilar, M.; Olivova, P.; Daly, A. E.; Gebler, J. C. Two-Dimensional Separation of Peptides Using RP-RP-HPLC System with Different PH in First and Second Separation Dimensions. J Sep Sci **2005**, 28 (14), 1694–1703. https://doi.org/10.1002/jssc.200500116.

41) Le, J.; Zhongqun, L.; Zhaoyan, W.; Yijun, S.; Yingjin, W.; Yaojie, W.; Yanan, J.; Zhanrong, J.; Chunyang, M.; Fangli, G.; Nan, X.; Lingyun, Z.; Xiumei, W.; Qiong, W.; Xiong, L.; Xiaodan, S. Development of Methods for Detecting the Fate of Mesenchymal Stem Cells Regulated by Bone Bioactive Materials. Bioact Mater **2021**, 6 (3), 613–626. https://doi.org/10.1016/j.bioactmat.2020.08.03.

42) Yang, F.; Shen, Y.; Camp, D. G.; Smith, R. D. High-PH Reversed-Phase Chromatography with Fraction Concatenation for 2D Proteomic Analysis. Expert Rev Proteomics **2012**, 9 (2), 129–134. https://doi.org/10.1586/epr.12.15.

43) Submitted, T. Proteomic Analyses of Biological Samples By Using Different Mass. **2018**, No. September.

44) Boyd, R.; Somogyi, Á. The Mobile Proton Hypothesis in Fragmentation of Protonated Peptides: A Perspective. J Am Soc Mass Spectrom **2010**, 21 (8), 1275–1278. https://doi.org/10.1016/j.jasms.2010.04.017.

45)Winter, D.; Seidler, J.; Hahn, B.; Lehmann, W. D. Structural and Mechanistic Information on C1 Ion Formation in Collision-Induced Fragmentation of Peptides. J Am Soc Mass Spectrom **2010**, 21 (10), 1814–1820. https://doi.org/10.1016/j.jasms.2010.06.020.

46) Kertesz-Farkas, A.; Reiz, B.; P. Myers, M.; Pongor, S. Database Searching in Mass Spectrometry Based Proteomics. Curr Bioinform **2012**, 7 (2), 221–230. https://doi.org/10.2174/157489312800604354.

47) Domselaar, G. Van. Laboratory 2.4: Mascot. University of Alberta **2004**, 1–17.

48) Tabb, D. L.; Eng, J. K.; Yates, J. R. Protein Identification by SEQUEST. 2001, No. October **2017**, 125–142. https://doi.org/10.1007/978-3-642-56895-4_7.

49) Leth-Larsen, R.; Lund, R. R.; Ditzel, H. J. Plasma Membrane Proteomics and Its Application in Clinical Cancer Biomarker Discovery. Molecular and Cellular Proteomics **2010**, 9 (7), 1369–1382. https://doi.org/10.1074/mcp.R900006-MCP200.

50) Cooper GM. The Cell: A Molecular Approach. 2nd edition. Sunderland (MA): Sinauer Associates; **2000**. Structure of the Plasma Membrane. Available from: https://www.ncbi.nlm.nih.gov/books/NBK9898/

51) The Medical Biochemistry Page, Biological Membranes and Membrane Transport Processes. https://themedicalbiochemistrypage.org/biological-membranes-and-membrane-transport-processes/

52) Elmahallawy, E. K.; Alkhaldi, A. A. M. Insights into Leishmania Molecules and Their Potential Contribution to the Virulence of the Parasite. Vet Sci **2021**, 8 (2), 1–21. https://doi.org/10.3390/vetsci8020033.

53) Liu, D.; Uzonna, J. E. The Early Interaction of Leishmania with Macrophages and Dendritic Cells and Its Influence on the Host Immune Response. Front Cell Infect Microbiol **2012**, 2 (June), 83. https://doi.org/10.3389/fcimb.2012.00083.

54) Proteomes, S.; Solis, N.; Cordwell, S. J. Chapter 4 Cell Shaving and False-Positive Control Strategies Coupled. **2016**, 47–55. https://doi.org/10.1007/978-1-4939-3676-2.

55) Elschenbroich, S.; Kim, Y.; Medin, J. A.; Kislinger, T. Isolation of Cell Surface Proteins for Mass Spectrometry-Based Proteomics. Expert Rev Proteomics **2010**, 7 (1), 141–154. https://doi.org/10.1586/epr.09.97.

56) Tiong, H. K.; Hartson, S.; Muriana, P. M. Comparison of Five Methods for Direct Extraction of Surface Proteins from Listeria Monocytogenes for Proteomic Analysis by Orbitrap Mass Spectrometry. J Microbiol Methods **2015**, 110, 54–60. https://doi.org/10.1016/j.mimet.2015.01.004.

57) Olaya-Abril, A.; Jiménez-Munguía, I.; Gómez-Gascón, L.; Rodríguez-Ortega, M. J. Surfomics: Shaving Live Organisms for a Fast Proteomic Identification of Surface Proteins. J Proteomics **2014**, 97, 164–176. https://doi.org/10.1016/j.jprot.2013.03.035.

58) Queiroz, R. M. L.; Charneau, S.; Motta, F. N.; Santana, J. M.; Roepstorff, P.; Ricart, C. A. O. Comprehensive Proteomic Analysis of Trypanosoma Cruzi Epimastigote Cell Surface Proteins by Two Complementary Methods. J Proteome Res **2013**, 12 (7), 3255–3263. https://doi.org/10.1021/pr400110h.

59) Rodríguez-Ortega, M. J.; Norais, N.; Bensi, G.; Liberatori, S.; Capo, S.; Mora, M.; Scarselli, M.; Doro, F.; Ferrari, G.; Garaguso, I.; Maggi, T.; Neumann, A.; Covre, A.; Telford, J. L.; Grandi, G. Characterization and Identification of Vaccine Candidate Proteins through Analysis of the Group A Streptococcus Surface Proteome. Nat Biotechnol **2006**, 24 (2), 191–197. https://doi.org/10.1038/nbt1179.

60) Dyballa, N.; Metzger, S. Fast and Sensitive Colloidal Coomassie G-250 Staining for Proteins in Polyacrylamide Gels. Journal of Visualized Experiments **2009**, No. 30, 2–5. https://doi.org/10.3791/1431.

61) Ishihama, Y.; Mann, M.; Extraction, G. STAGE (STop And Go Extraction) TIPS Desalting Procedure. **2003**, 17–18.

62) Queiroz, R. M. L.; Ricart, C. A. O.; Machado, M. O.; Bastos, I. M. D.; de Santana, J. M.; de Sousa, M. V.; Roepstorff, P.; Charneau, S. Insight into the Exoproteome of the Tissue-Derived Trypomastigote Form of Trypanosoma Cruzi. Front Chem **2016**, 4 (NOV), 1–10. https://doi.org/10.3389/fchem.2016.00042.

63) BoseDasgupta, S.; Ganguly, A.; Roy, A.; Mukherjee, T.; Majumder, H. K. A Novel ATP-Binding Cassette Transporter, ABCG6 Is Involved in Chemoresistance of Leishmania. Mol Biochem Parasitol **2008**, 158 (2), 176–188. https://doi.org/10.1016/j.molbiopara.2007.12.007.

64) Boozhmehrani, M. J.; Eslami, G.; Khamesipour, A.; Jafari, A. A.; Vakili, M.; Hosseini, S. S.; Askari, V. The Role of ATP-Binding Cassette Transporter Genes Expression in Treatment Failure Cutaneous Leishmaniasis. AMB Express **2022**, 12 (1). https://doi.org/10.1186/s13568-022-01419-5.

65) VaxiJen 2.0: a server for prediction of protective antigens, tumour antigens and subunit vaccines. Doytchinova IA, Flower DR. BMC Bioinformatics. **2007** Jan 5;8:4

66)Kumar, M.; Ranjan, K.; Singh, V.; Pathak, C.; Pappachan, A.; Singh, D. D. Hydrophilic Acylated Surface Protein A (HASPA) of Leishmania Donovani: Expression, Purification and Biophysico-Chemical Characterization. Protein Journal **2017**, 36 (4), 343–351. https://doi.org/10.1007/s10930-017-9726-x.

67) Pompa, A.; Walker, J. M. Unconventional Protein Secretion IN Series Editor. **2016**, 1459, 175–190. https://doi.org/10.1007/978-1-4939-3804-9.

68) Maclean, L. M.; O'Toole, P. J.; Stark, M.; Marrison, J.; Seelenmeyer, C.; Nickel, W.; Smith, D. F. Trafficking and Release of Leishmania Metacyclic HASPB on Macrophage Invasion. Cell Microbiol **2012**, 14 (5), 740–761. https://doi.org/10.1111/j.1462-5822.2012.01756.x.

69) Ulusan Bağcı, Ö.; Sadıqova, A.; Caner, A. Leishmania Major ve Leishmania Infantum Promastigot Formlarının Karşılaştırmalı Gen Ekspresyon Profilleri. Turkiye Parazitol Derg **2021**, 45 (2), 88–94. https://doi.org/10.4274/tpd.galenos.2021.66375.

70) Saha, A.; Bhattacharjee, A.; Vij, A.; Das, P. K.; Bhattacharya, A.; Biswas, A. Evaluation of Modulators of CAMP-Response in Terms of Their Impact on Cell Cycle and Mitochondrial Activity of Leishmania Donovani. Front Pharmacol **2020**, 11 (May), 1–12. https://doi.org/10.3389/fphar.2020.00782.

71) WoLF PSORT: Protein Subcellular Localization Prediction. https://wolfpsort.hgc.jp/

72)Escalona-Montaño, A. R.; Zuñiga-Fabián, M.; Cabrera, N.; Mondragón-Flores, R.; Gómez-Sandoval, J. N.; Rojas-Bernabé, A.; González-Canto, A.; Gutiérrez-Kobeh, L.; Pérez-Montfort, R.; Becker, I.; Aguirre-García, M. M. Protein Serine/Threonine Phosphatase Type 2C of Leishmania Mexicana. Front Cell Infect Microbiol **2021**, 11 (April), 1–11. https://doi.org/10.3389/fcimb.2021.641356.

73) Murungi, E.; Barlow, L. D.; Venkatesh, D.; Adung'a, V. O.; Dacks, J. B.; Field, M. C.; Christoffels, A. A Comparative Analysis of Trypanosomatid SNARE Proteins. Parasitol Int **2014**, 63 (2), 341–348. https://doi.org/10.1016/j.parint.2013.11.002.

74) Jespersen, M. C.; Peters, B.; Nielsen, M.; Marcatili, P. BepiPred-2.0: Improving Sequence-Based B-Cell Epitope Prediction Using Conformational Epitopes. Nucleic Acids Res **2017**, 45 (W1), W24–W29. https://doi.org/10.1093/nar/gkx346.

75) Pawlowic, M. C.; Zhang, K. Leishmania Parasites Possess a Platelet-Activating Factor Acetylhydrolase Important for Virulence. Mol Biochem Parasitol **2012**, 186 (1), 11–20. https://doi.org/10.1016/j.molbiopara.2012.08.005.

76) Rashidi, S.; Nguewa, P.; Mojtahedi, Z.; Shahriari, B.; Kalantar, K.; Hatam, G. Identification of Immunoreactive Proteins in Secretions of Leishmania Infantum Promastigotes: An Immunoproteomic Approach. Eastern Mediterranean Health Journal **2020**, 26 (12), 1548–1555. https://doi.org/10.26719/emhj.20.114.

77) Zhang, O.; Wilson, M. C.; Xu, W.; Hsu, F. F.; Turk, J.; Kuhlmann, F. M.; Wang, Y.; Soong, L.; Key, P.; Beverley, S. M.; Zhang, K. Degradation of Host Sphingomyelin Is Essential for Leishmania Virulence. PLoS Pathog **2009**, 5 (12). https://doi.org/10.1371/journal.ppat.1000692.

78) Chamakh-Ayari, R.; Bras-Gonçalves, R.; Bahi-Jaber, N.; Petitdidier, E.; Markikou-Ouni, W.; Aoun, K.; Moreno, J.; Carrillo, E.; Salotra, P.; Kaushal, H.; Negi, N. S.; Arevalo, J.; Falconi-Agapito, F.; Privat, A.; Cruz, M.; Pagniez, J.; Papierok, G. M.; Rhouma, F. B. H.; Torres, P.; Lemesre, J. L.; Chenik, M.; Meddeb-Garnaoui, A. In Vitro Evaluation of a Soluble Leishmania Promastigote Surface Antigen as a Potential Vaccine Candidate against Human Leishmaniasis. PLoS One **2014**, 9 (5), 1–12. https://doi.org/10.1371/journal.pone.0092708.

79) Devault, A.; Bãuls, A. L. The Promastigote Surface Antigen Gene Family of the Leishmania Parasite: Differential Evolution by Positive Selection and Recombination. BMC Evol Biol **2008**, 8 (1), 1–15. https://doi.org/10.1186/1471-2148-8-292.

80) Aldfer, M. M.; AlSiari, T. A.; Elati, H. A. A.; Natto, M. J.; Alfayez, I. A.; Campagnaro, G. D.; Sani, B.; Burchmore, R. J. S.; Diallinas, G.; De Koning, H. P. Nucleoside Transport and Nucleobase Uptake Null Mutants in Leishmania Mexicana for the Routine Expression and Characterization of Purine and Pyrimidine Transporters. Int J Mol Sci **2022**, 23 (15), 1–20. https://doi.org/10.3390/ijms23158139.

81) Corrales, R. M.; Sereno, D.; Mathieu-Daudé, F. Deciphering the Leishmania Exoproteome: What We Know and What We Can Learn. FEMS Immunol Med Microbiol **2010**, 58 (1), 27–38. https://doi.org/10.1111/j.1574-695X.2009.00608.x.

82) Mercado-Camargo, J.; Cervantes-Ceballos, L.; Vivas-Reyes, R.; Vivas-Reyes, R.; Vivas-Reyes, R.; Pedretti, A.; Serrano-García, M. L.; Gómez-Estrada, H. Homology Modeling of Leishmanolysin (Gp63) from Leishmania Panamensis and Molecular Docking of Flavonoids. ACS Omega **2020**, 5 (24), 14741–14749. https://doi.org/10.1021/acsomega.0c01584.

83) Al-jaghasi, A.; Abbady, A.; Al-khatib, S. Cloning, Expression and Purification Of. **2021**, 14 (2), 367–372.

84) de Mendonça, S. C. F.; Cysne-Finkelstein, L.; Matos, D. C. de S. Kinetoplastid Membrane Protein-11 as a Vaccine Candidate and a Virulence Factor in Leishmania. Front Immunol **2015**, 6 (OCT), 1–6. https://doi.org/10.3389/fimmu.2015.00524.

85) Sannigrahi, A.; Maity, P.; Karmakar, S.; Chattopadhyay, K. Interaction of KMP-11 with Phospholipid Membranes and Its Implications in Leishmaniasis: Effects of Single Tryptophan Mutations and Cholesterol. Journal of Physical Chemistry B **2017**, 121 (8), 1824–1834. https://doi.org/10.1021/acs.jpcb.6b11948.

86) Jackson, A. P. The Evolution of Amastin Surface Glycoproteins in Trypanosomatid Parasites. Mol Biol Evol **2010**, 27 (1), 33–45. https://doi.org/10.1093/molbev/msp214.

87) Ramakrishnan, S.; Docampo, R. Membrane Proteins in Trypanosomatids Involved in Ca2+ Homeostasis and Signaling. Genes (Basel) **2018**, 9 (6), 1–18. https://doi.org/10.3390/genes9060304.

88) Bouazizi-Ben Messaoud, H.; Guichard, M.; Lawton, P.; Delton, I.; Azzouz-Maache, S. Changes in Lipid and Fatty Acid Composition During Intramacrophagic Transformation of Leishmania Donovani Complex Promastigotes into Amastigotes. Lipids **2017**, 52 (5), 433–441. https://doi.org/10.1007/s11745-017-4233-6.

89) Hargrove, T. Y.; Wawrzak, Z.; Liu, J.; Nes, W. D.; Waterman, M. R.; Lepesheva, G. I. Substrate Preferences and Catalytic Parameters Determined by Structural Characteristics of Sterol 14α-Demethylase (CYP51) from Leishmania Infantum. Journal of Biological Chemistry **2011**, 286 (30), 26838–26848. https://doi.org/10.1074/jbc.M111.237099.

90) Carraro, C.; Id, P.; Reis, L. C.; Id, E. M. R.; Mary, C.; Id, O.; Id, C. C.; Valadares, E.; Levatti, D. C.; Umbelino, J.; Id, D. A. Parasitophorous Vacuole Biogenesis Via Cholesterol Homeostasis; **2019**.

91) Hallgren, J.; Tsirigos, K. D.; Damgaard Pedersen, M.; Juan, J.; Armenteros, A.; Marcatili, P.; Nielsen, H.; Krogh, A.; Winther, O. DeepTMHMM Predicts Alpha and Beta Transmembrane Proteins Using Deep Neural Networks. https://doi.org/10.1101/2022.04.08.487609.

92) Baghalabadi, V.; Razmi, H.; Doucette, A. Salt-Mediated Organic Solvent Precipitation for Enhanced Recovery of Peptides Generated by Pepsin Digestion. *Proteomes* **2021**, *9* (4). https://doi.org/10.3390/proteomes9040044

APPENDİX A

Chemicals, Devices and Bioinformatics Tools Used In The Study

Table A.1.List of Chemicals Used

Chemical Name	**Brand of Chemical**
Acetic Acid	Sigma Aldrıch
Ammonium Per Sulfate (APS)	Sigma
TEMED	AppliChem
Coomassie Brillant Blue G-250	Applichem Panreac
Aluminum sulfate 18 hydrate	Merck
SDS	AppliChem
Trisma base	Sigma
Glycine	AppliChem
Urea	Thermo Scientific
Silver Nitrate	Merck
DTT	Sigma
IAA	Sigma
Trypsin	Promega
TFA	Thermo Scientific
Ethanol	Sigma Aldrıch
Methanol	Sigma Aldrıch
Formaldehyde	Formaldehyde
C18 SPE cart	Finisterre C18
Acetonitrile (HPLC gradient)	Sigma Aldrıch
Ultrapure Water	Merck

Table A.2.List of Devices Used

Devices Name	**Brand of Devices**
NanoDrop	Thermo Scientific
pH meter	P-Selecta
Electrophoresis system	Bio-Rad
Power supply	Bio-Rad
Microcentrifuges	Sigma
Airstream/air filters	ESCO Scientific
Vortex	VELP SCIENTIFICA
Magnetic Stirrer	IKA- RCT classic
Speed-Vac	CHRIST Rotational Vacuum Concentrator RVC
Incubator	P SELECTA
Shaker	Infors HT
Micropipette	Eppendorf Research
HPLC column brand	Acclaim pepmap 100
HPLC	Dionex UltiMate 3000 RS Autosampler-Thermo Scientific
Orbitrap Fusion Mass Spectrometer	Thermo Orbitrap Fusion Tribrid

Table A.3.List of Bioinformatic Tools Used

Bioinformatic Tools	**Functions**	**Web site**
WoLF PSORT II	Protein Subcellular Localization Prediction	https://wolfpsort.hgc.jp/
Gene Ontology	Functions of Genes	http://geneontology.org/
TMHMM - 2.0	Transmembrane Topology Prediction	https://dtu.biolib.com/DeepTMHMM
SignalP - 6.0	Prediction of Signal Peptides and their cleavage sites	https://services.healthtech.dtu.dk/services/SignalP-6.0/
DeepLoc - 1.0	Prediction of eukaryotic protein subcellular localization	https://services.healthtech.dtu.dk/services/DeepLoc-1.0/
VaxiJen2.0	Prediction of Protective Antigens	http://www.ddg-pharmfac.net/vaxijen/VaxiJen/VaxiJen.html
IEDB.org	Immune Epitope Database	http://www.iedb.org/home_v3.php
BepiPred - 2.0	Prediction of potential linear B-cell epitopes	https://services.healthtech.dtu.dk/services/BepiPred-2.0/
ABCpred	Artificial neural network-based B-cell epitope prediction server	https://webs.iiitd.edu.in/raghava/abcpred/ABC_submission.html
AllerCatPro	Protein allergenicity potential prediction (version 2.0)	https://allercatpro.bii.a-star.edu.sg/
UniProt	Blast/Align/ Peptide search	https://www.uniprot.org/help/uniprotkb
Venn Diagram	Calculate and draw custom Venn diagrams	https://bioinformatics.psb.ugent.be/webtools/Venn/

APPENDIX B

All Proteins Identified in Results

Table B.1. Common Proteins in *L.Tropica, L.Major, L.Infantum, L.Donovani*

UniProtKB Code	Identified Protein	UniProtKB Code	Identified Protein
E9ADF0	BAR domain-containing protein	Q4Q587	Uncharacterized protein
A0A6L0XBL8	Cysteine peptidase Clan CA family C2 putative	Q4QF35	Proliferating cell nuclear antigen
E5LCR1	Acetyl-coenzyme A synthetase (EC 6.2.1.1)	A0A3S7WVE9	Mitogen-activated protein kinase kinase kinase, putative (Mitogen-activated_protein_kinase_kinase_kinase_putative/GeneDB:LmjF.19.0150) (Protein kinase domain family protein)
O02612	Histone H2B	A0A6L0XT25	Glycosomal_membrane_protein_-_putative
Q4Q0Z1	Paraflagellar rod component	A0A6L0XTK1	Eukaryotic_translation_initiation_factor_3_subunit_8_-_putative
A4HZF8	Putative RNA helicase	E9AXJ6	ATP synthase subunit beta (EC 7.1.2.2)
Q4Q1V0	Fructose-bisphosphate aldolase (EC 4.1.2.13)	Q4Q4E4	Paraflagellar rod component
Q4Q1F5	Dihydrolipoamide acetyltransferase component of pyruvate dehydrogenase complex (EC 2.3.1.-)	A0A3S7WR51	Glycerol-3-phosphate dehydrogenase [NAD(+)] (EC 1.1.1.8)
E9AXU6	Succinate--CoA ligase [ADP-forming] subunit alpha, mitochondrial (EC 6.2.1.5) (Succinyl-CoA synthetase subunit alpha) (SCS-alpha)	Q4QFP8	Putative small myristoylated protein-3
Q66NE0	Elongation factor 1B gamma (Elongation factor-1 gamma)	A0A3Q8IUV8	H(+)-transporting two-sector ATPase (EC 7.1.2.2)
A0A504Y1G6	methionine--tRNA ligase (EC 6.1.1.10)	A0A504XEQ3	2-oxoglutarate_dehydrogenase_subunit_putative
A0A3Q8IRB5	EF-hand domain pair family protein	E9AKI4	Uncharacterized protein
A0A381MFJ0	Plasma membrane ATPase (EC 7.1.2.1)	E9AUS1	Putative calpain-like cysteine peptidase
Q4Q833	Eukaryotic translation initiation factor 3 subunit E	A0A6L0X4B0	Phosphodiesterase (EC 3.1.4.-)
A0A6J8FPV3	Isoleucine--tRNA ligase (EC 6.1.1.5)	E9AGK7	Tubulin alpha chain
Q4QGE5	Metallo-peptidase, Clan MA(E), family 32	Q4QH17	Putative aminopeptidase (EC 3.4.11.-)
A0A6L0XPY9	Heat shock 70-related_protein_1 mitochondrial precursor putative	A0A6L0XVC5	CCR4-NOT transcription complex subunit 10 putative
A0A504X5R4	Pyruvate kinase (EC 2.7.1.40)	Q4Q8E6	Putative glucose-regulated protein 78
A0A504WXS7	Biotin/lipoate A/B protein ligase family protein	E9AFJ4	Glycosomal membrane protein
Q4Q090	Phosphoglycerate mutase (2,3-diphosphoglycerate-independent) (EC 5.4.2.12)	A0A3S7XBG7	Adenosylhomocysteinase (EC 3.3.1.1)

(Cont.on next page)

Table B.1. (cont.)

UniProtKB Code	Identified Protein	UniProtKB Code	Identified Protein
A0A6L0XSV1	Paraflagellar rod component putative	A0A3Q8IH18	Serine/threonine-protein phosphatase (EC 3.1.3.16)
Q4Q1M7	Phosphomannomutase (EC 5.4.2.8)	E9AT99	14-3-3 protein-like protein
A5JV96	Cytosolic tryparedoxin	A0A3S7X410	Peptidyl-prolyl cis-trans isomerase (PPIase) (EC 5.2.1.8)
Q4QIK0	Splicing factor ptsr1-like protein	A0A7G7LIC0	Actin
A0A504XWS5	60Kd inner membrane family protein	E9AE36	Putative paraflagellar rod protein 1D
E9ATG6	Glycine--tRNA ligase (EC 6.1.1.14) (Diadenosine tetraphosphate synthetase)	A0A3Q8IKG2	Zinc knuckle family protein (Zinc knuckle, putative) (Zinc_knuckle_putative/Pfam:PF00098/Pfam:PF13917/Pfam:PF14392/Pfam:PF13696)
E9AD27	Putative calpain-like cysteine peptidase	E9ARZ1	Putative 40S ribosomal protein S13
A0A504XGG2	Hypothetical protein conserved	E9AV90	Putative 40S ribosomal protein S23
Q4Q0X8	Proteasome regulatory ATPase subunit	A0A504XTQ3	Ribosomal L38e family protein
Q4Q3V3	Succinyl-CoA:3-ketoacid-coenzyme A transferase (EC 2.8.3.5)	A0A0M9GB44	Putative 60S ribosomal protein L23
A0A3S5H6Q3	Ubiquitin-conjugating enzyme family protein	A4I011	Putative 40S ribosomal protein L14
E9AEL4	Probable eukaryotic initiation factor 4A (ATP-dependent RNA helicase eIF4A)	Q4QJ39	Putative 60S ribosomal protein L19
A7LBL5	ADP/ATP translocase (ADP, ATP carrier protein)	E9AEA8	Putative 60S ribosomal protein L13
A0A3Q8IF94	Probable eukaryotic initiation factor 4A (EC 3.6.4.13) (ATP-dependent RNA helicase eIF4A)	Q9NGK2	40S ribosomal protein S24
A0A3S5H6Q8	AhpC/TSA family/Thioredoxin-like, putative (Thioredoxin-like family protein)	A0A3Q8IN17	60S ribosomal protein L13a, putative (60S_ribosomal_protein_L13a_putative (Ribosomal protein L13)
Q4Q0Q2	PUM-HD domain-containing protein	E9AN13	Putative ribosomal protein l35a
A0A3S5H821	Poly-zinc finger protein 2, putative (Poly-zinc finger_protein 2 putative	E9AYL4	Putative 60S ribosomal protein L35
A0A504XJ57	Hypothetical_protein_conserved	A0A3S7X558	40S ribosomal protein S2 (40S_ribosomal_protein_S2/GeneDB:LmjF.32.0450)
A0A504Y0K5	Mannosyl-oligosaccharide glucosidase (EC 3.2.1.106)	E9AVH4	40S ribosomal protein S6
A0A3S7X7R9	Hypothetical_protein_conserved	A0A3Q8IGZ0	40S ribosomal protein S16, putative (40S_ribosomal_protein_S16_putative) (Ribosomal protein S9/S16 family protein)
A0A6J8F523	14-3-3 protein_2 putative	Q9N9V3	Putative ribosomal protein L10

(Cont.on next page)

Table B.1. (cont.)

UniProt KB Code	Identified Protein	UniProtKB Code	Identified Protein
E9ACA6	Soluble NSF attachment protein SNAP putative	A0A3Q8IAA6	60S ribosomal protein L6
A0A504Y0F7	Histone H4	A0A504XI55	60S_ribosomal protein_L2_putative
A0A504X036	H(+)-exporting diphosphatase (EC 7.1.3.1)	A4HT78	60S ribosomal protein L7a
E9AN29	Putative nucleolar protein	E9APE6	Putative 40S ribosomal protein S4
A0A6J8FUI5	Hypothetical_protein_conserved	E9AL81	Putative 40S ribosomal protein S9
E9AHP6	Kinesin motor domain-containing protein	Q4Q504	60S ribosomal protein L27
Q4Q2T7	Alba domain-containing protein	P25204	40S ribosomal protein S8
E9AZQ9	Archaic_translocase_outer_mitochondrial_membrane_40_-_putative	A0A836FPG6	Ribosomal_L7Ae domain-containing protein
A0A3Q8IFS5	Hypothetical_protein_conserved	A0A3S7XBN1	60S ribosomal protein L18, putative
A0A504X3P8	Calmodulin	A0A3S7WWL2	60S ribosomal protein L9, putative (60S_ribosomal_protein_L9_putative/GeneDB:LmjF.21.1050/GeneDB:LmjF.30.3340)
A0A3Q8IGV2	COP9 signalosome, subunit CSN8 family protein	E9AJC9	40S ribosomal protein S7
Q4Q3T3	Metallo-peptidase, Clan MA(E), Family M32 (EC 3.4.17.19)	A0A504XGS2	Ribosomal protein S3
Q4Q5K7	Putative RNA binding protein	A0A504X354	Ribosomal protein L34e family protein
Q4QDQ2	Putative heat shock protein	A0A3S7X989	Ribosomal protein L11, N-terminal domain family protein
Q8WQX2	Inosine-uridine preferring nucleoside hydrolase (EC 3.2.2.1	P48157	60S ribosomal protein L11
E9AJG4	Probable eukaryotic initiation factor 4A (ATP-dependent RNA helicase eIF4A)	A0A504XZK6	60S_ribosomal_protein_L32/GeneDB:LmjF.21.1720 (Ribosomal protein L32 family protein)
Q07DU5	Peroxidoxin 2	E9ALD8	40S ribosomal protein S19-like protein
A0A6L0XK54	Glyceraldehyde-3-phosphate dehydrogenase (EC 1.2.1.12)	A0A504XVT2	Vacuolar (H+)-ATPase G subunit family protein
Q4QIQ9	Putative ATP-dependent DEAD/H RNA helicase	A0A504XKJ5	dimethylargininase (EC 3.5.3.18)
Q4QAC4	Transketolase (EC 2.2.1.1)	Q2QKR1	Actin severing and dynamics regulatory protein
E8NHQ7	Putative ATPase alpha subunit (WGS CADB00000000 data, contig 86)	E9B0B3	Glutamate dehydrogenase
A0A3Q8IIY4	Proteasome subunit beta	A0A504XG57	RNA pseudouridylate synthase family protein

Table B.2. All Proteins Identified in *L. Tropica , L. Major* and *L. Infantum*

UniProtKB Code	Identified Protein	UniProtKB Code	Identified Protein
A4IB12	Putative NADH-dependent fumarate reductase (EC 1.3.1.6)	E9AFD4	Probable eukaryotic initiation factor 4A (ATP-dependent RNA helicase eIF4A)
A0A504XH80	Hypothetical_protein_conserved	Q4Q3I9	Agglutinin domain-containing protein
Q4QG97	40S ribosomal protein S12	E9BJZ5	40S ribosomal protein S17, putative
A0A504Y8T0	Peptidase M3 family protein	E9ASW8	Putative nucleoside transporter 1
A0A381MH18	Paraflagellar rod protein 2C		

Table B.3. All Proteins Identified in *L.Tropica, L.Major and L.Donovani*

UniProtKB Code	Identified Protein	UniProtKB Code	Identified Protein
A4I341	Putative heat shock protein DNAJ	E9AHF3	Glutaredoxin-like protein
E9AHL7	Putative ADP-ribosylation factor	Q4QBW5	Condensin complex subunit 1
E9AZR2	40S ribosomal protein S14	Q9N852	TNFR-Cys domain-containing protein
Q8WT35	Histone H2A	Q4QCC9	Uncharacterized protein
Q4Q4I6	Heat shock protein 83-1	A0A504Y2G2	Ribosomal protein S10
E9ACC9	Ubiquitin-like domain-containing protein	Q4QGN0	alanine transaminase (EC 2.6.1.2)
Q4Q7X1	Glutamate dehydrogenase		

Table B.4. All Proteins Identified in *L.Tropica, L.Infantum* and *L.Donovani*

UniProtKB Code	Identified Protein	UniProtKB Code	Identified Protein
A0A6L0WS81	Cysteine--tRNA ligase (EC 6.1.1.16)	A0A5B8YY69	Glucose-6-phosphate isomerase (EC 5.3.1.9)
A0A6L0XQ80	Protein mkt1 putative	A0A3Q8IEB8	HMG (High mobility group) box family protein
A0A3Q8IEW8	Proteasome regulatory non-ATP-ase subunit 2, putative	A0A6J8FHB1	Epsin_putative/GeneDB:LmjF.25.0670
A4IDC1	Protein kinase	A4I2Y7	Phosphoenolpyruvate carboxykinase (ATP) (EC 4.1.1.49)
Q4QAC5	Putative 60S ribosomal protein L26	A0A6J8FAA6	Cyclic_nucleotide-binding_domain_containing_protein_putative
A0A504XFL7	WD domain, G-beta repeat family protein	A0A6L0XSK0	Kinetoplastid membrane protein-11

(Cont. on next page)

Table B.4. (cont.)

UniProtKB Code	Identified Protein	UniProtKB Code	Identified Protein
E9AUL9	Vesicle-fusing ATPase (EC 3.6.4.6)	Q4QGX9	Pyruvate, phosphate dikinase (EC 2.7.9.1)
A0A504XJA5	40S ribosomal protein S8	A0A6L0XKN3	Hypothetical_protein
Q4Q1I3	Methyltransferase (EC 2.1.1.-)	E9ARJ9	Putative ATP-dependent zinc metallopeptidase
A0A6J8FQF9	Ribosomal_protein_L3_putative	A0A6L0XPY4	Hypothetical_protein_-_conserved
A4HVL6	C3H1-type domain-containing protein	A0A3Q8IC74	Prefoldin subunit 3
Q4QGN9	Glucose-6-phosphate isomerase (EC 5.3.1.9)	Q4QJ67	Peptidyl-prolyl cis-trans isomerase (PPIase) (EC 5.2.1.8)
A0A504X048	Amidohydrolase family protein	A0A6J8FLW5	Asparagine--tRNA ligase (EC 6.1.1.22)
E9ANU8	Proteasome regulatory ATPase subunittcc1l8.3,putative	Q4Q0G9	Haloacid dehalogenase-like hydrolase-like protein
E9AU83	Putative ubiquitin fusion degradation protein	A0A836GZ07	CrtC domain-containing protein
Q4QGD8	Protein kinase A regulatory subunit	Q4QC71	Putative adenylate kinase (EC 2.7.4.3)
Q4Q702	Putative aldehyde dehydrogenase (EC 1.2.1.3)	Q4QAY0	Nucleoside 2-deoxyribosyltransferase
A0A6L0XJ26	Ubiquitin carboxyl-terminal hydrolase (EC 3.4.19.12)	A0A3Q8IEP3	Tyrosine-protein phosphatase domain-containing protein
Q4Q6F6	Biotin/lipoate protein ligase-like protein (EC 6.3.4.15)	A0A6L0XLL9	MORN repeat family protein
E9AES2	Putative aldose 1-epimerase (EC 5.1.3.3)	A0A504X198	PCI domain family protein
P27892	Histone H2A.2	A4HYC9	Putative proteasome regulatory non-ATP-ase subunit
A4I9N5	Putative proteasome regulatory non-ATP-ase subunit 11	A0A3S5H6V9	Nucleolar RNA binding protein, putative
F1DAX4	Methyltransferase (EC 2.1.1.-)	A0A3S5H6R4	N-terminal_conserved domain of Nudc./CS domain containing protein putative
A0A836HJ30	PHR domain-containing protein	A0A504Y0L5	Eukaryotic translation initiation factor 3 subunit E (eIF3e)
E9ANF6	Serine--tRNA ligase (EC 6.1.1.11) (Seryl-tRNA synthetase)	Q4Q7J1	Importin subunit alpha
A0A504XM54	Ribosomal protein S17 family protein	A0A0R6Y3Z1	Rhodanese domain-containing protein
E9AFD0	Putative ubiquitin-activating enzyme e1	A0A6L0XJN6	FACT complex subunit
A0A3Q8IBA7	Metallo-peptidase, Clan MG, Family M24	A4HYA8	4-coumarate:coa ligase-like protein (EC 6.2.1.12)
A0A6J8FBD6	Elongation factor 1-alpha	Q4QGA9	Alba domain-containing protein
Q4QBV0	Putative 40S ribosomal protein S15	Q4Q5X1	Putative aldehyde reductase
E9AKQ4	Putative 5'-3' exonuclease (EC 3.1.11.-)	A4HWN5	Ribonucloprotein
Q4QEI9	Elongation factor 1-alpha	H6V7N3	Ascorbate peroxidase (EC 1.11.1.11) (Ascorbate_peroxidase/GeneDB:LmjF.34.0070)
E9ALD1	Proteasome regulatory non-ATP-ase subunit, putative	Q4Q1X7	Putative 40S ribosomal protein S10
E9AYZ5	Putative isovaleryl-coA dehydrogenase (EC 1.3.99.-)	E9ANP0	40S ribosomal protein S5
A0A3S7X9N0	Cys/Met metabolism PLP-dependent enzyme family protein	Q4QBD8	Putative NADP-dependent alcohol dehydrogenase (EC 1.1.1.1)

(Cont. on next page)

Table B.4 (cont.)

UniProtKB Code	Identified Protein	UniProtKB Code	Identified Protein
E9B9H0	Amastin surface glycofamily protein	A0A3Q8IF50	RNA recognition motif family protein (RNA-binding protein, putative)
A0A6L0WJR8	Leishmanolysin (EC 3.4.24.36)	Q4Q455	Cysteine conjugate beta-lyase,aminotransferase-like protein
E9AE86	Aminopeptidase (EC 3.4.11.-)	E9AHZ7	Phosphoglycerate mutase (2,3-diphosphoglycerate-independent) (EC 5.4.2.12)
P21620	Inosine-5'-monophosphate dehydrogenase (IMP dehydrogenase)	Q4QJ42	Putative glutamine synthetase (EC 6.3.1.2)
E9ALX8	FGE-sulfatase domain-containing protein	A0A6J8FGW4	Glutamate--tRNA ligase (EC 6.1.1.17)
E9AEJ6	FCP1 homology domain-containing protein	A0A3S7WQY1	Dehydrogenase-like protein (Short chain dehydrogenase family protein)
E9ARW1	Putative fibrillarin	A0A504XMF5	Protein kinase domain family protein
A0A3Q8IC41	Proteasome subunit alpha type	A0A6L0XPB6	Transmembrane 9 superfamily member
E9AFP0	Polyadenylate-binding protein (PABP)	A0A3Q8IHK1	AP-1 complex subunit gamma
Q94593	Leishmanolysin (EC 3.4.24.36)		

Table B.5. All Proteins Identified in *L.Major, L.Infantum* and *L.Donovani*

UniProtKB Code	Identified Protein	UniProtKB Code	Identified Protein
A4I041	Peroxidoxin	A0A6L0WY88	Dynein_heavy_chain_-_putative
A4HWU2	Putative fucose kinase (EC 2.7.1.52)	A0A504XDW6	Guanine nucleotide-binding protein subunit beta-like protein
A0A504XEX2	Calpain cysteine protease family protein	A0A504XXL8	Adaptin N terminal region family protein
A0A3S5H732	60S ribosomal protein L10a, putative	Q4Q1W6	Putative ribosomal protein L24
E9BL65	Serine/threonine-protein kinase, putative	E9B7A8	40S ribosomal protein S7
Q4Q063	CCT-theta		

Table B.6. All Proteins Identified in *L.Tropica, L.Major*

UniProtKB Code	Identified Protein	UniProtKB Code	Identified Protein
E9ADB2	Arginine--tRNA ligase (EC 6.1.1.19) (Arginyl-tRNA synthetase)	E9ARL1	Putative elongation factor Tu
E9ALW3	ANK_REP_REGION domain-containing protein	E9AVV4	Uncharacterized protein
A0A6L0WQR7	ADP/ATP translocase (ADP, ATP carrier protein)	A0A3Q8I931	Tb-292 membrane associated protein-like protein
A0A3Q8IGG5	Scd6-like Sm domain containing protein, putative	E9AV44	Proteasome regulatory non-ATP-ase subunit 5,putative
Q4Q6L9	Putative cytoskeleton-associated protein CAP5.5	E9AFK0	Putative 60S ribosomal protein L27A/L29
Q4QB61	Uncharacterized protein	E9BL24	Rab GDP dissociation inhibitor
Q4QD56	Peptidylprolyl isomerase (EC 5.2.1.8)	Q4QI58	Stress-induced protein sti1
Q4Q7Y4	Putative heat-shock protein hsp70		

Table B.7. All Proteins Identified in *L.Tropica, L.Infantum*

UniProtKB Code	Identified Protein	UniProtKB Code	Identified Protein
A4HX87	ATP pyrophosphate-lyase (Adenylyl cyclase)	A0A504XCI8	Proteasome regulatory subunit C-terminal family protein
P62883	Guanine nucleotide-binding protein subunit beta-like protein (Antigen LACK)	Q4Q9Y5	RRM domain-containing protein
E9B5C7	Glucose-6-phosphate 1-epimerase (EC 5.1.3.15)	E9ADY9	Putative carnitine/choline acetyltransferase (EC 2.3.1.-)
A0A504XAQ2	Eukaryotic translation initiation factor 3 subunit 7 (eIF-3) family protein	A0A6J8FMG6	Heat_shock_protein_putative/GeneDB:LmjF.33.2390
A0A6J8FHL8	Peptidase_m20/m25/m40_family-like_protein/GeneDB: LmjF.31.1890	A0A3Q8IRQ9	Adenylyl cyclase-associated protein, putative (Adenylyl_cyclase-associated_protein_putative/Gene DB:LmjF.36.5590)
A0A504X2U7	Calreticulin	Q4QDF8	Putative histone H3 variant
A0A504X8V9	Dynein light chain type 1 family protein	A0A3Q8IDK7	Glutathione peroxidase
A0A504XNH3	Malate dehydrogenase (EC 1.1.1.37)	A4HXZ5	Proline--tRNA ligase (EC 6.1.1.15) (Prolyl-tRNA synthetase)
A4HWV9	Core histone H2A/H2B/H3/H4 family protein (Histone H3) (Histone_H3)	A0A6L0XHZ8	TPR-repeat protein putative
A0A6L0XUN7	Protein disulfide-isomerase (EC 5.3.4.1)	E9AD89	Mitochondrial RNA binding protein 1

(Cont.on next page)

Table B.7. (cont.)

UniProtKB Code	Identified Protein	UniProtKB Code	Identified Protein
A0A3Q8ID36	Succinate dehydrogenase [ubiquinone] flavoprotein subunit, mitochondrial (EC 1.3.5.1)	A0A504XKY3	Similarity to endo-1-like protein
E9B217	RAP domain-containing protein	O97220	TPR_REGION domain-containing protein
E9ARJ3	V-type proton ATPase subunit C	A0A6J8FKM5	Amidohydrolase_putative/GeneDB:LmjF.30.1380
P36982	Protein phosphatase 2C (PP2C) (EC 3.1.3.16)	E9ANH0	Proteasome subunit alpha type
A0A6L0XUF9	Magnesium chelatase	E9AU65	Putative homocysteine S-methyltransferase (EC 2.1.1.10)
A0A504X6X3	Prolyl endopeptidase (EC 3.4.21.-)	A4HY74	ANK REP REGION domain-containing protein
E9AG07	ANK_REP_REGION domain-containing protein	A6N857	Glycogen synthase kinase 3 short
Q70GE8	Thiol-dependent reductase 1 (EC 1.8.5.1)	E9ALU9	TPR_REGION domain-containing protein
Q4QDR6	Mitochondrial carrier protein	E9B078	Putative 40S ribosomal protein S17
A0A504XWS4	Microtubule-associated protein	E9AVW1	Ribonucleoside-diphosphate reductase small chain,putative
A0A6L0XSH3	Cysteine desulfurase (EC 2.8.1.7)	A0A6L0Y0J3	Cell_cycle sequence binding phosphoprotein (RBP33) putative
A0A504X2R6	Eukaryotic translation initiation factor 3 subunit L (eIF3l)	A0A504XK79	La domain family protein
E9ADX4	Tryparedoxin	E9AWK3	Ankyrin/TPR repeat protein
Q4Q271	Stress-inducible protein STI1 homolog	A0A6J8FRU2	Clathrin heavy chain
Q4Q5J5	Phenylalanine--tRNA ligase (EC 6.1.1.20)	Q4QIG6	Putative Qc-SNARE protein

Table B.8. All Proteins Identified in *L.Tropica* and *L.Donovani*

UniProtKB Code	Identified Protein	UniProtKB Code	Identified Protein
A0A6L0XU71	RNA recognition motif.	E9AZL0	Putative propionyl-coa carboxylase beta chain (EC 6.4.1.3)
Q4QFF2	Putative ribonucleoprotein p18, mitochondrial	A0A6L0XHY4	phosphoglucomutase (alpha-D-glucose-1,6-bisphosphate-dependent) (EC 5.4.2.2)
A0A6L0WL20	Proteasome subunit beta	I3VJS9	6-phosphogluconate dehydrogenase, decarboxylating (EC 1.1.1.44)
A4HV47	Proteasome subunit beta	E9AWK2	Guanine nucleotide-binding protein beta subunit-like protein
E9ATU8	VHS domain-containing protein	E9AK98	Putative RNA-binding protein

(Cont. on next page)

T

Table B.8. (cont.)

UniProtKB Code	Identified Protein	UniProtKB Code	Identified Protein
E9ARJ6	Putative RNA binding protein	A0A3S7WTV8	Dihydroorotate dehydrogenase (fumarate) (EC 1.3.98.1)
E9AWI3	V-type proton ATPase subunit a	A4I8S1	Uncharacterized protein
Q4Q1M1	Chaperonin HSP60, mitochondrial	Q25292	Heat-shock protein
A0A504XQC8	deoxyribose-phosphate aldolase (EC 4.1.2.4) (2-deoxy-D-ribose 5-phosphate aldolase)	E9AKE5	Protein kinase domain-containing protein
A0A6L0Y1Z8	CCT-eta	E9AKM8	Surface antigen-like protein
Q4Q3J4	Malate dehydrogenase (EC 1.1.1.37)	A0A504WZS1	Aldo/keto reductase family protein
Q5EXB3	Phosphotransferase (EC 2.7.1.-)	E9BMT4	Cytochrome c oxidase VIII (COX VIII), putative
Q4QEJ8	Separase	A0A6L0WIT3	Hypothetical protein conserved
Q306G5	Prohibitin	E9AKM1	V-type proton ATPase subunit
E9AUK9	Putative serine/threonine protein phosphatase 2A regulatory subunit	A4HTC3	Putative nucleolar RNA-binding protein
E9AJT7	Putative 26S protease regulatory subunit (EC 3.6.4.8)	A0A504XQ51	Aminopeptidase (EC 3.4.11.-)
Q4QFT9	CS domain-containing protein	Q4QDX3	Probable citrate synthase, mitochondrial (EC 2.3.3.16)
E9AZP7	Putative RNA binding protein rbp16	Q4Q127	Eukaryotic translation initiation factor 3 subunit I (eIF3i)
A0A504XXH7	Tubulin/FtsZ family, GTPase domain protein	A0A3S7X190	Serine/threonine-protein phosphatase (EC 3.1.3.16)
A0A6J8FER6	Translation_initiation_factor_eIF-2B_subunit_epsi lon_putative/GeneDB:LmjF.28.2290	E9AM02	Putative lipophosphoglycan biosynthetic protein
E9AWD4	Guanine nucleotide-binding protein subunit beta-like protein	A0A6J8F3S0	Splicing factor ptsr1-like protein/GeneDB:LmjF.07.0870
E9ALU2	Putative carnitine/choline acetyltransferase (EC 2.3.1.-)	A0A6L0XS69	NADH-dependent_fumarate_reductase_-_putative
A0A3Q8IJB4	Dynein-associated roadblock protein-like protein (Roadblock/LC7 domain family protein)	A0A504WUV1	CS domain-containing protein
A4I5F6	pyridoxal kinase (EC 2.7.1.35)	A0A6L0Y094	H/ACA ribonucleoprotein complex subunit
Q9NL75	Aspartate carbamoyltransferase (EC 2.1.3.2)	E9AML1	Calmodulin
A4I5C0	Adenosine kinase (AK) (EC 2.7.1.20)	Q6Y9R4	phosphogluconate dehydrogenase (NADP(+)-dependent, decarboxylating) (EC 1.1.1.44)
A4IB24	Putative ubiquitin-conjugating enzyme e2 (EC 6.3.2.19)	E9AWQ4	Kinesin-like protein
Q4QGK1	Putative surface antigen protein 2	A0A504XII7	Transketolase, pyrimidine binding domain family protein
A0A3S7X3V5	Guanine nucleotide-binding protein subunit beta-like protein	A0A3S7X5X3	Nucleoside diphosphate kinase (EC 2.7.4.6)
Q4QI04	VTC domain-containing protein	Q4QCD1	ARID domain-containing protein
R9VW71	Thiol-specific antioxidant	Q4QD53	Inosine-5'-monophosphate dehydrogenase (EC 1.1.1.205)

(Cont.on next page)

Table B.8. (cont.)

UniProtKB Code	Identified Protein	UniProtKB Code	Identified Protein
Q4QFL8	Phosphopyruvate hydratase (EC 4.2.1.11)	Q4Q7R2	Aspartate--tRNA ligase (EC 6.1.1.12) (Aspartyl-tRNA synthetase)
A0A504XQJ1	T-complex protein 1 subunit gamma	A0A504X1D9	Dullard-like phosphatase domain protein
A0A504XL37	Cyclic nucleotide-binding domain family protein	Q4Q0T5	Paraflagellar rod component
A0A6L0XP96	Regulatory_subunit_of_protein_kinase_a-like_protein	A4HTH1	Inositol phosphosphingolipid phospholipase C-Like
Q4Q0D3	Putative chaperone protein DNAj	E9AV54	DUF4833 domain-containing protein
Q4Q6V0	Putative CAS/CSE/importin domain protein	A0A3Q8IE32	Antimony resistance marker of 58 kDa
Q8IFU8	Glutathionylspermidine synthase preATP-grasp family protein	E9APL7	HIT domain-containing protein
A4I7N8	U6 snRNA-associated Sm-like protein LSm4	A0A3S7WQQ4	Uncharacterized protein family (UPF0160), putative
A4HZH7	Putative cytochrome c oxidase subunit VI	Q4QDS1	Proline--tRNA ligase (EC 6.1.1.15) (Prolyl-tRNA synthetase)
A0A6J8F5T1	Nudix_hydrolase_2_putative/GeneDB:LmjF.05.0090	A4I752	Putative lipase (EC 3.1.1.3)
Q4Q821	Vacuolar proton pump subunit B (V-ATPase subunit B)	E9B2W9	Dynein heavy chain 7, axonemal
A0A504WUD0	RuvB-like helicase (EC 3.6.4.12)	E9ATX4	Putative phosphoribosylpyrophosphate synthetase
Q4Q5G2	Putative proteasome regulatory non-ATP-ase subunit	Q4Q8L3	Leishmanolysin (EC 3.4.24.36)
Q4QCE4	Putative ATP-binding cassette protein subfamily E,member 1	Q4Q289	Proteasome subunit beta
E9AC38	Putative fatty acyl CoA syntetase 1 (EC 6.2.1.3)	A0A3S5H6Y2	Ubiquitin-fold modifier 1
A0A6L0WU94	Ubiquitin-like modifier-activating enzyme 5	Q4Q7A2	Succinyl-CoA:3-ketoacid-coenzyme A transferase (EC 2.8.3.5)
E9AUQ9	Putative calpain-like cysteine peptidase (EC 3.4.22.-)	A0A3S7XBB9	Oxoglutarate dehydrogenase (succinyl-transferring) (EC 1.2.4.2)
A0A6J8FD12	Transketolase (EC 2.2.1.1)	A0A6J8FC72	Hypothetical protein conserved
A0A6J8FEU0	60S acidic ribosomal protein P0	A0A6L0XVT7	Small nuclear ribonucleoprotein E (snRNP-E) (Sm protein E)
A0A6L0XFP1	Protein transport protein Sec24C putative	A4IE04	Haloacid dehalogenase-like hydrolase putative
Q4Q449	Putative mitogen activated protein kinase	A0A504X686	DREV methyltransferase family protein
Q4QHC9	Nuclear transport factor 2 (NTF-2)	A0A3S7WP36	LSM domain containing protein, putative
E9AMQ7	Phosphoglycerate mutase family protein	E9APT8	Putative immunodominant antigen (Tc40 antigen-like)
Q4Q270	Leucine carboxyl methyltransferase 1 (EC 2.1.1.233)	E9AST1	Proteasome subunit alpha type
A4HX10	Tyrosyl or methionyl-tRNA synthetase-like protein (EC 6.1.1.-)	E9AUV8	Putative aminoacylase (EC 3.5.1.14)
E9AQA0	NAD-specific glutamate dehydrogenase (EC 1.4.1.2)		

Table B.9. All Proteins Identified in *L.Major* and *L.Infantum*

UniProtKB Code	**Identified Protein**	**UniProtKB Code**	**Identified Protein**
A0A3S7X4X3	ADP-ribosylation factor	Q9NGR0	6-phosphogluconate dehydrogenase, decarboxylating (EC 1.1.1.44)
E9AC48	Sidoreflexin	A4HZB2	Phosphotransferase (EC 2.7.1.-)
E9ALJ0	Nucleosome assembly protein (NAP)	E9ALW4	Putative ribosomal protein L1a
Q4Q9A0	J domain-containing protein	E9B1S0	H(+)-exporting diphosphatase (EC 7.1.3.1)
A0A504X8M0	Cation diffusion facilitator transporter family protein	Q4Q361	D-isomer specific 2-hydroxyacid dehydrogenase-like protein
A0A6L0XMH7	Unconventional_ myosin	A0A6J8FNW4	Tubulin beta chain
E9AP80	Putative 60S ribosomal protein L18	A0A3Q8ICH0	60S ribosomal protein L21, putative (60S_ribosomal_protein_L21_putative/GeneDB:LmjF.16.0460/ GeneDB:LmjF.34.3650)
Q4Q8N2	Sacchrp_dh_NADP domain-containing protein	A0A504WWC4	Ubiquitin-related modifier 1 homolog
Q4Q3H3	Putative eukaryotic translation initiation factor 5	E9AZL5	40S ribosomal protein S26
A0A3Q8IDW6	60S ribosomal protein L36, putative	A0A504XNX0	C-8_sterol_isomerase-like_protein/GeneDB:LmjF.29. 2140

Table B.10. All Proteins Identified in *L.Major* and *L.Donovani*

UniProtKB Code	**Identified Protein**	**UniProtKB Code**	**Identified Protein**
Q4Q1R2	Clathrin heavy chain	Q4Q2D5	Ubiquitin carboxyl-terminal hydrolase (EC 3.4.19.12)
A4I3Y3	Putative acyl-CoA dehydrogenase (EC 1.3.99.3)	Q4Q5T1	WD REPEATS REGION domain-containing protein
E9AL77	MPN domain-containing protein	E9AYX8	Putative small GTP-binding protein Rab1
E9ARU9	Uncharacterized protein	A0A504X3J8	Alkyldihydroxyacetonephosphate_synthase/ GeneDB:Lm jF.30.0120
E9AKJ6	Surface antigen-like protein	E9AD26	Putative calpain-like cysteine peptidase (EC 3.4.22.-)
A0A6L0XHB0	Heat shock 70-related_protein 1 mitochondrial precursor putative	A0A6L0XYI1	Coatomer subunit beta'

Table B.11. All Proteins Identified in *L.Infantum* and *L.Donovani*

UniProtKB Code	Identified Protein	UniProtKB Code	Identified Protein
E9B063	Dihydrolipoamide acetyltransferase component of pyruvate dehydrogenase complex (EC 2.3.1.-)	A4HRH2	Putative long-chain-fatty-acid-CoA ligase (EC 6.2.1.3)
E9ADT0	Putative U-box domain protein	A0A3Q8IED6	Thioredoxin-like family protein (Tryparedoxin 1, putative) (Tryparedoxin_1_putative/GeneDB:LmjF.29.1160)
A0A6L0XJF0	Guanine nucleotide-binding protein subunit beta-like protein	Q5QQ43	Isopentenyl-pyrophosphate isomerase (Putative isomerase (EC 5.3.3.2))
A4I076	Putative NADP-dependent alcohol dehydrogenase (EC 1.1.1.1)	Q9XZY0	Surface antigen-like protein
Q4Q3L5	Coatomer subunit beta'	A0A6L0XSD0	Threonine--tRNA ligase (EC 6.1.1.3) (Threonyl-tRNA synthetase)
A0A6L0X1D8	Eukaryotic translation initiation factor 4 gamma type 2 putative	Q4QDF0	Malate dehydrogenase (EC 1.1.1.37)
Q8WT38	Histone H2A	E9APJ0	Putative chaperonin TCP20
Q4Q1W2	Putative short chain 3-hydroxyacyl-CoA dehydrogenase (EC 1.1.1.35)	A0A3Q8INK5	3-oxo-5-alpha-steroid 4-dehydrogenase family protein (Enoyl-CoA reductase, putative (EC 1.3.99.5)) (Enoyl-CoA_reductase_putative/GeneDB:LmjF.25.1770)
A0A6L0XL90	Trifunctional_enzyme_alpha_subunit_-_mitochondrial_precursor-like_protein	A4HU34	Mitochondrial RNA binding protein 2
A0A504XM61	Emp24/gp25L/p24 family/GOLD protein	E9AW11	Mannose-1-phosphate guanyltransferase (EC 2.7.7.13)
A0A3Q8IK52	Nucleolar protein 10	A0A504X311	Oxidoreductase_putative/GeneDB:LmjF.36.4170 (Zinc-binding dehydrogenase family protein)
A0A504Y865	Dehydrogenase E1 component family protein	A0A6J8F9L4	Hypothetical_protein_conserved
A0A3S7X6B0	Heat shock protein 83-17 (Heat_shock_protein_83-17/GeneDB:LmjF.33.0365)	E9ACC0	Voltage-dependent anion-selective channel
Q9UAB4	Proteasome subunit alpha type	A0A6L0XWT5	Glyceraldehyde-3-phosphate dehydrogenase (EC 1.2.1.12)
E9BRZ9	Short chain dehydrogenase, putative	Q4Q171	Oxoglutarate dehydrogenase (succinyl-transferring) (EC 1.2.4.2)
A0A3Q8IFX2	Fructose-bisphosphate aldolase (EC 4.1.2.13)	A0A504XH18	peptidylprolyl isomerase (EC 5.2.1.8)
Q4QBF5	Putative endoribonuclease L-PSP (Pb5)	A0A7S5YR77	Leishmanolysin (EC 3.4.24.36)
Q4Q7L5	Adenosine kinase (AK) (EC 2.7.1.20) (Adenosine 5'-phosphotransferase)	A0A8F0ZVD2	Heat shock protein 70
E9AP89	Palmitoyltransferase (EC 2.3.1.225)	A0A504WWD3	Cytosolic carboxypeptidase-like protein 5 (ATP/GTP-binding protein-like 5) (Protein deglutamylase CCP5)
A0A504XY16	Ribosomal protein S28e family protein	A0A3S7X782	Small nuclear ribonucleoprotein Sm D2 (Sm-D2) (snRNP core protein D2)
A0A3Q8IL41	Histone H2A	A4HW83	Putative tyrosyl-tRNA synthetase (EC 6.1.1.1)
Q4Q1J6	Putative developmentally regulated GTP-binding protein 1	A0A3Q8IET6	60S ribosomal protein L30 (60S_ribosomal_protein_L30/GeneDB:LmjF.35.0240) (Ribosomal protein L7Ae/L30e/S12e/Gadd45 family protein)
A0A3S7X5D6	T-complex protein 1 subunit epsilon (CCT-epsilon)	Q4QEG9	Cystathionine beta-synthase (EC 4.2.1.22)

(Cont. on next page)

Table B.11. (cont)

UniProt KB Code	Identified Protein	UniProt KB Code	Identified Protein
A0A3S5H6J5	Pretranslocation protein, alpha subunit, putative	A0A1X9JSW4	Rab GDP dissociation inhibitor
E9AGQ5	Putative heat shock protein	Q4Q9F9	Methenyltetrahydrofolate cyclohydrolase (EC 3.5.4.9)
A0A3S7X954	60S ribosomal protein L32 (60S_ribosomal_protein_L32/GeneDB:LmjF.35.2050) (Ribosomal protein L32 family protein)	A0A504X9D8	NAD binding domain of 6-phosphogluconate dehydrogenase family protein
A0A3Q8IJ66	non-specific serine/threonine protein kinase (EC 2.7.11.1)	E8NHE3	Putative 60S ribosomal protein L28 (Uncharacterized protein LmxM_11_1130a_1)
E9AQQ4	Sucrose-phosphate synthase-like protein (EC 2.4.1.14)	A0A504XGP8	Short chain dehydrogenase family protein (Short_chain_dehydrogenase_putative/Pfam:PF00106)
E9BA59	Uncharacterized protein	E9AGG4	Soluble promastigote surface antigen PSA-50S (Surface_antigen_protein_2_-_putative)
E9AQL0	Uncharacterized protein	A0A6J8FNV8	24-dienoyl-coa reductase fadh1 putative/GeneDB:Lm jF.33.0830
A0A504XRL9	Myb-like domain-containing protein	A0A3Q8IDH3	D-lactate dehydrogenase-like protein (FAD linked oxidase, C-terminal domain family protein)
E9BIZ3	Stress responsive A/B Barrel Domain family protein (Stress responsive A/B Barrel Domain, putative) (Stress_responsive_A/B_Barrel_Domain_putative/Pfam:PF07876)	E9BHM0	Aldehyde dehydrogenase, mitochondrial
A0A3S7X1J0	Conserved plasma membrane protein	A0A504Y515	Ribonucleoside-diphosphate reductase (EC 1.17.4.1)
Q4QAB9	HMG CoA synt N domain-containing protein	E9ACG7	Multifunctional fusion protein [Includes: L-glutamate gamma-semialdehyde dehydrogenase (EC 1.2.1.88) (L-glutamate gamma-semialdehyde dehydrogenase); Delta-1-pyrroline-5-carboxylate dehydrogenase (P5C dehydrogenase)]
A0A504X640	Ribosomal L27e family protein	A4IAI3	Uncharacterized protein
A0A836GCC3	Protein transport protein SEC23	A4HU18	Elongation factor-1 gamma
A4I291	Putative thimet oligopeptidase (EC 3.4.24.15)	Q4Q6Z5	Glyceraldehyde-3-phosphate dehydrogenase (EC 1.2.1.12)

Table B.12. All Proteins Identified only in *L.Tropica*

UniProtKB Code	Identified Protein	UniProtKB Code	Identified Protein
A4HU23	Uncharacterized protein	A4I4T9	Uncharacterized protein
A0A6J8FDD5	Intraflagellar transport protein 172 putative/GeneDB:LmjF.21.0980	A0A504Y3H9	Biotin-requiring enzyme family protein
Q4QD30	C2 domain-containing protein	A0A3S7WX98	Intraflagellar transport protein-like protein (Intraflagellar_transport_protein-like_protein/Gen eDB:LmjF.22.1370)
B9UX70	Myosin	A0A3Q8IJ16	AMMECR1, putative (AMMECR1_putative/Pfam:PF01871)
Q4Q0C1	Rab-GAP TBC domain-containing protein	A0A3S7X2Q8	Nucleosome assembly protein (NAP) family protein (Nucleosome assembly protein (NAP), putative) (Nucleosome_assembly_protein_(NAP)_putative /Pfam:P F00956)
A0A504Y3V5	Uncharacterized protein	A0A504XBE1	Formate--tetrahydrofolate ligase (EC 6.3.4.3)
E9ARX7	Kinesin-like protein	E9AVR2	NADH-cytochrome b5 reductase (EC 1.6.2.2)
A4I1F5	Putative cytochrome c oxidase VII	A0A3Q8INE9	40S ribosomal protein S25
A4IA52	Uncharacterized protein	Q4QE21	Activator of Hsp90 ATPase N-terminal domain-containing protein
E9AUG4	ER membrane protein complex subunit 10	A4I8C5	AMP-activated protein kinase glycogen-binding domain-containing protein
A0A504Y1S7	Tubulin binding cofactor C family protein	E9BEZ5	Guanine nucleotide-binding protein subunit beta-like protein
Q4QIB4	Translation initiation factor-like protein	Q4Q3M9	Uncharacterized protein
Q4QGA2	Uncharacterized protein	E9ASE9	Proteasome subunit beta
E9BJ67	Snf7,putative (Snf7 putative / Pfam:PF03357)	A0A3S7X9Z4	Calmodulin-like protein (Calmodulin-like protein/ GeneDB:LmjF.35.3890)
Q4QJ54	Leucine zipper transcription factor-like protein 1	Q4QDJ1	Putative RNA binding protein
A0A504XLV8	Protein kinase domain family protein	E9AYV2	Putative calpain-like cysteine peptidase
A4I2N5	Putative calpain-like cysteine peptidase (EC 3.4.22.-)	A0A3S7WRR3	SKP1-like protein (SKP1-like protein/GeneDB:LmjF.11.1210) (Skp1 family, dimerization domain protein)
A4HYA0	ATP-binding cassette protein subfamily F, member 2	Q4QHR3	Mitochondrial RNA binding protein 2
Q4Q3C1	Elongation factor 1-beta (Elongation factor 1B beta) (Putative translation elongation factor 1-beta)	Q4QFL5	Bifunctional NAD(P)H-hydrate repair enzyme (Nicotinamide nucleotide repair protein) [Includes: ADP-dependent (S)-NAD(P)H-hydrate dehydratase (EC 4.2.1.136) (ADP-dependent NAD(P)HX dehydratase); NAD(P)H-hydrate epimerase (EC 5.1.99.6) (NAD(P)HX epimerase)]
A0A504XMT4	C2 domain-containing protein	E9B2S6	Putative ATP-dependent zinc metallopeptidase (EC 3.4.24.-)
Q27685	Phosphoglycerate kinase, glycosomal (Phosphoglycerate kinase C) (EC 2.7.2.3)	A0A504X297	Nucleoside transporter family protein
A4I1K8	Uncharacterized protein	Q4Q7N0	Putative 4-methyl-5(Beta-hydroxyethyl)-thiazole monophosphate synthesis protein
A5JV97	Mitochondrial tryparedoxin	E9AVG6	Proteasome subunit alpha type
E9APD3	Proteasome regulatory ATPase subunit 2, putative	A0A504XGZ0	Uncharacterized protein

(Cont. on next page)

Table B.12. (cont.)

UniProt KB Code	Identified Protein	UniProt KB Code	Identified Protein
A4HY48	Kinesin-like protein	A0A3Q8ILE8	Beta-Casp domain family protein (Cleavage and polyadenylation specificity factor, putative)
A4I9C2	Transcription elongation factor-like protein	A4I4I4	ATP-dependent Clp protease subunit, heat shock protein 100 (HSP100), putative
E9AHR9	Uncharacterized protein	E9AC56	Putative mitochondrial processing peptide beta subunit (EC 3.4.24.64)
A0A3Q8IF98	Uncharacterized protein	E9AQD5	Uncharacterized protein
A4I5D2	Putative ubiquitin fusion degradation protein 2	A4I2S5	Guanine nucleotide-binding protein subunit beta-like protein
E9BU19	Succinate--CoA ligase [ADP-forming] subunit beta, mitochondrial (EC 6.2.1.5)	Q4QHC3	Putative small GTP-binding protein Rab11
A4HYX0	DUF1935 domain-containing protein	Q8I8E1	Protein disulfide-isomerase (EC 5.3.4.1)
Q4Q3I5	IFT81 calponin homology domain-containing protein	A0A6J8FBI7	Ubiquitin_hydrolase_putative/GeneDB:LmjF.15.1300
E9AVP9	Prefoldin 5-like protein	Q4Q3I7	CS domain-containing protein
A4I7R1	Uncharacterized protein	A4IB15	ARMET C-terminal domain-containing protein
A0A6J8FC81	Protein_kinase_putative/GeneDB:LmjF.25.1520	E9AN41	Protein transport protein SEC23
A4HXG5	Putative META domain containing protein	E9AQH1	Eukaryotic translation initiation factor 1A,putative
P14700	Membrane antigen containing repeating peptides (Clone 39)	E9AHA3	Aminopeptidase P1 putative
Q4Q760	Uncharacterized protein	A4ICM0	Putative short chain 3-hydroxyacyl-CoA dehydrogenase (EC 1.1.1.35)
E9AG49	Putative arginine N-methyltransferase (EC 2.1.1.-)	A0A6L0XBQ0	Nucleosome assembly protein (NAP) putative
E9ALV6	EF-hand domain-containing protein	A0A504XJD8	DEAD/DEAH box helicase family protein
Q4QFF3	Putative replication Factor A 28 kDa subunit	E9B1J0	Putative calpain-like cysteine peptidase
Q4Q4N5	Myosin XXI	Q4Q636	Putative 3,2-trans-enoyl-CoA isomerase,mitochondrial (EC 5.3.3.8)
Q4Q970	Heat shock protein 70-related protein	Q4QDU3	UTP--glucose-1-phosphate uridylyltransferase (EC 2.7.7.9)
Q4QJI9	RNA helicase (EC 3.6.4.13)	A0A504Y112	Isochorismatase family protein
Q4QCJ6	Uncharacterized protein	D2EAU8	Heat shock protein 70 (Heat-shock protein 70 kDa)
E9APX8	Uncharacterized protein	Q4QDM4	Uncharacterized protein
A4HX92	Cystathionine beta-synthase	A0A6L0XDE6	T-complex protein 1 subunit delta
Q4Q5B4	Uncharacterized protein	A0A836FRZ1	Adenylate kinase active site lid domain-containing protein
Q86PL7	Putative small GTP-binding protein	E9AQS9	Coatomer subunit delta
E9B8L4	Guanine nucleotide-binding protein subunit beta-like protein	A0A504Y4S8	Leucine-rich repeat-containing protein 51
E9B6I9	Probable eukaryotic initiation factor 4A	Q07DU7	Peroxidoxin 1
A0A3S5H830	Fibrillarin (Fibrillarin family protein)	A4I2W4	Uncharacterized protein

(Cont. on next page)

Table B.12. (cont.)

UniProtKB Code	Identified Protein	UniProtKB Code	Identified Protein
A0A504Y2T4	Hypothetical protein conserved	A0A6J8FD18	Glycerophosphoryl_diester_phosphodiesterase family putative/Pfam:PF03009
A0A504Y2T4	Hypothetical protein conserved	Q4Q848	CS domain-containing protein
E9B689	Uncharacterized protein	Q4Q794	Putative intraflagellar transport (IFT) protein
A4I1B5	Putative eukaryotic initiation factor 5a	Q4Q889	Arf-GAP domain-containing protein
Q4Q7K6	Putative p22 protein	Q5I5K8	Glyoxalase II
E9B3W0	Uncharacterized protein	E9B0J8	GrpE protein homolog
A0A6L0WPZ1	Proteasome subunit alpha type	Q4QF62	60S acidic ribosomal protein P2
E9ARZ6	Putative nucleosome assembly protein	A4I8X4	Guanylate kinase (EC 2.7.4.8)
A0A504XPR9	Short chain dehydrogenase family protein	A4I4S9	Putative GTP-binding protein
E9ACH4	FYVE-type domain-containing protein	Q4QF36	Glutamine--tRNA ligase (EC 6.1.1.18)
E9AS51	4-coumarate:coa ligase-like protein (EC 6.2.1.12)	A4I048	Mannose-1-phosphate guanylyltransferase (EC 2.7.7.13)
Q4QIC2	Putative small ubiquitin protein	A0A504XD89	NADH:flavin oxidoreductase / NADH oxidase family protein (NADH:flavin_oxidoreductase/NADH_oxidase_putative/ GeneDB:LmjF.12.1130)
Q4Q7S1	Putative phosphatase 2C (EC 3.1.3.16)	E9AHI1	Uncharacterized protein
Q4Q3Q4	Transcription elongation factor-like protein	Q4Q5I2	T-complex protein 1 subunit epsilon (CCT-epsilon)
A0A504XHA2	Microtubule associated protein (MAP65/ASE1 family) protein	E9ARC2	Uncharacterized protein
Q4FX34	Aspartate aminotransferase (Putative aspartate aminotransferase (EC 2.6.1.1))	Q4Q9I8	Glycine cleavage system P protein (EC 1.4.4.2)
A4HS40	Adenylosuccinate lyase	E9AEW4	Threonine--tRNA ligase (EC 6.1.1.3) (Threonyl-tRNA synthetase)
Q4Q5B2	Uncharacterized protein	E9AM64	Uncharacterized protein
Q4Q5F6	CRAL-TRIO domain-containing protein	E9ACJ7	RRM domain-containing protein
P43152	Hypoxanthine-guanine phosphoribosyltransferase (HGPRT) (HGPRTase) (EC 2.4.2.8)	E9ASA8	Guanine nucleotide-binding protein subunit beta-like protein
A4I213	Putative asparagine synthetase a (EC 6.3.1.1)	A4I0R1	Signal recognition particle 54 kDa protein
E9AP92	Peptidase M24 domain-containing protein	E9AT61	Uncharacterized protein
Q4Q259	Elongation factor 2	A4IAT4	3-dehydrosphinganine reductase (EC 1.1.1.102)
Q4Q6Q6	C3H1-type domain-containing protein	E9AFQ3	FYVE-type domain-containing protein

Table B.13. All Proteins Identified only in *L.Major*

UniProtKB Code	Identified Protein	UniProtKB Code	Identified Protein
E9AZ36	ATP-dependent RNA helicase (EC 3.6.4.13)	Q4QIE0	Superoxide dismutase (EC 1.15.1.1)
E9AVY4	Putative 40S ribosomal protein L14	A4I7N3	Putative ras-related rab-4
E9AUU6	Putative 40S ribosomal protein S11 (Ribosomal protein S11 homolog)	Q9U0V4	Uncharacterized protein L7836.08
E9ARI8	Aconitate hydratase (Aconitase) (EC 4.2.1.3)	E9AS23	Malate dehydrogenase (EC 1.1.1.37)
E9APE5	Nucleobase transporter	A4HUM7	Cilia- and flagella-associated protein 45
Q4QD99	Uncharacterized protein	A0A836KG59	Uncharacterized protein
Q9U9R3	Peptidyl-prolyl cis-trans isomerase (PPIase) (EC 5.2.1.8)	Q4Q6F8	Uncharacterized protein
A4I9P1	Elongation factor 1-beta	A4IBV3	Polyadenylate-binding protein (PABP)
O43941	Protein phosphatase-2C	E9AV93	T-complex protein 1 subunit delta
A4ICW8	Elongation factor 2	E9AZ16	Putative intraflagellar transport protein IFT88
Q4QJG7	Trypanothione reductase (EC 1.8.1.12) (N(1),N(8)-bis(glutathionyl)spermidine reductase)	A4ICP1	40S ribosomal protein S18, putative
E9AX51	Putative 60S ribosomal protein L26	Q4Q5Q0	Uncharacterized protein
Q1X7K5	Excreted/secreted protein 27	E9ADA0	DUF4201 domain-containing protein
Q4Q4X1	AMP-activated protein kinase glycogen-binding domain-containing protein	A0A6L0XNB0	Malate dehydrogenase (EC 1.1.1.37)
O00912	Histone H4	A0A836GST7	Aminotransferase class I/classII domain-containing protein
E9ACQ0	Surface antigen-like protein	E9AHX6	Putative polyubiquitin
P90628	Cathepsin L-like protease	E9ALJ9	Aminopeptidase (EC 3.4.11.-)
E9ADA4	CCT-beta	E9AXZ8	Putative 60S ribosomal protein L7
E9AGU0	Phosphoglycerate kinase (EC 2.7.2.3)	A4HXQ8	Uncharacterized protein
Q9BIC6	ATP-dependent 6-phosphofructokinase (ATP-PFK) (Phosphofructokinase) (EC 2.7.1.11) (Phosphohexokinase)	E9ATK7	Paraflagellar rod component
A0A504XNY7	Fumarate hydratase (Fumerase) family protein	E9AZ89	Putative trypanothione synthetase (EC 6.3.1.9)
A0A504YAJ5	Adenylate kinase family protein	P08148	Leishmanolysin
E9AKK2	Dipeptidyl peptidase 3 (EC 3.4.14.4) (Dipeptidyl aminopeptidase III) (Dipeptidyl peptidase III)	Q6S4V7	Dihydrolipoyl dehydrogenase (EC 1.8.1.4)
E8NHP1	WGS CADB00000000 data, contig 73	Q4QI73	Uncharacterized protein
A0A381MN47	Histone H2A	A4I3X4	Vacuolar proton pump subunit B
Q4QIM8	Calcium-transporting ATPase (EC 7.2.2.10)	A4I058	Endoribonuclease L-PSP (Pb5), putative

(Cont. on next page)

Table B.13. (cont.)

UniProtKB Code	Identified Protein	UniProtKB Code	Identified Protein
Q4Q8Z6	Uncharacterized protein	Q9U1E1	Nucleoside diphosphate kinase (EC 2.7.4.6)
A4HX65	Separase	A4I8F2	T-complex protein 1 subunit alpha (CCT-alpha)
Q4Q843	Putative glycoprotein 96-92	Q4QFI9	C2 domain-containing protein
E9AXQ3	Pyruvate dehydrogenase E1 component subunit beta (EC 1.2.4.1)	A4HV06	40S ribosomal protein S5
Q4Q6R3	5-methyltetrahydropteroyltriglutamate--homocysteine S-methyltransferase (EC 2.1.1.14)		

Table B.14. All Proteins Identified only in *L.Infantum*

UniProtKB Code	Identified Protein	UniProtKB Code	Identified Protein
A4I6L8	Putative sodium stibogluconate resistance protein	A4I565	Putative phosphatase 2C
E9AYE0	Proline dehydrogenase (EC 1.5.5.2)	E9AKX0	Uncharacterized protein
E9AEU1	Putative NADH-dependent fumarate reductase (EC 1.3.1.6)	A4I993	Putative carboxypeptidase
A0A6L0XE70	Hypothetical_protein_-_conserved	A0A504X1R9	Aspartate carbamoyltransferase
E9ASX8	Chaperonin HSP60, mitochondrial	E9APR0	Elongation of fatty acids protein (EC 2.3.1.199) (Very-long-chain 3-oxoacyl-CoA synthase)
Q4Q698	Thiolase protein-like protein (EC 2.3.1.16)	E9AYK4	Kynureninase (EC 3.7.1.3) (L-kynurenine hydrolase)
A0A6J8FFW9	Calpain like cysteine peptidase putative	A0A6L0XSU1	p450 reductase putative
E9AEB3	ATP-dependent 6-phosphofructokinase (ATP-PFK) (Phosphofructokinase) (EC 2.7.1.11) (Phosphohexokinase)	E9AP30	SET domain-containing protein
E9AWE0	T-complex protein 1 subunit gamma	A0A3Q8IDP8	Small myristoylated protein 4, putative
A0A6L0XHI9	BRO1-like_domain/ALIX_V-shaped_domain_binding_to_HIV_-_putative	Q4Q6G8	Putative sodium stibogluconate resistance protein
E9ANS9	Mitochondrial SSU ribosomal protein 29	E9ANS8	Proteasome subunit beta
A4HX21	Uncharacterized protein	A4I9I5	Malate dehydrogenase (EC 1.1.1.37)
E9AQ01	Lysine decarboxylase family protein	E9AQB1	60S acidic ribosomal protein P2
E9B0P0	Putative RNA-binding protein	A0A504X2D9	Aminotransferase class I and II family protein
Q18L74	Aspartate aminotransferase (EC 2.6.1.1)	E9AWN2	Ubiquitin carboxyl-terminal hydrolase (EC 3.4.19.12)
A0A6J8FB09	Hypothetical_protein_conserved	E9ASU7	Uncharacterized protein

(Cont. on next page)

Table B.14. (cont.)

UniProtKB Code	Identified Protein	UniProtKB Code	Identified Protein
E9B2J9	Putative RNA binding protein	A4I1U2	Glycine cleavage system P protein (EC 1.4.4.2)
Q4Q6K7	Mevalonate kinase (MK) (EC 2.7.1.36)	A4IE50	Paraflagellar rod component
E9APX4	Ubiquitin-60S ribosomal protein L40	A4IBK1	Probable eukaryotic initiation factor 4A (ATP-dependent RNA helicase eIF4A)
A0A504YC88	Guanine nucleotide-binding protein subunit beta-like protein	D0VX00	coproporphyrinogen oxidase (EC 1.3.3.3)
A0A6J8FCK2	Hypothetical_protein_conserved	A0A6L0WP97	Squalene monooxygenase (EC 1.14.14.17)
E9ANQ5	WD REPEATS REGION domain-containing protein	E9BNK4	Ubiquitin-conjugating enzyme protein, putative
E9AEX1	Arginase (EC 3.5.3.1)	E9BTP4	Uncharacterized protein
Q4QHJ8	Putative mitogen-activated protein kinase	Q27684	Phosphoglycerate kinase, cytosolic (EC 2.7.2.3) (Phosphoglycerate kinase B)
E9AGF0	Alanine transaminase (EC 2.6.1.2)	A4HSG7	Uncharacterized protein
A4I574	Aspartate--tRNA ligase	A5JV94	Trypanothione synthetase
E9B0Q4	Pyridoxal kinase (EC 2.7.1.35)	E9ARF7	UDP-galactopyranose mutase (EC 5.4.99.9)
E9AV58	Hypoxanthine phosphoribosyltransferase (EC 2.4.2.8)	A4HYI8	Putative ATP-dependent zinc metallopeptidase
A0A8B6T4S3	Mannose-6-phosphate isomerase (EC 5.3.1.8)	A0A6L0XWR5	HECT-type E3 ubiquitin transferase (EC 2.3.2.26)
A4HUX3	Putative aminopeptidase	A0A504XDW3	D-lactate_dehydrogenase-like_protein/GeneDB:LmjF. 27.2020/GeneDB:LmjF.29.0280 (FAD linked oxidase, C-terminal domain family protein)
E9AR65	Putative ATP-dependent RNA helicase	A0A6L0XL78	Nuclear segregation protein putative
E9AGQ4	Rab7 GTP binding protein_putative/ GeneDB:LmjF.18. 0890 (Ras family protein)	Q71S85	Superoxide dismutase (EC 1.15.1.1)
Q4VKK7	Mitochondrial tryparedoxin peroxidase	Q4JHN0	Stress-induced protein 1 (Stress-induced protein sti1) (Stress-induced_protein_sti1/GeneDB:LmjF.08 .1110) (TPR repeat family protein)
Q4Q277	Histone H4	A4IAE6	Phosphomannomutase-like protein
E9AL95	Flavoprotein subunit-like protein	A4HVP9	Nucleobase transporter
A0A836GU99	Separase	A4HSY0	Putative acyl-coenzyme a dehydrogenase (EC 1.3.99.3)
P39050	Trypanothione reductase (TR) (EC 1.8.1.12) (N(1),N(8)-bis(glutathionyl)spermidine reductase)	A0A504XSL4	Carnitine/choline_acetyltransferase_putative/Gene DB:LmjF.29.1310 (Choline/Carnitine o-acyltransferase family protein)
A0A6J8FFA7	Hydrophilic_acylated_surface_protein_b/GeneDB:Lmj F.23.1060/GeneDB:LmjF.23.1070	A0A0U2N0C0	Rab1
A0A504WXZ6	Calpain cysteine protease family protein (Cytoskeleton-associated protein CAP5.5 putative/G eneDB:LmjF.31.0440)	E9ALK3	Putative GTP-binding protein
A4I384	Proteasome subunit beta	A0A836GYM4	Inosine-5'-monophosphate dehydrogenase (EC 1.1.1.205)
Q4QH08	ATPase ASNA1 homolog (EC 3.6.-.-) (Arsenical pump-driving ATPase homolog) (Arsenite-stimulated ATPase)	E9ANJ4	mRNA (guanine-N(7))-methyltransferase (EC 2.1.1.56)
P40285	Histone H3		

Table B.15. All Proteins Identified only in *L.Donovani*

UniProtKB Code	Identified Protein	UniProtKB Code	Identified Protein
Q4QHW4	Putative serine peptidase family S51, peptidase E	Q4QJ76	Guanine nucleotide-binding protein subunit beta-like protein
A0A3S7X885	NAD dependent epimerase/dehydratase family protein	A0A504X0P2	Mitochondrial RNA_binding complex 1 subunit putat ive/GeneDB:LmjF.31.0630
A0A6L0WQB3	Adenylosuccinate synthetase (AMPSase) (AdSS) (EC 6.3.4.4) (IMP--aspartate ligase)	A4I504	FoP duplication domain-containing protein
A4HZE4	Uncharacterized protein	A4HYQ5	CTP synthase (EC 6.3.4.2) (UTP--ammonia ligase)
A4I077	Cytochrome c oxidase subunit 10, putative (EC 1.9.3.1)	A4IA34	DNA/RNA-binding protein Alba-like domain-containing protein
P69201	Ubiquitin-60S ribosomal protein L40	E9BGP3	Fumarate hydratase (EC 4.2.1.2)
Q4Q9B1	Putative aspartate--ammonia ligase (EC 6.3.1.1)	E9ATA6	Putative 60S ribosomal protein L22
A0A504X1I4	NIPSNAP family protein (NIPSNAP_putative/Pfam:PF07978)	A4HZF7	Uncharacterized protein
E9AU90	Similarity to endo-1-like protein	A4ID97	Thioredoxin domain-containing protein
A0A836KKC1	Thioredoxin domain-containing protein	E9AMI0	Uncharacterized protein
A4I7B5	Phosphoglycan beta 1,3 galactosyltransferase 5	Q4QCD7	19S proteasome regulatory subunit
A0A6L0WS83	CPSF A subunit region containing protein putative	B5APK4	Triosephosphate isomerase (EC 5.3.1.1)
E9AQZ4	Cystathionine beta-synthase (EC 4.2.1.22)	E9AMS8	Prolyl endopeptidase (EC 3.4.21.-)
A4HXI2	Alpha/beta hydrolase family protein	E9AMC6	Putative small ubiquitin protein
Q9GNZ8	Proteasome subunit alpha type	A0A504X883	Uncharacterized protein
Q4QEZ7	Protein tyrosine phosphatase PRL-1 (EC 3.1.3.48)	E9AQ24	Lysine--tRNA ligase (EC 6.1.1.6) (Lysyl-tRNA synthetase)
A0A3S5H6N5	40S ribosomal protein S12	Q4QA84	Electron transfer flavoprotein subunit beta (Beta-ETF)
E9BRS2	40S ribosomal protein S3a	A0A6L0XXI5	Protein kinase putative
A0A504X4I9	NLI interacting factor-like phosphatase family protein	Q4Q7J9	Uncharacterized protein
Q4QEI2	ATP pyrophosphate-lyase (Adenylyl cyclase)	A2CIQ8	Fumarate hydratase (EC 4.2.1.2)
Q4QBS0	Putative NADH-cytochrome b5 reductase (EC 1.6.2.2)	Q4QIZ9	Proteophosphoglycan ppg4
A0A5B8YYQ3	Phosphopyruvate hydratase (EC 4.2.1.11)	A0A6L0XS13	Glycerol kinase (EC 2.7.1.30)
E9AFM5	Uncharacterized protein	A4I4S8	Uncharacterized protein
Q4Q1B2	Fibrillarin	Q4QEM2	Paraflagellar rod protein 2C
E9ATS7	Methylthioribose-1-phosphate isomerase (M1Pi) (MTR-1-P isomerase) (EC 5.3.1.23)	A0A3Q8IIZ5	Enoyl-CoA hydratase/isomerase family protein (Enoyl-CoA hydratase/isomerase-like protein) (Enoyl-CoA_hydratase/isomerase-like_protein/GeneDB:LmjF.29.2310)
A0A6L0XVZ8	RNA-binding_protein_42_(RNA-binding_motif_protein_42)_-_putative	Q4Q4Z3	Putative proteasome regulatory non-ATP-ase subunit
A0A504XPT0	Serine/threonine-protein kinase TOR (EC 2.7.11.1)	Q4QBY1	BAG domain-containing protein
A4IBT3	Putative PRP8 protein homologue	A4IC62	PRORP domain-containing protein

(Cont. on next page)

Table B.15. (cont.)

UniProtKB Code	Identified Protein	UniProtKB Code	Identified Protein
A4I3S2	X-Pro dipeptidyl-peptidase (S15 family) protein	Q4QDG9	Uncharacterized protein
E9APE3	Adenylosuccinate synthetase (AMPSase) (AdSS) (EC 6.3.4.4) (IMP--aspartate ligase)	A0A504Y6F1	Uncharacterized protein
E9B654	Pru domain-containing protein	E9ANH2	Uncharacterized protein
A0A504X8Z3	Prohibitin	E9BC31	Uncharacterized protein
A0A504XLI3	Thioredoxin family protein	A0A3Q8IT45	RNA binding protein, putative (RNA recognition motif family protein) (RNA_binding_protein_putative/GeneDB:LmjF.32.0750)
A4HYC1	Phenylalanine--tRNA ligase (EC 6.1.1.20) (Phenylalanyl-tRNA synthetase beta subunit)	E9AK23	Uncharacterized protein
A0A836KD97	Pyruvate kinase (EC 2.7.1.40)	E9AUD8	Cytochrome c oxidase subunit I
E9AG04	Putative GTP-binding protein	A0A504WWZ5	Splicing factor 3B subunit 1 family protein
A0A6J8FNK4	Importin_beta-1 subunit_putative/GeneDB:LmjF.34.0 490	A0A3S5H7U1	Tubulin beta chain
A4I9L9	Putative importin beta-1 subunit	E9AWY5	Putative mismatch repair protein
E9AD61	Uncharacterized protein	Q4QHL1	Dehydrogenase-like protein
A0A3G4YQ88	60S acidic ribosomal protein P2	A0A6L0XKU2	Hypothetical_protein_-_conserved
E9ACN8	Amidohydrolase 3 domain-containing protein	A0A504XZW5	Aconitate hydratase (Aconitase) (EC 4.2.1.3)
Q4QCU9	Acireductone dioxygenase (Acireductone dioxygenase (Fe(2+)-requiring))	A4IAR1	Uncharacterized protein
A0A3Q8IJQ7	Dynein light chain roadblock	Q4Q8A8	Putative hydrolase, alpha/beta fold family
A0A3Q8IUI3	Uncharacterized protein	E9ADE3	BRO1 domain-containing protein
E9APJ2	Pyrroline-5-carboxylate reductase (EC 1.5.1.2)	E9AEI8	Coatomer subunit zeta
A0A3Q8IH05	Malate dehydrogenase (EC 1.1.1.37)	Q4QEE6	Phenazine biosynthesis-like protein
Q4Q286	Cell differentiation protein-like protein	A0A6L0XK73	Translocon-associated protein beta (TRAPB)/Domain of uncharacterized function DUF11/CARDB putative
A4HU29	Putative eukaryotic translation initiation factor 2 subunit	Q4QJF7	Skp1 domain-containing protein
E9BR42	RuvB-like helicase (EC 3.6.4.12)	E9BUF2	Uncharacterized protein
E9AC37	Uncharacterized protein	Q4QAH5	Uncharacterized protein
A0A504XZ66	Intra-flagellar transport protein 57 family protein	A0A504X156	Peptidylprolyl isomerase (EC 5.2.1.8)
E9B0J0	Uncharacterized protein	E9ALP7	Putative paraflagellar rod protein 1D
A0A836GVP8	PCI domain-containing protein	E9APJ1	Putative 60S ribosomal protein L44
A0A6L0XNG6	2-oxoisovalerate dehydrogenase subunit alpha (EC 1.2.4.4)	E9BTS3	Chaperonin HSP60, mitochondrial
E9BJB9	Small nuclear ribonucleoprotein protein, putative	E9ANQ6	Putative eukaryotic release factor 3

(Cont. on next page)

Table B.15. (cont.)

UniProt KB Code	Identified Protein	UniProt KB Code	Identified Protein
A4HSE3	RNA helicase (EC 3.6.4.13)	A0A6J8FQA7	60S ribosomal protein L29
A0A6L0XML6	Nucleolar protein putative	A0A504XDQ2	Proteasome activator PA28 C-terminal domain-containing protein
E9ARM6	Putative rab7 GTP binding protein	E9B5K8	ER membrane protein complex subunit 1
D0AB09	Glycylpeptide N-tetradecanoyltransferase (EC 2.3.1.97)	E9AJV2	Helicase ATP-binding domain-containing protein
E9ANW2	NOT2/NOT3/NOT5 C-terminal domain-containing protein	Q4Q5M1	Uncharacterized protein
A0A6L0WP94	D-3-phosphoglycerate dehydrogenase-like protein	A0A504X2M5	Zinc_finger_protein_family_member_putative /GeneDB:LmjF.36.0730
A0A6J8FDZ2	Calpain-like cysteine peptidase putative/ GeneDB:L mjF.27.0490/GeneDB:LmjF.27.0500/G eneDB:LmjF.27.0510	A0A836KG08	Protein kinase domain-containing protein
A4HWN7	Phosphodiesterase (EC 3.1.4.-)	Q4Q9H4	Putative 60S ribosomal protein L7
A0A3Q8IUV8	H(+)-transporting two-sector ATPase (EC 7.1.2.2)	A4IDZ7	Putative kinetoplast-associated protein