MASS SPECTROMETRY-BASED COMPARATIVE PROTEOMIC ANALYSIS OF DRUG RESISTANT AND NONRESISTANT STRAINS OF PARASITE *Trichomonas vaginalis*

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

MASTER OF SCIENCE

in Chemistry

by
Begüm ÖZYAĞCI

December 2018
İZMİR

We approve the thesis of **Begüm ÖZYAĞCI**

Examining Committee Members:

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

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

Doc. Dr. Ferda SOYER DÖNMEZ
Department of Molecular Biology and Genetics, İzmir Institute of Technology

28 December 2018

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

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

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

ACKNOWLEDGMENTS

First of all, I would like to express my deepest thanks to Prof. Dr. Talat YALÇIN for his guidance, understanding, encouragement, and motivation throughout this thesis. It was a great opportunity and honour for me to work at Biological Mass Spectrometry and Proteomics Laboratory. I consider myself lucky to have a supervisor like him because he has expertise not only on mass spectrometry but also sharing his life experience whenever I need. I would like to thank to Prof. Dr. Ahmet ÖZBİLGİN for his support and understanding. I am also very grateful to İbrahim Çavuş for his cooperation and patience for isolation of samples throughout the study.

I am appreciated to The Scientific and Technological Research Council of Turkey (TÜBİTAK) for the financial support with research project (Project No:215Z257)

Further, I would exclusively like to thank to Dr. Melike DİNÇ for her help, support, contributions, and sincere. Also, I would particularly like to Dr. Melda Z. GÜRAY for her sharing her experience and knowledge, motivation, kindness. My special thanks go to my lovely sister Nazlı GÜVENÇ for her companionship, moral support, understanding, favour on the way we walk during this tough period. Besides, I would like to thank Seren ŞEN for her friendship, helps, cheer and, warmth. It has been a pleasure working with my all laboratory friends and I am grateful to each of them for memorable moments and sharing.

Most importantly, I owe my endless gratitude to my family; my mother Nurgün ÖZYAĞCI, my father Ahmet R. ÖZYAĞCI and, my brother Hazar ÖZYAĞCI. I feel like luckiest person in the world thanks to their love, compassion, motivation and extremely patience to me through all my life.

ABSTRACT

MASS SPECTROMETRY-BASED COMPARATIVE PROTEOMIC ANALYSIS OF DRUG RESISTANT AND NONRESISTANT STRAINS OF PARASITE *Trichomonas vaginalis*

Trichomonas vaginalis (*T. vaginalis*) causes sexually transmitted disease, trichomoniasis. Acute infections can result in cervical cancer, pelvic inflammatory disease, HIV-1 infection, preterm births, and low birth weight. Metronidazole, an antibiotic, is the standard treatment of the disease. However, in some cases *T. vaginalis* shows resistance to metronidazole thus treatment fails. Nevertheless, resistance mechanism of the parasite is not fully understood.

Mass spectrometry has become an important tool in proteomic analysis with the emergence of soft ionization techniques instead of traditional protein identification and sequencing methods. In this study, a comparative gel-free proteomic analysis based on mass spectrometry with soft ionization technology was performed in two sets for resistant and non-resistant strains of *T. vaginalis* parasite isolated from clinical cases. Total proteins were digested by in solution digestion method and separated with high pH reversed-phase liquid chromatography. After fractions were concatenated, each fraction of sample was analyzed by reversed-phase liquid chromatography electrospray ionization tandem mass spectrometry (LC-ESI-MS/MS) and proteins were identified by database search.

Common and differential proteins between resistant and two sensitive trichomoniasis samples were compared according to their molecular function. Results indicate that ferredoxin 5 and hydroxyl amine reductase are differential proteins with iron-sulfur cluster binding activity identified for only resistant strain. Thioredoxin reductase, alcohol dehydrogenase, superoxide dismutase, pyruvate:ferredoxin oxidoreductase are studied proteins in the literature and also identified as common proteins in all strains for this study. These proteins might have a role in drug resistance mechanism of *T. vaginalis*.

ÖZET

Trikomonas vajinalis PARAZİTİNİN İLACA DİRENÇLİ VE DİRENÇSİZ SUŞLARININ KÜTLE SPEKTROMETRESİ TEMELLİ KARŞILAŞTIRMALI PROTEOMİK ANALİZİ

Trikomonas vajinalis (*T. vajinalis*) cinsel yolla bulaşan trikomoniyaz hastalığına sebep olur. Akut enfeksiyonları servikal kanser, pelvik inflamatuar hastalığı, HIV1 enfeksiyonu, erken doğum ve düşük doğum ağırlığı ile sonuçlanabilir. Bir antibiyotik olan metronidazol, hastalığın standart tedavisidir. Yine de bazı durumlarda *T. vajinalis* metronidazole direnç gösterir, bu yüzden tedavi başarısız olur. Buna rağmen parazitin direnç mekanizması tam olarak anlaşılamamıştır.

Kütle spektrometrisi, geleneksel protein tanımlama ve sıralama yöntemleri yerine yumuşak iyonizasyon tekniklerinin ortaya çıkmasıyla proteomik analizde önemli bir araç haline gelmiştir. Bu çalışmada, klinik vakalardan izole edilen *T. vajinalis* parazitinin dirençli ve dirençsiz suşları için iki set halinde yumuşak iyonizasyon teknolojisi ile kütle spektrometresine dayalı karşılaştırmalı jel içermeyen proteomik analizi yapıldı. Toplam proteinler, çözelti içinde parçalama methoduyla parçalandı ve yüksek pH ters fazlı sıvı kromatografisiyle ayrıldı. Fraksiyonlar birleştirildikten sonra, numunenin her bir fraksiyonu ters fazlı sıvı kromatografisi elektrosprey iyonizasyon tandem kütle spektrometresi (LC-ESI-MS / MS) ile analiz edildi ve proteinler veri tabanı araştırması ile tanımlandı.

Dirençli ve iki dirençsiz trikomoniasis örnekleri arasındaki ortak ve farklanan proteinler moleküler fonksiyonlarına göre karşılaştırıldı. Sonuçlar, ferredoxin 5 ve hidroksilamin reduktazın yalnızca dirençli suş için belirlenmiş diferansiyel proteinler olduğunu göstermektedir. Tioredoksin redüktaz, alkol dehidrojenaz, süperoksit dismutaz, piruvat:ferredoksin oksidoredüktaz, literatürde incelenen proteinlerdir ve ayrıca bu çalışma için bütün suşlarda ortak proteinler olarak tanımlanmaktadır. Bu proteinlerin, *T. vaginalis*'in ilaç direnç mekanizmasında bir rolü olabilir.

TABLE OF CONTENTS

LIST OF FIGURES

Figure **Page**

LIST OF TABLES

Table **Page**

CHAPTER 1

MASS SPECTROMETRY AND PROTEOMICS

1.1. Mass Spectrometry

Mass spectrometry (MS) technique is based on examination of gas-phase ions using mass spectrometer which measures mass-to-charge ratio (m/z) of gaseous ions. Fragmentation mechanism of gaseous ions provides information about molecular structure and allows quantification of atoms and molecules.

History of MS extends back a long time. Joseph John Thomson found out charge-to-mass ratio (e/m) of an electron by deflecting cathode rays in the magnetic and electric field of high-vacuum cathode ray tube in 1897. He estimated the mass of electron indirectly and took 1906 Nobel Prize in Physics with his discovery. These experiments lay the foundations of mass spectrometry. Thomson built first a mass spectrometer for mass measurement of charged atoms with the assistance of Francis Aston who improved Thomson's instrument further to discover isotopes and studied to enhance resolving power at the beginning of 20^{th} century. Mass spectrometry was introduced commercially to the science world by 1940s.

McLafferty, Klaus Biemann, and Carl Djerassi enlightened fragmentation mechanisms of unknown organic molecules to determine structural information. Researches of these three scientists are the cornerstones of biological application of MS. Although small organic molecules were started to be analyzed by 1988s, gas phase ionization of large macromolecules like larger proteins, complex carbohydrates, and nucleic acids without decomposition and higher fragmentation was a challenge at the time. Fast atom bombardment (FAB), thermospray ionization and plasma desorption (PD) were novel techniques that put into practice were not processed well for protein ionization. Introduction of soft ionization techniques pioneered the revolution of biological MS while keeping sample integrity and creating gas phase ions without extensive fragmentation[1-2]. Mass spectrometers have actively participated in a variety of scopes such as forensic, toxicology, metabolomics, proteomics, clinical and, pharmaceutical analyses.

1.1.1. Instrumentation

Essential components of mass spectrometers are illustrated in Figure 1.1. After the sample is introduced into the system using a direct insertion probe or chromatographic methods, the first principle is generating gas phase ions from neutral analyte molecules by means of ionization source (e.g. ESI, MALDI). Ions are accelerated and focused with the electric/magnetic field and separated based on m/z ratio in the mass analyzer (e.g. quadrupole, ion trap, time of flight (TOF)). Detector (photomultiplier, electron multiplier, multichannel plate) records the quantity of ions that hit a surface for each m/z ratio. Lastly, ion currents induced by charged particles were displayed in the form of a mass spectrum (a plot shows intensity against m/z ratio data) by the processing system. Mass spectrometers have to be operated under high vacuum (10^{-5} -10^{-7} Torr) values to maintain low pressure. Otherwise, charged particles of an analyte can react with other gaseous molecules which generate a complex spectrum. Therefore, increased mean free path is essential to avoid neutralization of analyte ions before reach to the detector[3-4].

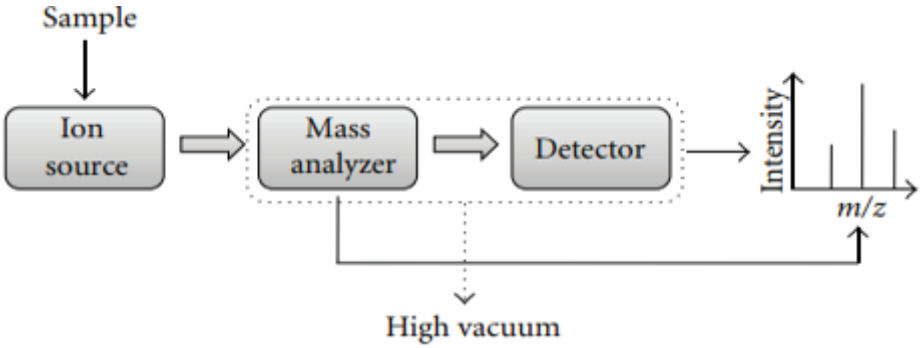


Figure 1.1. The basic components of mass spectrometer
(Source: Banerjee, et al. 2012)

1.2. Sample Ionization Methods

The most important part of mass spectrometry is to convert analyte molecules into ionic species through losing or gaining of hydrogen, electron capture or ejection prior to analysis. Some of the ionization techniques are as follows; electron impact (EI), chemical ionization (CI), field ionization and desorption (FI/FD), fast atom bombardment (FAB), and plasma desorption (PD). The main advance occurred with the introduction of

electrospray ionization (ESI) and matrix-assisted laser desorption/ionization (MALDI) which were started to analyze non-volatile and thermally unstable compounds. These are soft ionization techniques that keep sample integrity with lesser fragmentation[4-5].

1.2.1. Electrospray Ionization (ESI)

ESI is a liquid ionization method under atmospheric pressure to produce gas phase ions. The first development of electrospray ionization was introduced by Malcolm Dole who aimed to find out molecular mass of macromolecules in 1968[6]. John Fenn who was awarded Nobel Prize in 2002 coupled mass spectrometer with electrospray to analyze small molecules and ions (proteins and peptides) by 1980s[7]. Thus, gas phase ions of large macromolecules like proteins and nucleic acids are analyzing using the soft ionization method without degradation of an analyte.

Analyte and organic solvent (methanol, acetonitrile, isopropanol) mixture is injected into the capillary with a continuous flow from syringe pump or chromatographic method by applying an electric field between the counter electrode and the capillary (1-6 kV) to produce charged droplets under atmospheric pressure. The sample must be prepared as clean as possible without contaminants like buffer and salt which might be adsorbed on the capillary tube of ESI, otherwise no signal is acquired. Positive and negative charges start to be separated through the electric field. When the capillary is positively charged (positive ion mode), positive ions start to migrate counter electrode. At the tip, positive ions start to accumulate as droplets which form a Taylor cone. The opposite occurs in the negative ion mode. Droplets are subjected to heated nebulizing nitrogen gas (N_2) which does not form covalent bonds with the analyte ions to evaporate the solvent. Droplets with the same amount of charge start to shrink while solvent is evaporating. When droplet surface reaches to Rayleigh instability limit, coulombic forces become greater than the surface tension of droplets and coulombic fission takes place which produces smaller charged particles. Figure 1.2 shows that generated a series of multiply charged analyte ions in the gas phase are directed with optics into mass analyser[4].

ESI tends to produce multiply charged ions of biomolecules. Especially, tryptic peptides include proton accepting sites at N-terminal amino group and C-terminal site of arginine or lysine residue. Different charge states of peptides or proteins are distributed through the mass spectrum of ESI. Thus, mass range of ESI is unlimited and larger

proteins can be observed at lower m/z ratios. Moreover, the ion source of MS can be coupled with liquid chromatography (LC) due to its liquid state ionization property. ESI tandem (multistage mass analyses of ions) mass spectrometry can be operated with quadrupole, ion trap and time-of-flight mass analyzers and in some cases in combination. This variation provides versatile approaches for proteomic problems[10].

1.2.2. Matrix Assisted Laser Desorption/Ionization (MALDI)

MALDI is a solid-phase ionization technique of large molecular weight, thermally labile, and non-volatile macromolecules such as proteins, oligosaccharides, oligonucleotides or synthetic polymers. In 1988, Michael Karas and Franz Hillenkamp introduced this term which generated doubly charged molecular ions from proteins above 10kDa molecular mass range[11]. The analyte is dissolved in a volatile solvent and mixed with a large excess of UV-absorbing matrix solution as 1:1000 ratio. Matrix is a weak organic acid which has chromophore property and absorbs laser energy, acts as a proton donor or acceptor during desorption. The mixture is placed onto a target plate and analyte molecules are allowed to dry and crystallize uniformly with a matrix solution. The target plate is placed into the mass spectrometer and irradiated with a pulsed UV-laser, generally nitrogen laser at 337 nm or neodymium-doped yttrium aluminum garnet (Nd: YAG) at 355 nm. Laser pulses start to heat the matrix which is ablated and sublimated into gas phase from solid phase. As matrix is transferred into gas phase, internal energy is also transferred into the analyte molecules and provides ionization without fragmentation. Ionized analyte molecules are accelerated and directed towards mass analyzer using electric field (Figure 1.3)[4, 9].

Sample preparation is a critical step to produce a high-quality spectrum for MALDI. Matrix-analyte composition and preparation are crucial obtaining a perfect uniform crystal structure with shot-to-shot reproducibility. A matrix in small quantities causes to start aggregation and oligomerization of analyte. Therefore, a huge amount of analyte should be in the matrix-analyte mixture. The ideal matrix must be miscible and co-crystallized with the analyte at ambient conditions, have vacuum stability, small molecular weight and ability to absorb laser energy which is emitted from a laser at a specific wavelength. Preferred matrices are α-cyano-4-hydroxycinnamic acid (CHCA) for peptides, 3,5-dimethoxy-4-hydroxy-cinnamic acid (sinapinic acid) for proteins and

2,5-dihydroxy benzoic acid (DHB). Positive or negative ions are created according to operated positive or negative ion mode. While MALDI produces mainly singly or doubly charged ions by protonation in positive mode, negative ions are detected in negative mode. In addition, MALDI is more tolerant to contaminants (salts, buffers, surfactants, organic additives) than ESI up to moderate concentrations levels[12].

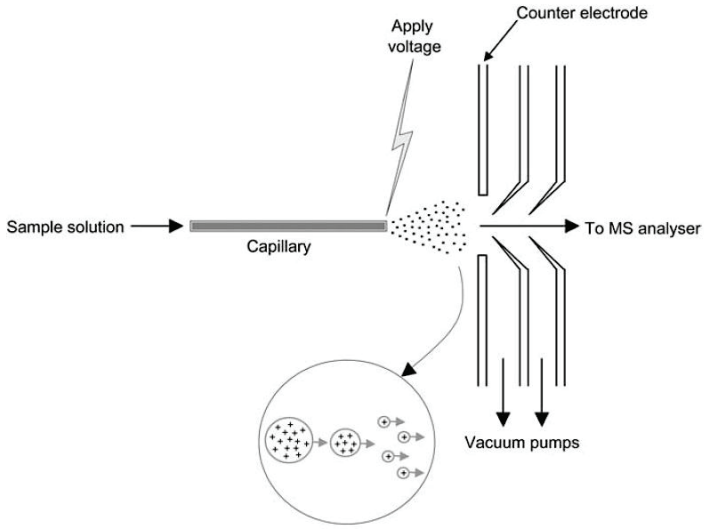


Figure 1.2. Schematic diagram of an ES interface and the ES process
(Source: Lane 2005)

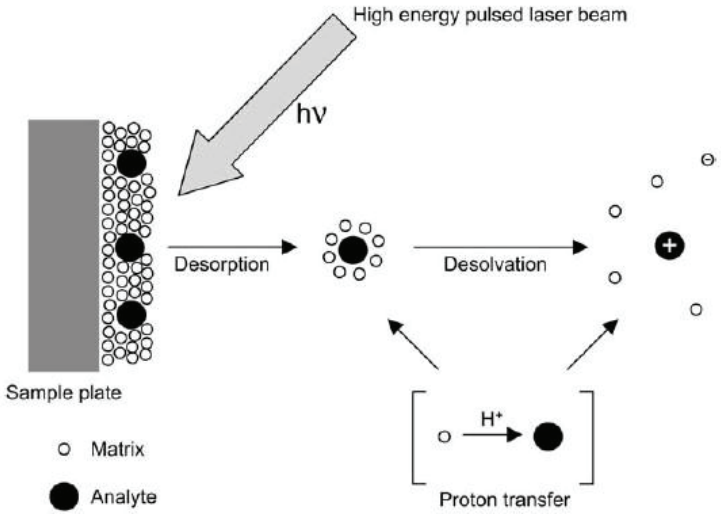


Figure 1.3. The MALDI process
(Source: Lane 2005)

1.3. Mass Analyzers

Mass analyzers which separate ions based on their m/z ratios using a magnetic or electric field, are essential components of mass spectrometers. Ions of different m/z ratios are transmitted and focused on a detector according to their kinetic energy, velocity and momentum to obtain resolved masses of ions. Several types of mass analyzers are utilized for proteomic studies; quadrupole (Q), ion trap (IT), time-of-flight (TOF), Orbitrap, and Fourier transform-ion cyclotron resonance (FT-ICR). Mass range, mass accuracy, resolution (ability to separate two distinct mass signals), transmission (ratio of number of ions reaching the detector and a number of ions produced in the ion source), speed and, ability to perform tandem MS (MS/MS or MS^n) are main characteristics of analyzers. Among them, FT-ICR has an advantage of better resolution and accuracy while the ion trap is capable of multiple mass measurements (MS^n). Ion trap which was performed for the experiments of the thesis will be explained in detail.

1.3.1. Quadrupole (Q)

Physicist Paul Wolfgang described quadrupole mass analyzer by 1950s for the first time. Quadrupole consists of four parallel, electrical metal rods along the central axis (Figure 1.4)[10]. Direct current (DC) and alternating radio frequency (RF) voltages are applied to the opposite rods and one pair is positively charged and the other is negatively charged. Generated ions in the ion source are focused through the central axis between quadrupole rods and the magnetic field is created by applied RF and DC voltages. Each of m/z values has a set amount of DC/RF voltage. Therefore, when specific DC and RF voltages are introduced, only ions with specific m/z values follow a trajectory in oscillating between the rods and reach to the detector. Ions with greater or lesser m/z ratio collide to rods, follow a wrong trajectory and fail to pass through the quadrupole. The voltage applied is proportional to the m/z ratio value of ions that are allowed to pass quadrupole; higher the voltage means ions with higher m/z ratio values will reach to detector[9, 12]. Quadrupoles are low cost, robust, easy to use mass analyzers. However, they have limited mass range, resolving power and capability for MS/MS measurements. The last one can be overcome with the attachment of extra quadrupole (triple quadrupole).

The first and third quadrupoles are responsible for scanning and second quadrupole is a collision cell where daughter ions of selected one are obtained.

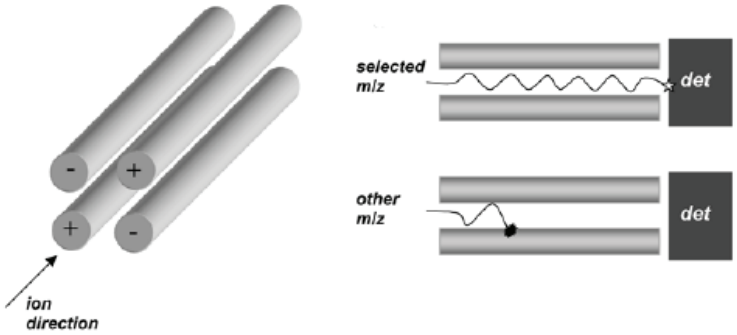


Figure 1.4. A quadrupole mass analyser
(Source: Liebler 2002)

1.3.2. Ion Trap (IT)

Ion traps are mass analyzers with oscillating RF field which is induced to trap ions in two (2D) and three dimensions (3D). Thus, ion traps have two versions as 2D or 3D. The first one is 3D ion trap which is also known as quadrupole ion trap (QIT), was described by Paul and Steinwedel in 1960 and then modified by Stafford and co-workers in 1984[4, 13]. QIT comprises one top, one bottom electrode (end electrodes), and one ring electrode at the center of the trap which forms a closed loop to trap ions (Figure 1.5)[10]. All m/z values get into analyzer at the same time. The ring electrode is subjected to RF voltage to generate an oscillating electric field, unlike quadrupole which also uses DC voltage. Ions start to oscillate in a horizontal eight-shaped trajectory based on their m/z values. Besides, ions proceed to repel each other so their trajectories expand as a function of time. Helium gas, cooling gas, helps to keep ions in the middle of the trap and removes their excess energy to avoid ion losses because of the collision. At the same time, to scan mass range RF voltage and the small amount of RF voltage are applied to the ring electrode and end electrodes, respectively. When the voltage is applied to the ring electrode, energy and oscillation frequency of ions start to increase. Ions with different m/z values have different voltages to leave the trap. Once frequency reaches to the frequency of end electrode, excessive oscillating motion destabilizes ions except for precursor ion and causes ejection outside of the trap. After precursor ion oscillates and

collides by He gas, fragmentation takes place. The most significant property is that the ion trap has the capability for MS^n measurements. RF voltage on the ring electrode is ramped to induce further fragmentation of MS/MS fragments. MS^n measurements provide detailed information about fragmentation mechanism[9-10].

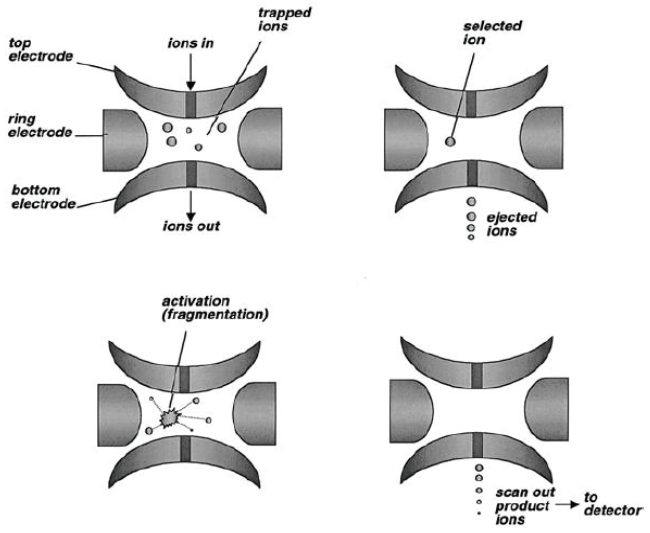


Figure 1.5. Schematic representation of an ion-trap MS instrument
(Source: Liebler 2002)

Besides, linear ion trap (LIT) was developed. This time, ions can be reflected to forward and backward because they were designed as four-rod quadrupole ending in lenses. As the quadrupolar field is induced through radial direction, ends of the trap are subjected to the electric field. LITs have high ion trapping efficiency, sensitivity, and dynamic range.

1.3.3. Time-of-Flight (TOF)

Time-of-light mass analyzers are another kind of analyzers and were described by Stephens in 1946. The first commercial instrument was linear TOF mass spectrometer which was designed by Wiley and McLaren in 1955. The primary step was coupling MALDI to TOF analyzer which improved analysis of synthetic polymers, polymer/biomolecule conjugates as well as biomolecules[4]. The main idea of TOF analyzers is the acceleration of ions under electric field and separation of them according

to their velocities in field-free drift tube which has no magnetic or electrical field inside. As the potential is applied between the electrode and an extraction grid, ions are accelerated with the same kinetic energy forward to the flight tube. Ions start to be dispersed based on their velocities in time within field-free region before reach to the detector, and the heavy ions reach to the detector later because of their lower velocity while arrival time was earlier for lighter ions due to their higher velocity. Mass-to-charge ratios were determined by taking into consideration flight time of ions in the field free region[10].

High resolution and mass accuracy are significant parameters of MALDI-TOF instrument for protein identification studies. However, various velocities of ions with the same m/z values is the reason of resolution problem. There are two solutions: the first one is reflectron and the other is delayed extraction. Reflectron focuses ions of same m/z value which have different velocities (uneven energy distribution); thus, they reach to the detector in an equal time interval. Decelerating voltage provides a larger pathway for faster ion which has higher kinetic energy to catch the slower ion. Therefore, they hit the detector at the same time and individual ions are resolved better Figure 1.6)[9]. Another application is time delayed extraction in which a delay (nanoseconds to microseconds) is created between ionization and ion extraction, and then a voltage pulse is applied to accelerate slow ions and to catch the faster at the detector[9].

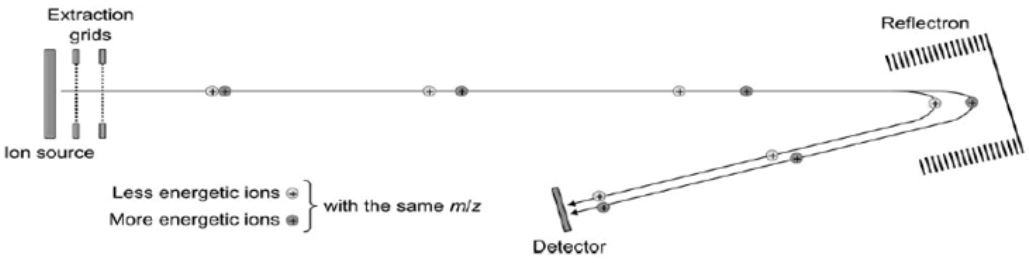


Figure 1.6. Basic components of a reflectron TOF analyzer
(Source: Liebler 2002)

Furthermore, post-source decay (PSD) is one of the techniques to operate with reflectron TOF systems. Herein, voltage is modulated to detect fragment of peptide ions during ionization and acceleration through the tube. PSD spectra of peptide demonstrates peptide immonium ions (H_2N^+ = CHR), which is an indicator of specific amino acid to identify peptide sequences[10].

1.3.4. Fourier Transform Ion Cyclotron Resonance Mass Analyzer (FTICR)

The principle of this analyzer is based on ionization, mass analysis, and detection in an ion trap as a cubic cell which comprises a strong magnetic field inside. The cell consists of three plates: transmitter, receiver, and trapping plates. Figure 1.7 demonstrates the working principle of FT-ICR analyzer. Firstly, ions are injected into the cell along the same plane of the constant magnetic field (3-7 Tesla). When, they are trapped inside the cell they move in cyclotron motion (circular orbit) which is perpendicular to the magnetic field and specific for every m/z value. RF potential is applied to transmitter plate to detect ions, and every ion absorbs energy from the pulse signal. RF signal contains specific frequency for each ion and ions move coherently in larger orbits along receiver plate. Various frequencies are operated for simultaneous detection of ions with different m/z ratio. Mass-to-charge ratio of an ion is determined by measuring and transforming its time domain spectrum into a regular mass spectrum using Fourier transform[9]. Although FT-ICR analyzer is expensive and complicated, it has high mass accuracy, resolution, sensitivity and mass range which makes it a significant tool for analytical proteomics.

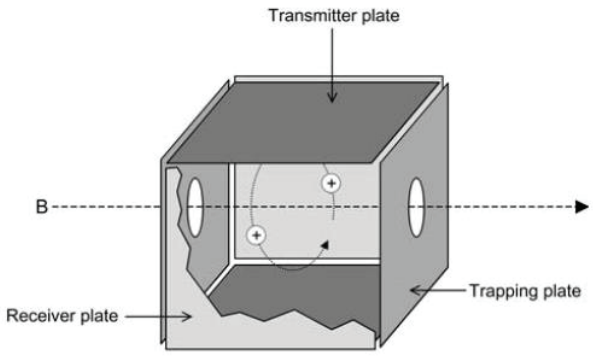


Figure 1.7. Schematic diagram of FTMS analyzer
(Source: Lane 2005)

1.3.5. Orbitrap

Orbitrap, electrostatic trap, was invented by Makarov in 1996 and took place commercially on the market in 2005. There are two electrodes; while the outer electrode has a bent shape and is divided into two with a small gap, the inner electrode has spindle-

shape illustrated in Figure 1.8. Only DC voltage is applied, and ions oscillate around central electrode as complex spirals along the z-axis. The critical property of orbitrap is that frequency is independent of kinetic energy and only dependent on m/z value. Current of ions is measured, then transformed into individual frequencies by FT and mass spectra is obtained[4]. When compared to FT-ICR, Orbitrap has moderate cost, complexity and similar high performance.

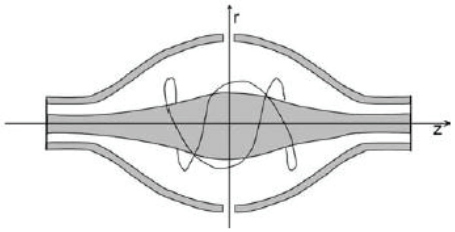


Figure 1.8. The electrostatic trap or orbitrap
(Source: Hoffman, et al. 2007)

1.4. Ion Detectors

The detector is the last part of a mass spectrometer. Ion beam passes through an analyzer which separates ions according to their m/z ratio strike to the detector and electrically detected. There are various types of detectors which are chosen based on required detection sensitivity, speed and chemical stability. Conventional detectors are Faraday cup, electron multiplier (EM), and multichannel plate (MCP). The desired detector must have some properties like low noise, fast response, extensive dynamic range, long term stability, low cost, and high amplitude[14].

The first type is a *Faraday cup* which consists of a metal cup with a conducting electrode. When ions hit to the surface of the metal cup, they induced an ion current due to collisions as they are neutralized. Thus, the discharge current is amplified and detected.

Electron multiplier (EMP) is another detector type which has opposite charge of ions and increases the speed of ions by using a range of dynodes. The first dynode is excited with the emission of secondary electrons and multiplied in a cascade by hitting other dynodes. A metal anode collects and detects all electrons.

Multichannel plate (MCP) is a kind of electron multiplier which was designed with several independent channels. When each ion hit to the surface of the channel,

electrons are produced from the surface. A huge number of electrons are generated with a collision between secondary electrons and surface of the channel in a parabolic trajectory. Every time more and more electrons are emitted and detected.

1.5. Proteomics

Genomics, transcriptomics and proteomics are omic technologies to detect genes, mRNA, and proteins in a biological sample, respectively (Figure 1.9). These technologies have many applications for physiological processes, disease diagnosis, prognosis, and biomarker discovery[15]. The genome is the total set of DNA and genomics is a discipline which deals with sequence information of genomes in organisms. Genome only involves information, in order to function it must be expressed. The transcriptome is a complete set of RNA in a cell or organism. Transcriptional regulation of genes is measured to study of gene expression in transcriptomics. Wilkins has proposed the term proteome which is defined as a full set of proteins encoded by genome in a cell, tissue or organism. The third omic technology is proteomics which is a large-scale study in molecular biology and related about the investigation of composition, structure, function, the interaction of proteins[16]. To determine the sequence of the human genome and to understand the genes, the international Human Genome Project was started. Whereas human genome is assumed to have 35000 genes, there might be millions of proteins if post-genomic processes occur like post-translational modification in protein level. mRNA levels do not directly reflect protein levels in a cell on account of turnover, degradation and protein modifications that is not apparent from DNA sequence. Variety of proteins cannot be explained by genomics, so proteomics which is an advanced step after genomics in biology is necessary to explain cell function from the point of protein level[17]. It is more complicated than genomics because while genome is static and identical for all cells in an organism, the proteome is dynamic that expression level changes with time, cell type, modifications, and environmental conditions. Thus, protein function, localization and protein-protein interaction can be explained by proteomics[18]. The primary objective of proteomics is to detect the progression of a disease, to catch a biomarker and to determine therapeutic targets. Protein composition was monitored to find out cancer tumors[19], Alzheimer's disease[20], and Down syndrome[21] in the last years.

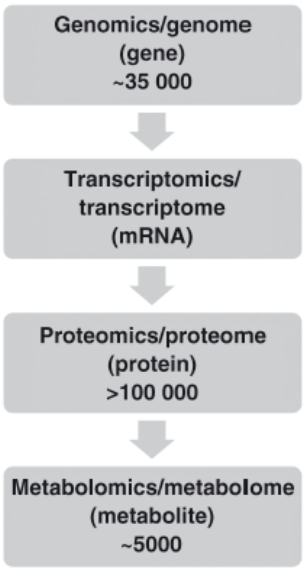


Figure 1.9. Omic sciences and their interaction

(Source: Horgan, et al. 2011)

1.5.1. Types of Proteomics

Proteomics is classified into three groups which are named as expression, structural, and functional proteomics.

Expression proteomics compares qualitative and quantitative protein expression level between different conditions like comparison of protein differences between healthy and diseased cell. Two-dimensional gel electrophoresis (2D-GE), mass spectrometry are the techniques to identify novel proteins for signal transduction and to specify disease-related proteins. Structural proteomics can give information about the three-dimensional shape of proteins, the function of protein complexes and protein-protein interactions. X-ray crystallography and NMR spectroscopy are the technologies to determine the structure of the protein. The functional group is selected and characterized to get information about protein signaling, disease mechanism, and protein-drug interactions in functional proteomics[17-18].

1.6. Mass Spectrometry-Based Proteomics

MS has evolved a preferred tool for identification of proteins, determination of post-translational modifications, researches on protein expression and protein interactions

in the biological field. Workflow of typical MS-based proteomic studies can be categorized into some steps. Firstly, protein mixture is separated and purified into less complex components which are achieved mainly by two-dimensional polyacrylamide gel electrophoresis (2D-PAGE) or chromatographic techniques. After separation, intact proteins or peptides which are obtained from digestion of proteins by an enzyme are directed to a mass spectrometer to gather information about molecular mass and sequence details. Lastly, gathered data were searched against protein database or processed by the algorithm for protein identification or modifications[17].

1.6.1. Separation Techniques Before the Mass Analysis

High concentration and purity for sample preparation which is free of contaminants (at least tolerable) is the key of successful MS-based proteomic analysis. It is critical to resolving complex protein mixture since complex cell lysate includes contaminating molecules such as lipids, carbohydrates, salts, detergents, nucleic acids which restrain proper MS signal and accurate identification of proteins. Two main proteomic approaches for separation are classified as gel-based, 2D-PAGE and gel-free, LC.

1.6.1.1. Electrophoretic Separation

An electrophoresis is an analytical tool which is based on the migration of charged particles in an electric field. 2D-GE is a most preferred powerful technique to separate complex protein mixtures from cells, tissues or other biological samples as individual proteins. In the first dimension, proteins are separated according to their isoelectric point (pI) by isoelectric focusing (IEF). Then, focused proteins are resolved based on their molecular masses by sodium dodecyl sulfate-polyacrylamide gel electrophoresis (SDS-PAGE) in the second dimension (Figure 1.10)[22-23]. Each band of spot pattern on the gel represents one protein and intensity of individual band points out how much the cell produced that protein. Thus, parallel separation and analysis of thousands of proteins are possible. 2D-GE firstly was introduced by O'Farrell[22] in 1975 and this first version includes carrier ampholyte-containing gel in narrow tubes for the

first-dimension analysis. Then, the appearance of immobilized pH gradients (IPG) gels instead of carrier ampholytes facilitated resolution and reproducibility of IEF[22].

Protein solubilization is a necessary step for gel-electrophoresis. Native structure of proteins resists to solubilization, so denaturation is crucial to support solubility of proteins. Denaturation means breaking of all bonds such as disulfide bonds, hydrogen bonds, ionic interactions, hydrophobic interactions, and van der Walls forces. Otherwise, proteins start to aggregate and precipitate. Rehydration buffer is required for solubilization, denaturation of proteins and hydration of IPG strips. Major components are urea or urea/thiourea as a denaturating agent, zwitterionic detergent CHAPS or ampholytes as a solubilizing agent, and dithiothreitol (DTT), dithioerythreitol (DTE) as reducing agent. However, the isoelectric point of protein and ionic strength of solution should not be affected by rehydration buffer. When the strip is hydrated with buffer, protein solution is loaded into the strip under voltage. Then, voltage is increased and maintained for several hours to focus proteins. Depend on the surrounding, proteins carry either positive, negative or zero net charge which indicates they are amphoteric molecules. The total charge of the protein is a summation of all charges on not only amino acid residues but also amino and carboxyl terminal. The isoelectric point (pI) is the point when net charge of the protein is zero. Proteins are positively charged if pH values are below their pI and negatively charged if pH values are under their pI[22]. Under the influence of electric field, protein moves along pH gradient to the point where its net charge is zero and then stops migration. Therefore, focusing effect of IEF derives separation of proteins at their pIs despite small charge differences. After IEF process, strips were applied onto SDS-polyacrylamide gel in which separation is based on molecular weights, independent of charges. Polyacrylamide gels are produced regarding free radical polymerization reaction between acrylamide and cross-linking agent N, N'-methylene-bis-acrylamide with control of initiator- catalyst ammonium persulphate-N, N, N', N'-tetramethylethylenediamine (TEMED)[23]. Pore size can be modified by changing the concentration of acrylamide and a cross-linking agent. The aim is to remove the effect of the charge in the second dimension. Protein molecules are treated and covered with sodium dodecyl sulfate to disrupt the secondary, tertiary, quaternary structure of proteins and negatively charged linear polypeptide chains are generated. Mobility starts to depend on just size and the uniform charge is formed. Thus, proteins are distinguished according to their size while they are dragged along gel. Separated proteins are visualized with different staining procedures; silver staining, Coomassie

staining and, Sypro Ruby staining. Differential proteins are compared between two different conditions and excised from the gel to digest enzymatically (mainly with trypsin enzyme). Then, they are analyzed with proteomic approaches to reveal disease related protein, stress impacts, and toxic influences. Cell differentiation, drug discovery, cancer research and, protein purification are other principle applications of 2D-GE[22, 24].

Despite advantages, 2D-SDS-PAGE is time consuming due to excess number of protein spots and has poor reproducibility when visualization of stained gels is compared. Also, the small dynamic range of staining allows the only visualization of abundant proteins. Also, low soluble proteins, membrane proteins, are not apparent on the gel. Moreover, co-migration of proteins and biases in pI and molecular weight are problematic issues to obtain information.

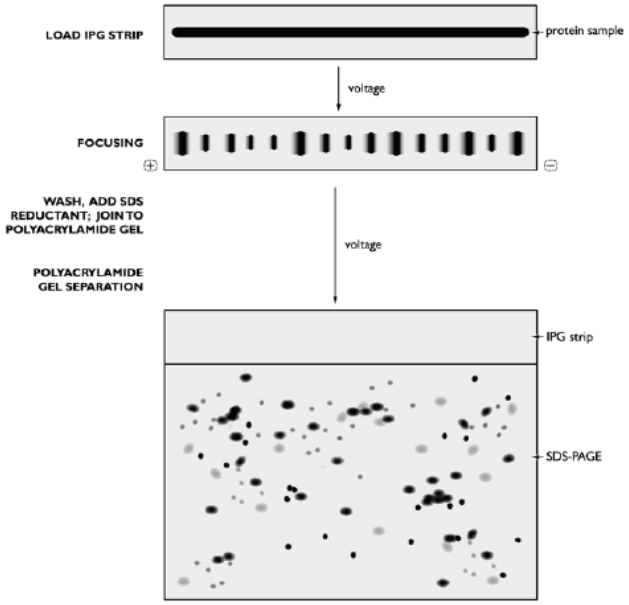


Figure 1.10. Schematic representation of 2D-SDS-PAGE
(Source: Berkelman 2002)

1.6.1.2. Chromatographic Techniques

The idea of all chromatographic techniques is associated with the interaction of sample and stationary phase which is followed by elution with the mobile phase to disrupt the interaction between sample and stationary phase. Separation in LC rests upon retarding time of compounds in a column. LC step eliminates ion suppression in MS,

allow enrichment of low abundance peptides and proteins and removes salt from the sample[23]. Properties like high recovery, reproducibility, resolving power, speed and compatibility with ESI make high performance liquid chromatography (HPLC) attractive alternative to SDS-PAGE in proteomics. Also, minimal sample preparation is another advantage when compared to SDS-PAGE. Reversed-phase (RP), ion-exchange, affinity and size exclusion chromatography are the kinds of chromatographic separations and can be used alone or in combination. RP separation is based on hydrophobic interaction between sample and stationary phase which is composed of covalently bonded nonpolar hydrophobic groups to stationary silica particles. Elution is accomplished by a gradual increase of hydrophobicity of mobile phase through increasing concentration. Not only compatible with MS, but also the resolution for separation and analysis of proteins and peptides highlights that reversed-phase high-performance liquid chromatography (RP-HPLC) as a well-established tool in proteomic analyses[25].

Besides, RP-HPLC can separate similar polypeptides which only have one amino acid residue difference. The chromatographic approach produces a large amount of peptide fragments from protease digests of cell lysate. Digested proteins are separated with RP-HPLC and directly introduced into ESI-MS for identification. RP-HPLC is widely preferred for analyses of protein therapeutic products and natural protein and peptide studies[25]. Although one-dimensional LC has been applied as an economic technique, its application is restricted for proteomic studies which deal with thousands of proteins, so the analytical range of 1D-LC is exceeded because of insufficient peak capacity. Consequently, separation techniques with higher resolution are required for complex mixture analysis which contains thousands of peptides and proteins. Therefore, single LC-MS/MS is limited to identify a high number of proteins. Higher resolution, increased dynamic range and proteome coverage are developed and achieved with multidimensional chromatography or orthogonal chromatographic steps (2D) in which two different chromatographic techniques are coupled together[25-26].

In multidimensional techniques, ion-exchange chromatography (IEX) has been used for separation of peptides and proteins as the first dimension in 2D-LC. Then, fractions are separated with RP-HPLC and directed to ESI-MS analysis[27]. Besides, affinity chromatography as the first dimension has been used for phosphorylated and glycosylated proteins, while size-exclusion resin was integrated to RP-HPLC as a first dimension[28]. Generally, gel-free methods mean shotgun approach in which after protein

mixture is digested in solution, thousands of peptides are fractioned by LC, and identified by LC-MS/MS based on data dependent acquisition.

1.6.2. Protein Identification by Mass Spectrometry

There are two approaches for mass spectrometry analysis of proteins which is extracted from biological samples: bottom-up and top-down (Figure 1.11). Top-down approach involves the introduction of intact protein into the gas phase by ESI and followed by subsequent fragmentation in high-resolution mass spectrometer (generally FT-MS) for mass measurement of both protein and fragment ions. In this approach, intact protein is directly fragmented without prior enzymatic digestion. Protein identification relies on the comparison of masses of both protein and its fragment ions against sequence database search. Since it deals with whole mass, a top-down approach is capable of characterization of post translational modifications (PTMs) and protein isoforms[2, 29].

The bottom-up approach is based on analysis of chemically or enzymatically produced peptides rather than protein analysis. Peptide analysis has several benefits over protein analysis such as lower molecular mass, effective separation by LC, higher sensitivity and lesser charge states. The classical bottom-up approach, proteins are subjected to separation using gel electrophoresis or chromatographic techniques followed with digestion procedure. Trypsin is mostly applied proteolytic enzyme in bottom-up proteomics and it cleaves peptide bonds on C terminal of arginine or lysine residues which are well-distributed in protein and most abundant amino acids in human proteome[30]. Peptides are converted to the gas phase and analyzed with a mass spectrometer to determine peptide mass fingerprint (PMF) of each peptide. Then, collected data is transferred into protein database and peptide fingerprints are compared with each entry in the database. Alternatively, digested peptides can be further separated by RP-LC and introduced into a tandem mass spectrometer in which individual peptide ions are selected, fragmented and fragmentation pattern shows the partial amino acid sequence of peptides. Sequence database search assists to identify amino acid sequence from the pattern. On the other hand, a new method in which separation is not involved before digestion has revealed as a shotgun approach due to the emergence of tandem mass spectrometer[31].

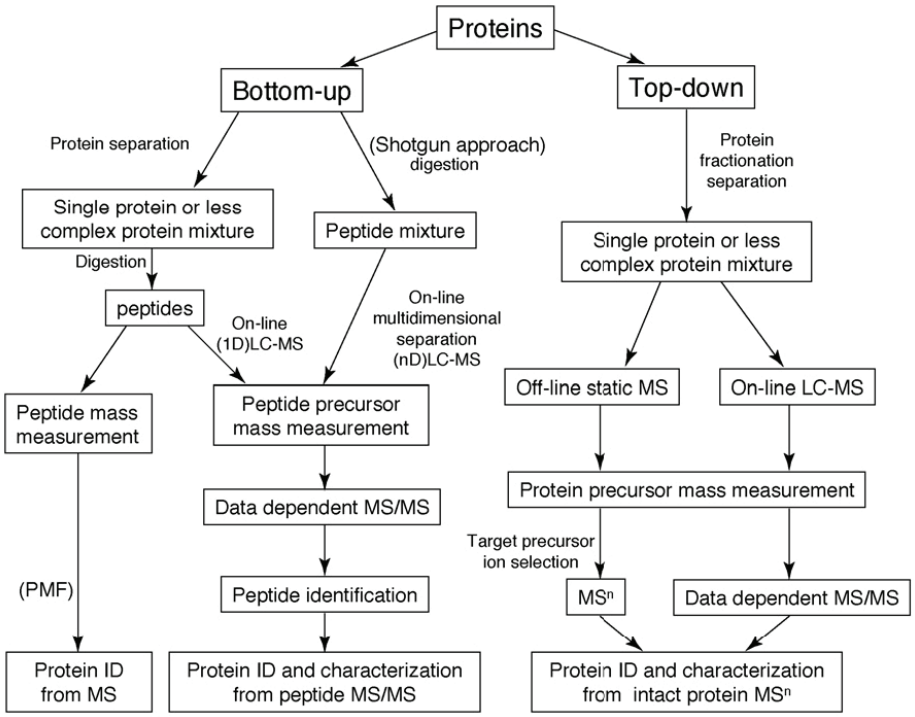


Figure 1.11. Strategies for MS-based protein identification and characterization (Source: Han, et al. 2008)

1.6.2.1. Shotgun Proteomics

Proteome analysis by *shotgun approach* relies on fractionation of sample before its analysis on the mass spectrometer. There is an analogy between shotgun proteomics and DNA sequencing, so shotgun method allows identification of proteins from a mixture using a tandem mass spectrometer. Proteins are digested into peptides by an enzyme like trypsin. Then, peptide mixture is separated by one dimensional or multidimensional HPLC and directly introduced into the mass spectrometer to measure m/z ratio of peptides. Interested peptides are selected and subjected to further fragmentation to produce MS/MS data. Resulted data were searched against a protein database to acquire peptide sequence information[32]. MS analysis and MS/MS fragmentation are performed consequently in shotgun proteomics because this approach is based on data dependent acquisition mode of the mass spectrometer. In this mode, the first MS survey scan of eluting parent ion serves to monitor peptide intensity and to identify potential targets for fragmentation. After a list is generated, parent ion from the list is automatically selected

to fragment at the second MS/MS scan. Therefore, the second MS/MS scan depends on the first MS survey.

Compared to other technologies (2D-SDS-PAGE and LC fractionation), shotgun approach simplifies sample handling, enables entire proteome analysis, increases data throughput and certainty of sequence information so adopted for protein analysis. However, there are some drawbacks. Before shotgun proteomic studies, the dynamic range of protein concentrations and the presence of a few highly abundant proteins should be considered. Another problem is the assignment of incorrect peptide sequences to MS/MS spectra. Thus, validation of an identified overall set of proteins by manual inspection of the MS/MS ion spectra is necessary. Another major challenge often encountered in shotgun proteomics is the protein inference problem. Due to the complexity of the peptide mixture, some sequences of identified peptides can be shared by several proteins in the database. This situation causes ambiguity for the identification of proteins. Even if most of the peptides are shared between two or more proteins, at least one peptide should be unique to that protein to differentiate it from other proteins. Increasing the number of unique peptide match, enhances the reliability of protein identification. However, a single peptide match is taken into consideration as a doubtful protein, so it is ignored[33-34].

The great number of peptides after protein digestion without any pre-fractionation step should be reduced to a measurable degree before LC-MS/MS analysis. Through this way, dynamic range and proteome coverage are increased compared to single LC-MS/MS. Multidimensional chromatography combined with MS has started to be used routinely for gel-free proteomic analysis. The basis of separation for multi-dimensional chromatography is orthogonality so each dimension applies different properties. Also, the effectiveness of 2D-LC depends on separation efficiency and compatibility of two separation dimensions[35]. Offline and online techniques are the categories of multidimensional HPLC[26]. The offline technique involves a fraction collection of samples at the end of the first dimension and re-injection onto second column. However, in online technique, two chromatographic methods are coupled to each other; eluent from first column is directly eluted onto the second column without fraction collection. Two columns are coupled together by column-switching system or multidimensional protein identification technology (MUDPIT). John Yate and colleges[28] coined the term MUDPIT. Here, an automated multidimensional LC which integrates strong cation exchange (SCX) resin and RP resin in the same capillary column were

described and combined with ESI MS online and database search for protein identification. However, SCX has some limitations like reduced sample recovery, low capability for resolving peptides and sample losses due to sample desalting which can have a negative impact on analytical sensitivity. Also, peptides tend to group during fractionation. By contrast with SCX, it is known that RP resolves peptides better and salt free buffers generates cleaner samples for LC-MS/MS analyses. However, RP for both dimensions of 2D-LC is not widely preferred due to orthogonality. Instead, it was showed that high-pH RPLC is semi-orthogonal compared to the low-pH option because of dramatic change of charge distribution. Peptides are uniformly covered in high-pH RPLC. Thus, RPLC-RPLC at different pH values represents orthogonality in proteomic approaches. Concatenation after the first dimension which is a combination of early, middle, late fractions eluted over equal time interval offers advantages regarding orthogonality between two separation and proteome coverage. It was pointed out that concatenated high-pH RPLC–low pH RPLC is an alternative to increase proteome coverage for gel free shotgun analyses[36].

1.6.3. Peptide Sequencing and Database Search

The traditional technique for identifying and sequencing of proteins is named as Edman degradation in which phenyl isocyanate is reacted with N-terminal amino group of polypeptide chain thus phenylthiocarbamyl derivatized residue is formed to perform hydrolysis of nearest amide bond. HPLC determines removed N-terminal amino acid. At the same time, the following residue is ready for the reaction, so cycle continues for adjacent amino acids. Even though Edman degradation provides accurate results, it can analyze pure protein or peptide samples, so its utilization is not convenient for complex protein mixture. But, it is unable to detect modified, or N-terminal blocked amino acids[37]. Western blotting which was introduced by Towbin in 1976 is another protein identification technique and enables antigen-antibody interaction to identify target protein in a complex protein mixture[38]. However, it requires a presumption of antibody and suffers from non-specific binding of the antibody. The technique gives qualitative or semi-quantitative information about the protein.

On the other hand, MS-based proteomics is capable of rapid sequencing and identification of post translational modifications. Tandem MS (MS/MS) has become a

significant key for proteomic studies. Multiple mass analyzers can be combined as the same kind of analyzers such as TOF/TOF, triple quadrupole (QqQ), or combined with different analyzers like Q/TOF in order to perform MS/MS analyses. After an ion with certain m/z (parent ion or precursor ion) is selected in the first analyzer, the ion is focused into collision cell and collide with an inert gas. During the collision, fragmentation of ion occurs as kinetic energy of precursor ion turns into vibrational energy. Lastly, the second mass analyzer measures m/z value of fragment ion. This process is frequently applied MS/MS technique for proteomics and known as collision induced dissociation (CID). Apart from that, fragmentation can repeat several times and ion trap and FT-ICR possess the advantage of MS^n measurements where n is the number of repeating MS experiments to obtain detailed information about the structure.

Acquired spectral data from tandem MS measurement is searched against theoretical data generated from a sequence database using search engines like Sequest, Mascot and, X! Tandem which is preferred to identify best matches of peptide sequence in protein database[39]. Thus, thousands of proteins can be determined through a single measurement. Among them, Mascot uses probability-based matching, and some parameters are required before execution of database search.

Firstly, a suitable database to be searched is selected such as the National Center for Biotechnology Information (NCBI) and Universal Protein Resource (UniProt). An enzyme that performs proteolytic digestion is also selected for the assumption of theoretical peptides in the database because the digestion procedure is also carried out for proteins in the database. Missed cleavage sites for the enzyme are defined since it is widely accepted that there are not only peptide products of ideal digestion. Fixed modifications in which all amino acids are considered modified and variable modifications in which all amino acids are considered with or without modification are indicated as other important parameters. Moreover, precursor (MS) tolerance, fragment (MS/MS) tolerance and, peptide charge must be assigned. When the search is completed, a result of the protein list with score numbers is obtained. Higher score number means confident and reliable matches for identification[40]. Another critical step is to search the data not only in standard target database but also in decoy database which resembles target database in length and amino acid distribution as reversed or shuffled form. False discovery rate (FDR) is an assessment of data quality and described as the proportion of wrong peptide spectrum matches (PSM) among the identifications. Therefore, lower values of FDR (e.g., %5) are preferable for proteomic studies. Besides, the percolator is

an optional filtration to differentiate positive and negative peptide matches to decrease false discovery rate[41].

CHAPTER 2

Trichomonas vaginalis AND SEXUALLY TRANSMITTED DISEASE TRICHOMONIASIS

2.1. Introduction to *Trichomonas vaginalis* and Trichomoniasis

The genus Trichomonas is a parasitic flagellated anaerobic protozoan[42]. Three species of Trichomonas are observable in human and they inhabit the large intestine, the mouth and the uregenital tract by *T. hominis, T. tenax* and *T. vaginalis*, respectively. *Trichomonas vaginalis* is a single-celled flagellated protozoan which was first discovered by Alfred Donné in 1836[43]. This organism possesses four flagella for the motility and one of the flagella attached to the part of the undulating membrane. A nucleus is located at the anterior end, and an axostyle passes through the nucleus by extending from the end of the body. The parasite does not have any known cyst form, it has only the trophozoite stage and divides by binary fission[43].

T. vaginalis does not possess respiratory metabolism due to the absence of mitochondria[43-44]. Therefore, energy metabolism of the parasite comprises glycolysis in cytosol which is followed by oxidative decarboxylation of pyruvate in double membrane organelle hydrogenosome for ATP and molecular hydrogen production (Figure 2.1)[42]. Glucose or glycogen is converted to phosphoenolpyruvate in the cytosol. Then, either cytosolic pyruvate kinase or phosphoenolpyruvate carboxykinase, malate dehydrogenase, hydrogenosomal or cytosolic malic enzyme are assigned to catalyze the formation of pyruvate from phosphoenolpyruvate[42]. Resulted pyruvate passes through the hydrogenosome and it is directly decarboxylated to Acetyl-CoA by hydogenosomal enzyme pyruvate:ferredoxin oxidoreductase (PFOR)[42, 45] . As a result of CoA transfer into succinate via acetate:succinate CoA transferase, acetate is produced as an end product. ATP is released due to catalysis of succinate thiokinase and CoA get involved in PFOR reaction cycle again[42, 46-47]. Also, electron carrier iron-sulfur protein, ferredoxin (Fd), captures electrons generated by PFOR. Because of coupling the electrons to hydrogen ions by hydrogenase which is responsible for catalyzing the oxidation and reduction of hydrogen, another product molecular hydrogen is generated[42, 48-49].

Besides, NAD(P) malic enzyme catalyzes decarboxylation of malate for pyruvate metabolize. NADH is then reoxidized by NAD:ferredoxin oxidoreductase to NAD^+ [42, 45].

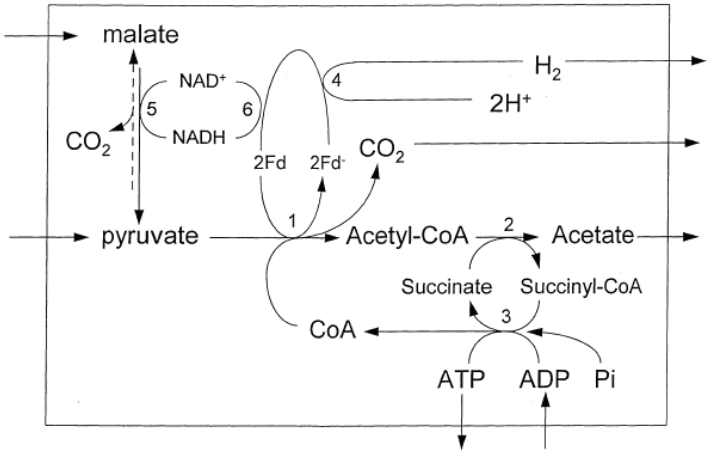


Figure 2.1. Schematic map of hydrogenosomal metabolism. 1, Pyruvate: ferredoxinoxidoreductase; 2, acetate:succinate CoA transferase; 3, succinate thiokinase (succinyl CoA synthetase); 4, hydrogenase; 5, NAD(P) malic enzyme (L-malate: NAD(P)-oxidoreductase [decarboxylating]; 6, NAD:ferredoxin oxidoreductase; Fd, ferredoxin.

Trichomonas vaginalis causes trichomoniasis, non-viral sexually transmitted infection (STI) in both women and men, by affecting human uregenital tract[50]. Trichomoniasis has an essential place among sexually transmitted diseases. According to the report of World Health Organization (WHO), the rate of encountered cases of the disease is increased by 11.2 % from 2005 (248.5 million) to 2008 (276.4 million)[51]. Although, trichomoniasis is endemic or potentially endemic to all countries, the rate of incidence is higher in sub-Saharan Africa, South and South East Asia[52]. In Turkey, the prevalence of the disease is not fully understood; in different groups positivity was observed at different rates because of regional epidemiological studies. Trichomoniasis transmission is possible with trophozoite stage. The organism turns its typical globular, pear-shaped form into flat, thin, ameboid shape after attaching to epithelial cells of genital tract[53] (Figure 2.2). Ameboid shape increases with the interaction of the parasite and the host cell. Adherence to this interaction causes inflammation associated with trichomoniasis[54]. Infection complaints in women differ from asymptomatic to severe acute inflammatory disease[55].

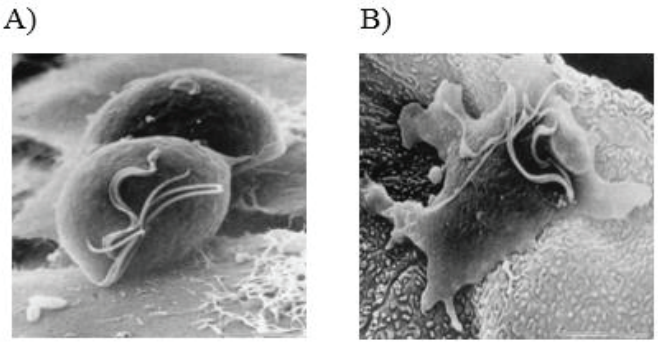


Figure 2.2. Scanning electron microscopic images of T. vaginalis trophozoite form (A) and ameboid form (B) (Source: Arroyo, et al. 1993)

Most encountered symptoms are abdominal pain, foul-smelling frothy discharge with green or yellow colour, itching, and burning. Despite the asymptomatic infection in men, sometimes urethritis, prostatitis, epididymitis is observable[43, 51, 55]. Acute infections increase the risk of pelvic inflammatory disease and cervical cancer. Furthermore, chronic infection of trichomoniasis predisposes to disease carrier of human immunodeficiency virus 1 (HIV-1) infection and exhibit the probability of adverse pregnancy outcomes like preterm birth and low birth weight[42, 50, 56].

One of the diagnostic methods is wet mount preparation which is based on microscopic visualization on a vaginal wet mount slide. Despite low cost and immediate diagnosis, the technique is insensitive for the diagnosis. The complexity of the direct immunofluorescent antibody staining brings about the problematic detection. Culture (modified Diamond's medium) is more sensitive than wet mount, however it is not widely used diagnostic method. *T. vaginalis* antigen detection via immunochromatographic assays and PCR assays are also prevalent methods for the detection[51, 57].

For the first time reference proteome map of *T. vaginalis* was studied using fresh clinical isolates to provide information about overall proteome profile rather than focusing on specific proteins by a combination of 2D-SDS-PAGE and MALDI-TOF/TOF MS[50]. Besides, a recent study points out changes in the MALDI-TOF MS parameter (modification of range setting 6-10kDa) allows it as a potential tool for *T. vaginalis* identification based upon proteome profile[51].

Metronidazole and tinidazole are 5-nitroimidazole drugs and introduced as antitrichomonal agent for treatment of trichomoniasis. Metronidazole is preferred for its ease of use, inexpensiveness and less side effect. Metronidazole enters cell via simple

diffusion. Usually the drug is inactive, and it should be in active form to be a toxic compound. Activation of metronidazole in the cell was given in Figure 2.3. It acts as an electron acceptor when enters hydrogenosome organelle and competes with hydrogenase for electrons. Consequently, the electrons generated by PFOR is directed to the drug by ferredoxin not to hydrogenase. By this way, hydrogen production is inhibited and reduction of the drug results with production of toxic nitro radical anion which binds and breaks the strands to DNA and induces cell death. Nitro radical is converted back to prodrug form in the presence of oxygen[42].

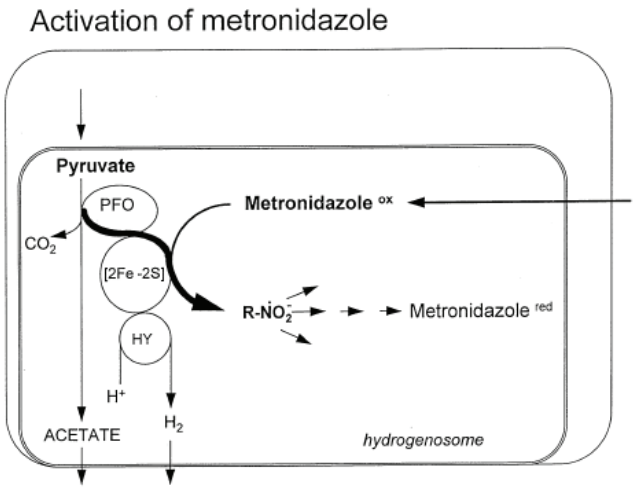


Figure 2.3. Scheme of metronidazole activation in the hydrogenosome
(Source: Kulda 1999)

2.2. Resistance to Treatment

However, sometimes treatment by metronidazole is ruled out and at least 5% of clinical cases stay resistant to the cure. Resistance rates change around the world, while this value is 2-5% in the United States, it exceeds %15 in Papua New Guine[58-59]. Metronidazole resistance has been studied to unravel the resistance mechanism of the parasite because it is a main problem for the treatment.

When trichomonads are exposed to sublethal concentrations of metronidazole both in vivo or in vitro, two types of resistance are appeared as either aerobic or anaerobic. Generally, metronidazole resistant clinical isolates show aerobic resistance in vivo when treated with metronidazole. Elevated levels of oxygen in hydrogenosome compete with

metronidazole because of its electron receptor property to catch ferredoxin-bound electrons which results with impaired activation of the drug and oxygen scavenging pathways because of decreased oxidase activity. That is why, this mechanism evolves increased futile cycling in which nitro radical oxidized back to its prodrug form and then reduced to superoxide anion. Instead of cell death, the process is followed with cell damage by superoxide anion. Also, it was pointed out that aerobic resistance mechanism is the result of defective ferredoxin levels[60]. Then, defective ferredoxin levels for clinical isolates which renders the aerobic resistance was showed as a reason for ferredoxin gene transcription regulation[61]. Even if laboratory induced resistance was associated with decreased ferredoxin levels with shrinking of hydrogenosomes, the situation was different for resistant clinical isolates which present not reduced transcription of ferredoxin with normal-sized hydrogenosomes[62]. Besides, disruption of a single ferredoxin gene did not lead to metronidazole resistance because seven ferredoxin gene was identified for hyrogenosomal signals[63].

On the other hand, anaerobic metronidazole-resistant strains can tolerate increased levels of the drug in vitro anaerobic conditions. The main reason for anaerobic resistance caused from eliminated drug activating pathways which results with low activation of metronidazole. Laboratory-induced resistant *T. vaginalis* strains showed down-regulation of PFOR and ferredoxin activity[64-65]. Also, hydrogenase activity was impaired, and hydrogen production was ceased. However, previously it has been reported that inhibiting hydrogenase activity induces cell death; therefore, multiple factors should be considered in drug resistance[58].

An alternative pathway for metronidazole reduction was determined as flavin enzyme thioredoxin reductase which is both target of nitroimidazoles and an enzyme for nitroimidazole activation. Nitroimidazole treatment disrupts thioredoxin activity and diminishes non-protein thiol levels[66]. Laboratory-induced high-level metronidazole resistant strains lack thioredoxin reducing activity due to lack of FAD cofactor; however, its activity is re-established again with the addition of FAD cofactor. Furthermore, it was pointed out that thioredoxin reductase activity was similar in both clinically metronidazole resistant and metronidazole susceptible *T. vaginalis* isolates[59]. Also, it has been demonstrated that free flavins can take place for the reduction of nitroimidazoles. Flavin reductase activity (described as NADPH oxidase) is capable of reducing oxygen to hydrogen peroxide, and its diminished activity might lead to impaired oxygen scavenging[59]. That is why, in contrast to metronidazole-susceptible cells, flavin reductase

activity was down-regulated or absent (free flavins were not reduced) in metronidazole resistant lines[59, 66].

Genome sequence of *T. vaginalis* was studied and 11 nitro-reductase genes were identified. It was hypothesized that single nucleotide polymorphism (SNP) which were ntr4$_{TV}$ and ntr6$_{TV}$ could cause loss of protein function and could be associated with metronidazole resistance. Thus, SNP might have a clinical role in identifying metronidazole resistance in refractory cases. However, further studies are necessary not only to investigate the role of ntr4$_{TV}$ and ntr6$_{TV}$ but also to determine a biomarker for clinical studies of metronidazole resistance independently of culture-dependent detections[65]. Another study was provided clues for clinical resistance such as reduced nicotinamide adenine dinucleotide phosphate (NADPH)-nitro-reductase, and NADH flavin oxidoreductase genes[63].

2.3. The Aim of The Thesis

The thesis aims to perform comparative proteome analysis of clinically isolated metronidazole resistant and sensitive strains of *T. vaginalis* by using shotgun proteomic approach. The results of this study can make a significant contribution to understand the mechanism of resistance. Moreover, results can be used to help the development of more effective therapeutic targets and improvement of current treatment methods.

CHAPTER 3

MATERIALS AND METHODS

3.1. The Growth of *Trichomonas vaginalis* Strains

In this study, protein samples were categorized into two and labelled as TV1 (sensitive-1), TV2 (sensitive-2) and ATCC 50143 (resistant), respectively. These protein samples were clinically isolated from *T. vaginalis* infected patients in Turkey.
Sensitive-1 and sensitive-2 are the clinical isolates from 45- and 42-years old women, respectively and the clinical isolates were sensitive to metronidazole treatment. Resistant one refers to metronidazole resistant clinical isolate of *T. vaginalis*. Parasite growth, isolation and protein extraction steps were performed by Prof. Dr. Ahmet Özbilgin and his team in Celal Bayar University, Faculty of Medicine. Cryotubes containing *T. vaginalis* strains in a liquid nitrogen tank were kept in the water bath at 37 ° C until they were melted. Parasite strains were first cultivated in Trypticase Yeast Extract Maltose (TYM) medium and incubated at 37 ° C. TYM medium was controlled for the growth.

3.2. Total Protein Extraction

When population reached to $2x10^6$ *T. vaginalis*/ml, the medium was centrifugated at 1200 rpm for 5 minutes and the supernatant was discarded. The remaining part was centrifuged again after 5 ml addition of phosphate-buffered saline (PBS). This washing process was repeated three times to remove the media. The pellet was transferred into TYM medium without the supplement of horse serum and incubated for 8 hours at 37 ° C. After three times washing process of cells with PBS, the supernatant was removed, and 1 ml of Mammalian Cell Lysis Reagent was added. Pellet with reagent was shaken for 10 minutes at room temperature. Then, cell lysate was centrifugated at 14.000 rpm for 15 minutes. The supernatant was received to another tube and stored at -20 ° C for the next analyses. Cold acetone precipitation protocol was applied for protein preconcentration and purification. In the protocol, 1 volume of protein sample was mixed with 4 volume of cold acetone and incubated at -20 ° C for 4 hours. Then, the mixture

was centrifugated at 14.000 rpm at 4° C for 15 minutes. The supernatant was discarded without disturbing the pellet at the bottom and resuspension buffer (7 M Urea, 2 M Thiourea, 0.1 M Tris-HCl / pH 7-8) was added to dissolve the pellet. Next, Bradford protein assay was performed in 96-well microplate (Greiner Bio-One) with Coomassie Plus Assay Kit (Thermo) to determine the protein concentration. 5 µl of different concentrations of bovine serum albumin (BSA) standard was mixed with 240 µl of Bradford reagent and 5 µl of resuspension buffer to obtain a calibration curve. 5 µl of the sample which has unknown concentration was added into 240 µl of Bradford reagent and 5 µl of water to find out concentration value from the calibration curve. Protein samples were diluted with a sufficient volume of resuspension buffer to arrange concentration values inside the calibration curve.

3.3. In-Solution Digestion

Proteins were converted to peptides with the help of in solution digestion method. Molecular weight cut-off (MWCO) filter was used to remove excess urea from the protein solution because urea might result in post translational modification during trypsin digestion at elevated temperature. Enough volume of resuspension buffer was added by pipette into 10 kDa MWCO filter (Merck Millipore) to contain 400 µg of total protein for each sample. 5 µl of 0.2 M Dithiothreitol solution (DTT) as a reducing agent, prepared in 50 mM Tris-HCl with pH 8, was added and incubated for 1 hour. Then, 20 µl of 0.2 M 2-Iodoacetamide (IAA) as an alkylating agent, prepared in 50 mM Tris-HCl with pH 8, was added and incubated for 1 hour. Lastly, 20 µl of 0.2 M DTT solution was added and incubated for 1 hour again for neutralization of IAA. Samples were centrifugated to remove urea and thiourea at 14.000 rpm for 20 minutes. In order to wash the samples, 200 µl of 50 mM Tris-HCl (pH 8) was added and centrifugated again. Trypsin solution was prepared in 50 mM Tris-HCl solution which contains % 10 acetonitrile (ACN) for each sample. The final concentration of trypsin was 20 µg/ml and 200 µl of the trypsin solution was added to each sample by pipetting gently. Prepared sample and trypsin mixture were transferred into low binding tubes and incubated overnight at 37° C. Injection volume of HPLC sample loop was 100 µl so the total volume of the sample was decreased about to 120 µl in vacuum centrifuge in case of sample loss during the injection.

3.4. High pH Reversed-Phase Chromatography with Fraction Concatenation and LC-MS/MS Analysis

Chromatographic separation was necessary to decrease the sample complexity due to protein digestion which was created a complex mixture of peptides. Therefore, HPLC separation was carried out before LC-MS/MS analysis to increase proteome coverage and dynamic range. Separation was performed using a SHIMADZU Prominence UFLC HPLC system and LC Solution Version 1.25 SP3 software.

Reverse phase C18 column with dimensions of 25 cm × 0.46 cm; 5 µm (Teknokroma Mediterraneasea 18) was preferred, and condition of the column temperature was set at 40 ° C for the analysis. Mobile phase A and B were prepared as 10 mM ammonium formate in water and 10 mM ammonium formate in 90% acetonitrile by adjusting pH to 10 with ammonium hydroxide, respectively. The injection volume of HPLC loop was 100 µl and the flow rate was operated as 0.5 ml/minute.

Samples were fractionated into 96 well plate (each well is 250 µl) by means of fraction collector (Sunchrom Micro Fraction Collector) during about 35 seconds/well. Binary gradient LC time programme for fractionation was given in Table 3.1. Following fractionation, eight wells combined by passing twelve fractions after each of them into 2 ml tubes. Thus, 96 fractions were reduced to 12 fractions for each kind of *T. vaginalis* protein digests in total, and eight well was pooled for each fraction. All fractions were completely evaporated in rotational vacuum concentrator for further analyses. For desalting and concentration of each fraction, ZipTip pipette tips (Merck Millipore) containing 0.6 µl C_{18} reversed-phase resin was used before mass analysis. The procedure was followed according to the user guide of ZipTip pipette tips which describes necessary steps to bind, wash and elute the peptide fractions on the resin. Before beginning ZipTip procedure, 20 µl of buffer (2% acetonitrile, 0.1% formic acid) was added into fractions. All fractions were waited during 15 seconds in an ultrasonic bath and for 5 seconds on vortex mixer three times, consecutively before centrifugation at 14.000 rpm for 10 minutes. Firstly, ZipTip pipette tips were treated with 10 µl of 100% acetonitrile three times to wet the resin. 10 µl of %1 formic acid was pipetted three times to condition the tip. Then, the sample was aspirated and dispensed 15 cycles for maximum binding to the tip without introducing air into the sample. The tip was washed by 10 µl of 1% formic acid three times and dispensed to waste. The peptides were eluted by 50% and 75%

acetonitrile to a total elution volume of 5 µl into a clean 0.5 ml low-bind microcentrifuge tube (Eppendorf). 25 µl of 0.1% formic acid was added into sample for adjustment of total volume to 30 µl.

Table 3.1. LC Time Programme

Time (min)	Mobile phase B
10.00	0%
10.00	5%
50.00	25%
57.00	30%
64.00	40%
71.00	80%
80.00	80%
90.00	0%
95.00	0%

Each fraction of the sample was analyzed by reversed-phase LC-ESI-MS/MS using Ultimate 3000 HPLC system (Dionex) coupled to LTQ XL ion trap mass spectrometer (Thermo Scientific). To operate analysis, tubing of the HPLC system was connected to electrospray. The following conditions and parameters were operated for HPLC process: column brand was Sigma Supelco Ascentis ® with the size of 15 cm × 500 µm; 2.7 µm at 25 ° C column oven; injection volume was 10 µl; flow rate was defined as 5 µl/minute.

Mobile phase A and B were prepared as 0.1% formic acid in water and 0.1% formic acid in acetonitrile, respectively. Throughout the experiment ion spray voltage for electrospray source was maintained at 3.00 kV. Capillary temperature and voltage were kept at 200° C and 6.00 V, respectively. In addition, LC-MS/MS time programme and data dependent acquisition parameters for mass spectrometry were set as in Table 3.2. Each kind of *T. vaginalis* protein digests which has 12 fractions was analyzed individually. When the first run of 12 fractions completed, the second run was performed for each fraction as a technical repeat. MS/MS raw data file of each fraction for two runs was converted to mascot generic file format (mgf) by using Thermo Proteome Discoverer 1.4. Then, all converted files were merged to a single file to make use of database search.

Table 3.2. LC-MS/MS time programme and data dependent acquisition parameters

Time (min)	Mobil Phase B
0.00	2%
4.00	2%
5.00	5%
6.00	8%
36.00	22%
39.00	35%
42.00	60%
45.00	90%
50.00	90%
55.00	2%
58.00	2%

Acquire time (minutes)	60
Scan events	15
Scan ranges (m/z)	400-1800
Activation type	CID
Isolation width (m/z)	1.00
Normalized collision energy	35
Activation time Q (ms)	30
Polarity	Positive
Data type	Centroid
Signal threshold	5000

3.5. Database Search

The name of the parasite *T. vaginalis* was written to taxonomy part under supporting data section of UniProt website for downloading *T. vaginalis* database which has 50.190 entries (www.uniprot.org). MS/MS data were extracted and searched against the UniProt database using Mascot Server 2.3 (www.matrixscience.com). The following parameters were operated: taxonomy was selected as all entries, enzyme was trypsin with allowing up to one missed cleavage, fixed modification was set as cysteine carbamidomethylation, oxidation of methionine was set as variable modification, peptide and MS/MS tolerances were 1.0 and 0.5 Da respectively, peptide charges were defined as +2 and +3. Monoisotopic mass was chosen, and the instrument was selected as Ion Trap. Decoy search which produces randomized or reversed sequences for database validation was performed by clicking on decoy checkbox. Percolator score option which aims to differentiate positive and negative peptide matches for decreasing percentage of false discovery rate (FDR) was selected and results were filtered. Venn diagram were drawn on the free website which was created by VIB from Ghent University to distinguish differential and common proteins between sensitive-1, sensitive-2, and resistant strains.

CHAPTER 4

RESULTS AND DISCUSSION

In this thesis, mass spectrometry-based qualitative shotgun proteomic analysis was performed to compare different proteins in drug-resistant and sensitive strains of parasite T. *vaginalis* which causes trichomoniasis. After protein extraction of one resistant and two sensitive strains (sensitive-1, sensitive-2), proteins were digested into peptides and they were fractioned chromatically using high pH-reversed phase chromatography which followed by reduction to 12 fractions with the concatenation of fractions. LC-MS/MS analysis was operated, and database search was carried out against the UniProt database using the Mascot search engine to expound on differential proteins.

Before protein level comparison of all samples, Mascot searches were operated to examine protein distribution. Default MS and MS/MS tolerance values of Mascot search engine is 1.2 Da and 0.6 Da, respectively. It is known that specification of too wide or narrow mass tolerance results in lower ion scores.

Although different parameters were tested, more appropriate tolerances were determined as 0.5 Da and 1.0 Da which were concluded with improved results. To introduce reliable peptide spectrum match (PSM) and to acquire minimal false discovery rate (FDR), target sequence database was also searched against randomized (decoy) database. Besides, results were again filtered for percolator scores to discriminate correct and incorrect spectrum. In this way, peptide matches were increased, sensitivity was improved, and FDR value was retained under 1% for results. All filtrations were evaluated in order to obtain the most reliable results.

6 shotgun proteomic results were obtained entirely from 2 set of samples, and each set involves resistant, sensitive-1, and sensitive-2 proteins. First, each set was compared separately to indicate different proteins. Afterward, it was tried to interpret differences in protein level between 2 sets and to explain what the reason for the difference might be. Venn diagrams were utilized to distinguish variation between drug resistant and sensitive samples (Figure 4.1). The diagrams were arranged for both characterized and uncharacterized proteins which were acquired after database search.

The left diagram demonstrates that identified total protein numbers are 345 for resistant, 364 for sensitive-1, and 386 for sensitive-2 strains in the first set. On the other

hand, identified protein numbers are 254 for resistant, 329 for sensitive-1, and 206 for sensitive-2 strains in the second set.

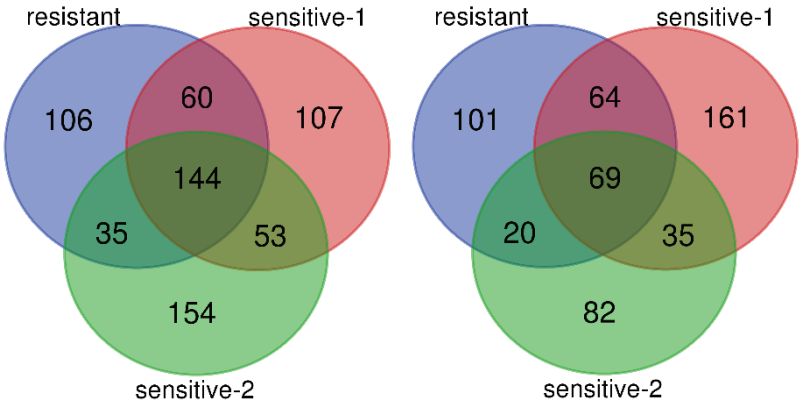


Figure 4.1. Venn diagrams showing the numbers of unique and overlapping proteins identified in first (left) and second (right) sets

Table 4.1, Table 4.2, and Table 4.3 represents differential proteins of resistant, sensitive-1 and, sensitive-2 strains for set 1, respectively. Table 4.4, Table 4.5, and Table 4.6 shows differential proteins of resistant, sensitive-1 and, sensitive-2 strains for set 2, respectively. Only characterized proteins in the database were demonstrated as differential protein lists with their score values and peptide match numbers. Uncharacterized proteins were also found for both sets. However, uncharacterized proteins were not included in the table because probably they were unstudied, and their function are unknown. Accession code of these uncharacterized proteins was given in Appendix A. Also, common proteins between resistant and sensitive-1 strains, resistant and sensitive-2 strains, and between all three strains were demonstrated including unknown proteins in Appendix B. Mascot search reports individual ion score which indicates identity or extensive homology for the search and this value was given as greater than 13 for all searches in which MS and MS/MS tolerance parameters were 0.5 Da and 1.0 Da, separately. Thus, the score values of all proteins were higher than 13. The identified proteins were classified based on their molecular function with Gene Ontology (GO) classification using UniProt database as follows: binding, catalytic activity, structural molecule activity, antioxidant activity, transporter and transmembrane transporter, enzyme regulator, transcriptional coregulator and corepressor, molecular function regulator, protein tag and, clathrin adaptor activity proteins.

Table 4.1. Characterized differential proteins of resistant strain in set 1

Accession Code	Protein Name	Score	Matches
A2EY23	Cytoplasmic heat shock protein 70, putative	1391	13
A2DHT2	Glyceraldehyde-3-phosphate dehydrogenase	2747	9
A2FVK7	Succinate--CoA ligase [ADP-forming] subunit beta, mitochondrial	391	9
O97108	Elongation factor 1-alpha (Fragment)	2111	8
Q03184	Succinate--CoA ligase [ADP-forming] subunit beta, hydrogenomal	1041	7
A2E1Z6	Ribosomal protein, putative	128	4
A2FCM7	Proteasome subunit alpha type	213	3
Q76KS7	Ribosomal protein S15a	1671	2
A2E0I5	Small GTP-binding protein, putative	391	2
A2E6S8	RNA-binding protein, putative	272	2
A2FL91	Ribosomal protein L5	237	2
A2FEV4	Chain A, Methionine Gamma-Lyase	231	2
A2DW27	Aminotransferase, class V family protein	220	2
A2E7V4	Ribosomal protein L18ae, putative	192	2
A2HLJ0	Cysteine protease, putative	168	2
A2E2L7	F-actin capping protein alpha subunit, putative	148	2
A2FIM7	Hydroxylamine reductase, putative	123	2
A2DRE4	Clathrin light chain	103	2
Q27096	TvhydB protein	97	2
Q4G290	GTP-binding protein YPTM2, putative	66	2
E5FCB1	Putative actin depolymerizing factor (Fragment)	59	2
A2EP03	60S ribosomal protein L18a	44	2
A2EBD1	60S ribosomal protein L6, putative	41	2
A2DNC2	Centrin, putative	39	2
A2FGT3	Aminopeptidase	28	2
A2DAU3	NAC domain containing protein	715	1
Q6IV59	Thioredoxin	454	1
A2DW84	Ribosomal protein L7Ae, putative	237	1
Q27117	Tubulin beta chain (Fragment)	206	1
A2G6W3	Histidine acid phosphatase family protein	196	1
A2EJW5	Flavodoxin-like fold family protein	156	1
A2E7I0	Aminotransferase, classes I and II family protein	148	1
A2EVR5	Ribosomal protein S14, putative	134	1
A2EP06	Pyridine nucleotide-disulphide oxidoreductase family protein	128	1
A2DIW6	Ribosomal protein L36e, putative	104	1
A2F078	Methionine aminopeptidase 2	91	1
A2G4E1	Ribosomal protein L13e, putative	83	1
A2FF20	Calcium-transporting ATPase	82	1

(cont. on next page)

Table 4.1. (cont.)

Accession Code	Protein Name	Score	Matches
A2DNV5	60S ribosomal protein L22-1, putative	82	1
A2F468	ATP-dependent RNA helicase p47, putative	77	1
A2DKF5	Ubiquitin family protein	68	1
A2EPZ2	Eukaryotic release factor 1, putative	64	1
A2EK65	Ribosomal protein S10p/S20e, putative	61	1
A2DIE4	40S ribosomal protein S16, putative	60	1
A2E3E3	60S ribosomal protein L7a, putative	51	1
A2E2R4	Phosphoglycerate mutase	49	1
A2FGS9	NifU-like domain containing protein	48	1
A2E1R6	60S ribosomal protein L32, putative	44	1
A2EMJ9	Ferredoxin 5	42	1
A2FDM1	H/ACA ribonucleoprotein complex subunit	41	1
A2DSC4	Calmodulin, putative	40	1
Q9XYL4	Putative NADPH-dependent butanol dehydrogenase	40	1
A2DFT9	Phosphofructokinase family protein	38	1
A2E7E7	CMGC family protein kinase	37	1
A2ED65	DJ-1 family protein	36	1
A2FGA5	Serine/threonine-protein phosphatase	31	1
A2DPT9	ZPR1-related zinc finger protein	31	1
A2EL04	Dolichyl-diphosphoolig accharide--protein glycosyltransferase subunit 1	30	1
A2EFM7	40S ribosomal protein S6	27	1
A2FBP0	Aldehyde dehydrogenase	26	1
A2EDI7	CK1 family protein kinase	23	1
A2E4U5	Maf protein	22	1
A2GGN6	Multidrug resistance protein, putative	21	1
A2ESJ6	Chaperonin subunit alpha1 CCTalpha, putative	20	1
A2FA79	Hypothetical hydrolase-like protein, putative	19	1
A2FR27	Dynamin central region family protein	18	1
A2FUT1	Putative tryptophanyl-tRNA synthetase	18	1
A2FPV5	Variant SH3 domain containing protein	17	1
A2EB65	G-protein alpha subunit, putative	16	1
A2EFC7	Major Facilitator Superfamily protein	15	1
A2E7D6	Leucine Rich Repeat family protein	15	1
A2DTZ0	Clan SB, family S8, subtilisin-like serine peptidase	14	1
A2DBM8	F-actin capping protein, beta subunit, putative	14	1
A2DXE8	Fibrillarin, putative	14	1
A2FDP9	Metalloenzyme superfamily protein	14	1
A2DH37	DEAD/DEAH box helicase family protein	13	1

Table 4.2. Characterized differential proteins of sensitive-1 strain in set 1

Accession Code	Protein Name	Score	Matches
A2DSF6	Elongation factor 1-alpha	2266	7
A2DRM7	Phosphofructokinase family protein	273	3
A2FZR4	NAC domain containing protein	642	3
A2EUY0	Ornithine carbamoyltransferase family protein	356	3
A2DNM4	Phosphoglycerate mutase	225	3
A2DAA4	Ribosomal protein, putative	195	3
A2E211	Ribosomal protein S15a, putative	607	2
A2E264	Ribosomal protein S14, putative	344	2
Q49P75	Cathepsin L-like cysteine proteinase	292	2
A2FQ07	Viral A-type inclusion protein, putative	227	2
A2D895	40S ribosomal protein S23, putative	183	2
A2DPX6	Ribosomal protein L5	117	2
A2DFD2	Dynamin central region family protein	108	2
A2DKP4	40S ribosomal protein S17-B, putative	102	2
A2DJE8	Nucleoside diphosphate kinase	71	2
A2FJT2	BTB/POZ domain containing protein	50	2
A2EZN7	Clan CA, family C1, cathepsin L-like cysteine peptidase	44	2
A2DMN0	Heat shock cognate protein, putative	38	2
A2EA92	60S ribosomal protein L11, putative	38	2
A2GUB1	Clan M-, family M49, dipeptidylpeptidase III-like metallopeptidase (Fragment)	31	2
A2DKF7	Ribosomal protein L13e, putative	27	2
A2FJ82	Hsp90 protein	21	2
A2EUZ9	Kelch motif family protein	21	2
A2E257	Ribosomal protein L7Ae, putative	421	1
A2E6R2	Ribosomal protein L29, putative	294	1
A2FNQ6	XYPPX repeat family protein	196	1
A2EJC2	DJ-1 family protein	166	1
A2E296	Ubiquitin family protein	127	1
A2DR67	Nitroreductase family protein	104	1
A2ETG5	Serine/threonine-protein ph phatase	79	1
A2EAX7	Ankyrin repeat protein, putative	53	1
A2FHJ2	AGC family protein kinase	51	1
A2EYL3	Flavodoxin-like fold family protein	50	1
A2DYP9	60S ribosomal protein L18, putative	48	1
A2E5T1	F-actin capping protein, beta subunit, putative	47	1
A2FRG0	Ribose 5-phosphate isomerase B family protein	44	1
A2DIJ0	XRP2, putative	41	1
A2FJT4	ENTH domain containing protein	40	1

(cont. on next page)

Table 4.2. (cont.)

Accession Code	Protein Name	Score	Matches
A2DXW7	Tryparedoxin peroxidase, putative	39	1
A2DD10	Actin, putative	34	1
A2EQY3	Cysteine synthases family protein	32	1
A2F7M8	NEDD8-activating enzyme E1 regulatory subunit	32	1
A2F136	HMG box family protein	30	1
A2EFL5	Galactokinase family protein	30	1
A2D8G1	60S ribosomal protein L19-3, putative	30	1
E5FCB6	Putative actin depolymerizing factor (Fragment)	30	1
A2FNG8	Cytochrome b5-like Heme/Steroid binding domain containing protein	28	1
Q5QGR5	Prolyl-tRNA synthetase (Fragment)	28	1
Q94832	HSP70 (Fragment)	27	1
A2FR65	Basic proline-rich protein, putative	26	1
A2ER18	Zinc finger in N-recognin family protein	25	1
A2G7J3	Ser/Thr protein phosphatase, putative	24	1
A2FII9	Mitochondrial carrier protein	23	1
A2DCH7	Ribosomal protein L10a	23	1
A2FCW4	64kDa iron hydrogenase, putative	22	1
A2EKG8	AMP-binding enzyme family protein	20	1
A2DSR2	STE family protein kinase	19	1
A2ECH6	SNF2 family N-terminal domain containing protein	19	1
A2DVE6	Methyltransferase, putative	17	1
A2EN74	DNA-directed RNA polymerase subunit beta	16	1
A2DZI2	Adaptin N terminal region family protein	14	1
A2E5L6	Epimerase/dehydratase, putative	14	1
A2D7S4	Phosphoprotein phosphatase, putative	14	1

Distribution of identified proteins according to molecular function was revealed as bar graphs in Figure 4.2 and Figure 4.3.

For detailed research, proteins of each molecular function that identified in the bar graphs were compared between all strains; resistant, sensitive-1 and, sensitive-2 strain for both sets. Identified proteins were demonstrated as Venn diagrams for binding, catalytic and, structural activity. The focus was on the proteins with catalytic and binding activities because the number of identified proteins in these groups were higher than other activity types for both sets. Thus, proteins with catalytic and binding activity in resistant strain might be related to drug resistance mechanism. Change in protein numbers are clearly observed for different molecular activities in Venn diagrams of two sets.

Table 4.3. Characterized differential proteins of sensitive-2 strain in set 1

Accession Code	Protein Name	Score	Matches
O96524	Alpha-actinin	3175	20
A2D9B6	AP65-3 adhesin	6410	15
Q27089	Pyruvate:ferredoxin oxidoreductase proprotein	3249	13
A2EVX8	Pyridine nucleotide-disulphide oxidoreductase family protein	1046	12
A2FLL6	Phosphoenol pyruvate carboxykinase, putative	1411	9
A2EJM3	Histidyl-tRNA synthetase family protein	306	8
F8QX91	Elongation factor 1-alpha (Fragment)	4155	7
A2DMN2	Malate dehydrogenase	3012	5
A2DBB7	40S ribosomal protein S4	614	5
A2DGQ5	Polyubiquitin, putative	1089	4
A2FVM1	60S acidic ribosomal protein P0	825	4
A2DHT0	QXW lectin repeat family protein	561	4
A2GAU8	Tryparedoxin peroxidase, putative	663	3
A2E4D0	Ribosomal protein, putative	585	3
Q6UJG4	Thioredoxin peroxidase	538	3
E5FCB7	Arp2/3 complex 34 kDa subunit (Fragment)	445	3
Q4KXQ3	14-3-3 protein	429	3
O15568	Succinate--CoA ligase [ADP-forming] subunit alpha, mitochondrial	290	3
A2E0F1	Valyl tRNA Synthetase, putative	243	3
A2EGX9	Triosephosphate isomerase	143	3
A2EP01	Enolase family protein	127	3
A2FZX1	Thioredoxin	120	3
A2G8D9	Glycine cleavage H-protein	120	3
A2F4M9	Amylo-alpha-1,6-glucosidase family protein	88	3
A2E0S8	4-alpha-glucanotransferase, putative	81	3
A2G7V0	ThiF family protein	237	2
A2FLG0	40S ribosomal protein SA	193	2
A2DHC3	Coronin	158	2
A2DLU0	Nicotinate phosphoribosyltransferase, putative family protein	115	2
A2ECU5	UTP--glucose-1-phosphate uridylyltransferase family protein	109	2
A2ECS1	Cysteinyl-tRNA synthetase family protein	93	2
A2FXM6	Clan CD, family C13, asparaginyl endopeptidase-like cysteine peptidase	89	2
A2EFE2	Lactate dehydrogenase isozyme 2, putative	88	2
A2D968	Aminotransferase, class V family protein	81	2
A2ESJ9	Thioredoxin peroxidase, putative	59	2
A2DFV0	Cysteine synthases family protein	50	2
A2D7Y6	Adenylate kinase family protein	46	2
A2FAG5	Adaptin N terminal region family protein	43	2
Q4G2D6	Ras-like GTP-binding protein YPT1, putative	40	2

(cont. on next page)

Table 4.3. (cont.)

Accession Code	Protein Name	Score	Matches
A2EA56	Galactokinase family protein	37	2
A2DNM5	Ribosomal protein L5	473	1
A2DFZ5	40S ribosomal protein S5, putative	252	1
A2DZL6	Purine nucleoside phosphorylase, putative	247	1
A2EXA4	NAC domain containing protein	217	1
A2FU96	AP complex subunit beta	189	1
A2D8Z0	Ribosomal protein L13e, putative	178	1
A2DLV9	Rac1, putative	161	1
A2EI94	Oligosaccharyl transferase STT3 subunit family protein	144	1
A2DM53	Ribosomal protein L22, putative	139	1
A2DCB1	Synaptobrevin family protein	119	1
A2EH55	MBOAT family protein	119	1
P49983	Adenylate kinase	118	1
A2DG97	Centrin, putative	108	1
A2E709	V-type proton ATPase subunit	100	1
A2FW82	Viral A-type inclusion protein, putative	92	1
A2EGK6	Inositol monophosphatase family protein	84	1
A2EG54	Ribosomal protein S19e, putative	82	1
A2G442	Ribosomal protein L32, putative	82	1
A2G527	Adseverin, putative	80	1
A2EGV2	Eukaryotic translation initiation factor, putative	79	1
A2D7L9	Ribosomal protein L7Ae, putative	74	1
A2ERK0	Dolichyl-diphosphooligosaccharide--protein glycosyltransferase 48 kDa subunit	66	1
A2DHY3	Cofilin/tropomyosin-type actin-binding protein	64	1
A2D8F3	GTP-binding nuclear protein	63	1
A2F412	Actin-related protein 2/3 complex subunit 5	61	1
A2F3Z2	Flavodoxin-like fold family protein	59	1
A2GJ01	Surface protein, putative	56	1
A2DIP8	SNARE protein, putative	56	1
A2FFS8	40S ribosomal protein S21	55	1
Q9NI37	Hydrogenosomal membrane protein 31	48	1
A2GEK7	PGP1, putative (Fragment)	47	1
A2FVV6	Ubiquitin family protein	45	1
A2F2D7	Pyridoxal kinase family protein	45	1
A2EHV1	Prefoldin subunit 3	43	1
A2ELA9	TPR Domain containing protein	38	1
A2FKC5	Hsp20/alpha crystallin family protein	36	1
A2DG67	60S ribosomal protein L6, putative	36	1
A2FT42	AP complex subunit sigma	35	1

(cont. on next page)

Table 4.3. (cont.)

Accession Code	Protein Name	Score	Matches
A2FT42	AP complex subunit sigma	35	1
A2EPF2	Proteasome/cyclosome repeat family protein	34	1
A2DV55	NAD+ synthetase family protein	32	1
A2F0F6	Dihydrolipoyl dehydrogenase	32	1
A2FGT0	PX domain containing protein	31	1
A2EHL4	SacI homology domain containing protein	28	1
A0A0K0	MNE9RNA-directed RNA polymerase	27	1
Q4G2A2	Ras-like GTP-binding protein RYL1, putative	27	1
A2EJJ9	Endoribonuclease L-PSP family protein	27	1
A2F1G2	Mrp, putative	26	1
A2FIH1	Renin binding protein, putative	26	1
A2DJS3	CAMK family protein kinase	25	1
A2DXW5	Calmodulin, putative	25	1
A2FA28	Ankyrin repeat protein, putative	24	1
A2E498	Branched-chain-amino-acid aminotransferase	22	1
A2DFL2	GP63-like	19	1
A2EKF3	Mitochondrial-type HSP70, putative	19	1
A2DEJ9	40S ribosomal protein S6	19	1
A2DXW1	Coatomer subunit beta'	18	1
A2ECT4	Pyridoxal-phosphate dependent enzyme family protein	18	1
A2D7K5	Myosin-cross-reactive antigen, putative	17	1
A2D908	Generic methyltransferase, putative	17	1
A2DGG3	Histone H3-1	16	1
A2EZW3	Haloacid dehalogenase-like hydrolase family protein	16	1
A2DZU8	Glucosamine--fructose-6-phosphate aminotransferase, isomerizing family protein	15	1
A2E4K0	Tubulin-tyrosine ligase family protein	14	1
A2FRG8	Mob1/phocein family protein	13	1
O15564	Methionine gamma-lyase	13	1

Common proteins with identified molecular activities between sensitive-1 and sensitive-2 strains can be considered as more reliable and responsible for showing sensitivity for drug treatment rather than differential proteins in these strains. That is why, the focus was on the common proteins between sensitive-1 and sensitive-2 strains instead of differential ones. During the study, it was dealt with shared proteins between sensitive-1 and sensitive-2 strains and differential proteins in resistant strain. Table 4.7 and Table 4.8 represent common proteins between sensitive-1 and sensitive-2 strains and differential proteins for resistant strain, respectively for binding activity.

Table 4.4. Characterized differential proteins of resistant strain in set 2

Accession Code	Protein Name	Score	Matches
A2DGG1	Actinin, putative OS	550	10
A2EY23	Cytoplasmic heat shock protein 70, putative OS	618	9
A2D9B6	AP65-3 adhesin OS	481	9
A2EBB4	Enolase 2, putative OS	983	7
Q27088	Pyruvate:ferredoxin oxidoreductase A OS	560	6
A2EW14	Fructose-1,6-bisphosphate aldolase, putative OS	228	5
A2E1Z6	Ribosomal protein, putative OS	81	4
A2GL34	Clathrin and VPS domain-containing protein (Fragment) OS	114	3
Q8WRQ8	Glucose-6-phosphate isomerase (Fragment) OS	26	3
A2DGL5	Ubiquitin, putative OS	253	2
Q76KS7	Ribosomal protein S15a OS	224	2
A2F6B9	Malic enzyme OS	206	2
A2EUY0	Ornithine carbamoyltransferase family protein OS	130	2
A2FEV4	Chain A, Methionine Gamma-Lyase OS	97	2
A2DDM9	14-3-3 protein OS	83	2
A2F5C1	Small GTP-binding protein, putative OS	66	2
A2DJS9	EF hand family protein OS	65	2
Q27875	Histone H4 OS	55	2
E5FCB6	Putative actin depolymerizing factor (Fragment) OS	28	2
A2E709	V-type proton ATPase subunit OS	28	2
A2DYA1	Protein kinase, putative OS	15	2
A2DL09	Chaperonin, putative OS	124	1
A2DWC1	DnaK protein OS	123	1
A2D8Z0	Ribosomal protein L13e, putative OS	105	1
A2D7V2	Cytidine deaminase OS	80	1
A2DAK7	Proteasome endopeptidase complex, putative OS	79	1
A2E4K8	Serine/threonine-protein phosphatase OS	68	1
A2G3J8	Tetraspanin family protein OS	65	1
A2E3S6	Haloacid dehalogenase-like hydrolase family protein OS	63	1
A2FIP8	L-lactate dehydrogenase, putative OS	62	1
A2EI52	Adenylate kinase family protein OS	58	1
A2FPE2	Oxidoreductase, aldo/keto reductase family protein OS	46	1
A2EZW4	6-phosphogluconate dehydrogenase, decarboxylating OS	44	1
P90622	Chaperonin 60 (Fragment) OS	44	1
A2DSC4	Calmodulin, putative OS	42	1
Q76KT0	Ribosomal protein L5 OS	42	1
A2EDY9	Glycerol-3-phosphate dehydrogenase [NAD(+)] OS	41	1
A2F2J0	Ribosomal protein L32, putative OS	40	1
A2FAG1	Glutamine synthetase, catalytic domain containing protein OS	39	1

(cont. on next page)

Table 4.4. (cont.)

Accession Code	Protein Name	Score	Matches
A2DYJ9	PCI domain containing protein OS	39	1
A2FR77	Phenylalanyl-tRNA synthetase, beta subunit family protein OS	33	1
A2FF20	Calcium-transporting ATPase OS	28	1
A2F763	Clan CA, family C19, ubiquitin hydrolase-like cysteine peptidase OS	25	1
A2DUS1	KE2 family protein OS	25	1
A2EPZ2	Eukaryotic release factor 1, putative OS	24	1
A2DRM7	Phosphofructokinase family protein OS	24	1
Q4G2C9	Ras-related protein Rab11C, putative OS	24	1
Q76KS9	Ribosomal protein L8 OS	23	1
A2FKI0	Translation initiation factor SUI1 family protein OS	22	1
A2DLQ5	Clathrin light chain OS	21	1
A2D8X5	Chaperonin subunit eta CCTeta, putative OS	21	1
A2D8F3	GTP-binding nuclear protein OS	19	1
A2EUN9	HMG box family protein OS	18	1
A2FKX4	CAMK family protein kinase OS	18	1
A2FDT5	Clan ME, family M16, insulinase-like metallopeptidase OS	18	1
A2DUV8	Ribosomal protein L23, putative OS	18	1
A2FGT3	Aminopeptidase OS	18	1
A2FIM7	Hydroxylamine reductase, putative OS	17	1
A2FXP4	Actin-related protein 2/3 complex subunit 4 OS	17	1
A2F6X9	Tctex-1 family protein OS	16	1
A2F719	Lysine--tRNA ligase OS	16	1
A2E6B9	SnoRNA binding domain containing protein OS	16	1
A2EIF8	Transketolase family protein OS	15	1
Q9NCN7	Eukaryotic release factor 3 GTPase subunit OS	15	1
A2G7K0	Lipase family protein OS	15	1
A2EGN0	Bromodomain containing protein OS	14	1
A2ECH6	SNF2 family N-terminal domain containing protein OS	14	1
A2DTZ0	Clan SB, family S8, subtilisin-like serine peptidase OS	14	1
A2E342	Adenosylhomocysteinase OS	14	1
A2FDS0	Ribosomal protein S3Ae, putative OS	13	1
A2DJE6	Phenylalanyl-tRNA synthetase, alpha subunit family protein OS	13	1

Table 4.9 and Table 4.10 show common proteins between sensitive-1 and sensitive-2 strains and differential proteins for resistant strain, respectively for catalytic activity.

Table 4.5. Characterized differential proteins of sensitive-1 strain in set 2

Accession Code	Protein Name	Score	Matches
O77068	Alpha actinin (Fragment)	2176	19
O76307	Endoplasmic reticulum heat shock protein 70 (Fragment)	327	11
Q4KY23	Pyruvate:ferredoxin oxidoreductase A-like protein	1667	8
A2E9H3	Pyrophosphate--fructose 6-phosphate 1-phosphotransferase 2	526	8
Q5EFD8	Enolase	1739	7
A2EBX0	Succinate--CoA ligase [ADP-forming] subunit beta, mitochondrial	381	5
A2F6S0	Fructose-1,6-bisphosphate aldolase	326	5
A2FAF3	Fructose-1,6-bisphosphate aldolase, putative	324	5
A2ELB9	Enolase family protein	284	5
Q27097	Polyubiquitin (Fragment)	596	4
Q5KTX1	Cytosolic-type hsp90 (Fragment)	281	4
A2DYG1	UTP--glucose-1-phosphate uridylyltransferase family protein	88	3
A2FT79	Proteasome subunit alpha type	102	2
A2DIE5	Chaperonin, putative	100	2
A2EFS6	Guanine nucleotide-binding protein beta subunit, putative	96	2
A2E7V4	Ribosomal protein L18ae, putative	95	2
Q5QGR0	Alanyl-tRNA synthetase (Fragment)	72	2
A2DHT0	QXW lectin repeat family protein	64	2
A2F136	HMG box family protein	60	2
A2ERB4	Seryl-tRNA synthetase family protein	51	2
A2E2R4	Phosphoglycerate mutase	43	2
A2DG67	60S ribosomal protein L6, putative	43	2
A2EJM3	Histidyl-tRNA synthetase family protein	41	2
A2FE23	Actin-related protein 2/3 complex subunit 4	41	2
A2D908	Generic methyltransferase, putative	38	2
A2DC16	Tubulin beta chain	33	2
A2EC78	Glucose-6-phosphate 1-dehydrogenase	33	2
A2E5T1	F-actin capping protein, beta subunit, putative	32	2
Q6UJG4	Thioredoxin peroxidase	27	2
A2FB71	Nascent polypeptide-associated complex subunit beta	243	1
Q4G2D2	Small GTP-binding protein, putative	187	1
A2DTR8	Ribosomal protein S13p/S18e, putative	151	1
A2FCA3	Cytidine deaminase	137	1
A2FGB9	Transcription factor BTF3, putative	121	1
A2EB65	G-protein alpha subunit, putative	88	1
A2FU96	AP complex subunit beta	84	1
A2EYY6	Oxidoreductase, aldo/keto reductase family protein	83	1
A2DST7	40S ribosomal protein S26	83	1
A2DVD8	Ribosomal protein S19, putative	83	1

(cont. on next page)

Table 4.5. (cont.)

Accession Code	Protein Name	Score	Matches
A2ENN9	Profilin	73	1
A2D7H9	Superoxide dismutase	73	1
A2FCW8	V-type proton ATPase subunit C	68	1
A2DW27	Aminotransferase, class V family protein	63	1
A2DR42	Chaperonin subunit zeta CCTzeta	62	1
A2FHI8	D-isomer specific 2-hydroxyacid dehydrogenase, putative	59	1
A2FDE0	Ribosomal protein S10p/S20e, putative	58	1
A2D7P3	Ribosomal protein L7Ae, putative	58	1
A2E4A6	DnaK protein	55	1
A2E1N5	Hybrid-cluster protein, putative	51	1
A2F1G9	Eukaryotic initiation factor 1A family protein	50	1
A2ER05	Obg-like ATPase 1	47	1
A2EU62	Purine nucleoside phosphorylase, putative	47	1
A2E256	60s Acidic ribosomal protein	43	1
A2ETD8	Chaperonin 60, putative	41	1
A2DJT4	Guanine nucleotide regulatory protein, putative	39	1
A2E6V5	Phosphorylase family protein	38	1
A2DK87	Phosphoribulokinase / Uridine kinase family protein	37	1
A2D895	40S ribosomal protein S23, putative	36	1
A2DIW1	Ribosomal protein L23, putative	32	1
A2G527	Adseverin, putative	32	1
A2E6G7	Clan M-, family M49, dipeptidylpeptidase III-like metallopeptidase	32	1
A2DA37	Peptidylprolyl isomerase	32	1
A2D900	TvhydB protein, putative	31	1
A2EU29	Serine/threonine-protein phosphatase	31	1
A2EFC7	Major Facilitator Superfamily protein	29	1
A2DJQ1	Ribosomal protein L24e, putative	28	1
A2DYA8	40S ribosomal protein S29, putative	27	1
A2G3F8	Ribulose-phosphate 3-epimerase	27	1
A2DDT1	Ribosomal protein L6, putative	26	1
A2G4G3	Auxin Efflux Carrier family protein	26	1
A2DIE4	40S ribosomal protein S16, putative	25	1
A2EKG8	AMP-binding enzyme family protein	25	1
A2E6Z8	Ribosomal protein L8, putative	25	1
A2DPT9	ZPR1-related zinc finger protein	25	1
A2ECR7	Ribosomal protein L15	24	1
A2FJL6	Glycerol-3-phosphate dehydrogenase [NAD(+)]	24	1
A2F8Y2	Clan MG aminopeptidase P-like metallopeptidase	24	1
Q5QGR5	Prolyl-tRNA synthetase (Fragment)	23	1

(cont. on next page)

Table 4.5. (cont.)

Accession Code	Protein Name	Score	Matches
A2DS48	PH domain containing protein	23	1
A2DIW6	Ribosomal protein L36e, putative	22	1
A2FSV6	Chaperonin subunit eta CCTeta, putative	22	1
A2E4F8	Adaptin N terminal region family protein	22	1
A2DEJ9	40S ribosomal protein S6	21	1
A2DA25	ATPase, AAA family protein	21	1
A2H0N5	2-dehydropantoate 2-reductase	18	1
A2DEZ9	Variant SH3 domain containing protein	18	1
A2FNN7	Cytosolic repetitive antigen, putative	17	1
A2D9P9	F/Y-rich N-terminus family protein	17	1
A2FED9	V-type proton ATPase subunit a	17	1
A2DV34	Dynein heavy chain family protein	16	1
A2G5B2	Surface antigen BspA-like	16	1
A2FCW4	64kDa iron hydrogenase, putative	16	1
A2FEM3	Haloacid dehalogenase-like hydrolase family protein	15	1
A2D7P7	Sm protein	15	1
A2DAE1	Clan MA, family M3, oligopeptidase A-like metallopeptidase	15	1
A2EI45	Kinesin-like protein	15	1
A2ER70	Domain found in IF2B/IF5 family protein	15	1
A2DBZ1	Ankyrin repeat protein, putative	14	1

Table 4.11 and 4.12 demonstrate common proteins between sensitive-1 and sensitive-2 strains and differential proteins for resistant strain, respectively for structural molecular activity. Identified proteins in the sets were labelled with a tick mark, whereas the negative mark was used for absent proteins in the sets. The presence of any protein in both sets shows the accuracy and reproducibility of the experiment. Common proteins observed in both sets are included at the top of the tables with yellow color. Besides, ferredoxin 5 protein has both catalytic and binding activity. It was identified in only resistant strain for set 1and labelled with green color.

According to Table 4.7 and Table 4.8, it is observable that common proteins of sensitive strains with binding activity are not the same for two sets and different proteins present in both sets. However, six proteins were differentially identified as eukaryotic release factor (putative), chain A methionine gamma-lyase, hydroxylamine reductase (putative), clathrin light chain and aminopeptidase, respectively for resistant strain in both sets.

Table 4.6. Characterized differential proteins of sensitive-2 strain in set 2

Accession Code	Protein Name	Score	Matches
Q27102	AP65-3 adhesin	1890	12
O00827	Cytoplasmic heat shock protein 70	930	12
A2DSV0	Ribosomal protein S3Ae, putative	173	6
Q27090	Hydrogenosomal malic enzyme subunit B proprotein	119	6
A2DHT2	Glyceraldehyde-3-phosphate dehydrogenase	329	5
A2DLD6	Malate dehydrogenase:SUBUNIT	95	5
A2HXI6	Uncharacterized protein (Fragment)	310	3
A2FBC9	4-alpha-glucanotransferase family protein	149	3
A2EA89	Valyl tRNA Synthetase, putative	87	3
A2DV55	NAD+ synthetase family protein	53	3
A2DGQ5	Polyubiquitin, putative	1193	2
A2DMN2	Malate dehydrogenase	567	2
A2G5E4	Ribosomal protein L32, putative	416	2
A2FSH1	60S ribosomal protein L18a	200	2
A2FQ69	Adenylate kinase	175	2
A2DYG8	Calcium-transporting ATPase	127	2
A2EL72	Clan MH, family M20, peptidase T-like metallopeptidase	72	2
A2E2L7	F-actin capping protein alpha subunit, putative	51	2
A2EVX8	Pyridine nucleotide-disulphide oxidoreductase family protein	44	2
A2F0H0	Transaldolase	19	2
Q3LS71	Ras family protein	176	1
A2FEH4	Oxidoreductase, aldo/keto reductase family protein	143	1
A2FFS8	40S ribosomal protein S21	140	1
A2DY01	Fructose-1,6-bisphosphate aldolase, putative	109	1
A2E6S8	RNA-binding protein, putative	92	1
A2DKP4	40S ribosomal protein S17-B, putative	82	1
A2DDW2	Ribosomal protein L7Ae, putative	72	1
A2DP65	Pyruvate kinase	72	1
A2F5J7	Chaperonin 60, putative	68	1
A2FS14	RanBP1 domain containing protein	67	1
A2HLJ0	Cysteine protease, putative	66	1
A2E6R2	Ribosomal protein L29, putative	56	1
A2ELR1	S-adenosylmethionine synthase	41	1
A2E6Z9	60S ribosomal protein L30, putative	40	1
A2E7Z7	Helicase conserved C-terminal domain containing protein	39	1
A2EXA4	NAC domain containing protein	39	1
A2DKF5	Ubiquitin family protein	35	1
A2D962	Glycosyltransferase family 28 C-terminal domain containing protein	33	1
A2G0N8	Coronin	32	1

(cont. on next page)

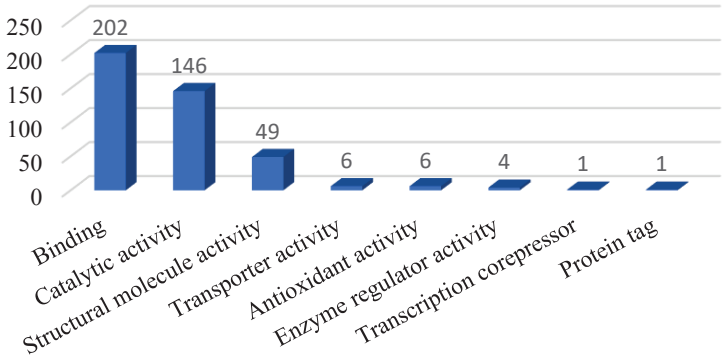
Table 4.6. (cont.)

Accession Code	Protein Name	Score	Matches
A2DHS6	Isochorismatase family protein	32	1
A2F2E6	Adenylate and Guanylate cyclase catalytic domain containing protein	28	1
A2F3B3	Small GTP-binding protein, putative	27	1
A2F106	UDP-glucose 6-dehydrogenase	26	1
A2DFI8	AP complex subunit beta	26	1
A2DI34	CAMK family protein kinase	25	1
A2E0I2	Adaptin N terminal region family protein	25	1
A2FDP9	Metalloenzyme superfamily protein	25	1
A2FX58	Centrin, putative	24	1
A2EQL6	Hypothetical UPF0311 protein CAC3321, putative	22	1
Q27117	Tubulin beta chain (Fragment)	21	1
A2ECS1	Cysteinyl-tRNA synthetase family protein	21	1
A2DUA7	QXW lectin repeat family protein	21	1
A2EN51	Ribosomal protein S19e, putative	20	1
A2ERK0	Dolichyl-diphosphooligosaccharide--protein glycosyltransferase 48 kDa subunit	19	1
A2EL04	Dolichyl-diphosphooligosaccharide--protein glycosyltransferase subunit 1	19	1
A2FGT0	PX domain containing protein	19	1
A2DLU0	Nicotinate phosphoribosyltransferase, putative family protein	18	1
A2EZN4	BTB/POZ domain containing protein	17	1
A2DQX2	DnaK protein	17	1
A2EK65	Ribosomal protein S10p/S20e, putative	15	1
A2GKT2	Ankyrin repeat protein, putative	15	1
A2DNM4	Phosphoglycerate mutase	14	1
A2D919	40S ribosomal protein S16, putative	14	1

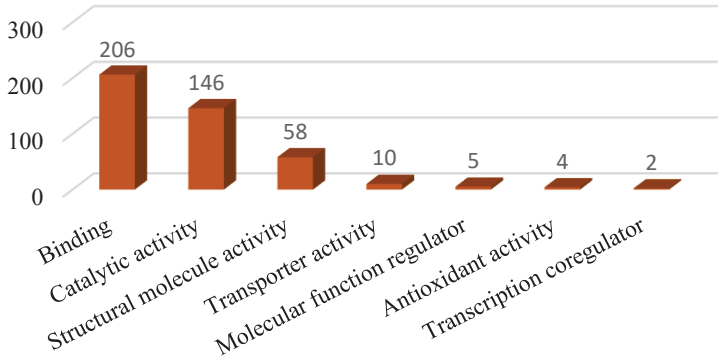
Although each differential protein in resistant strain might be responsible for resistance, however; identified six binding proteins of resistant strain can have higher effect for resistance mechanism because of identification in both sets which shows accuracy and repeatability of these proteins.

As shown in the Table 4.9, common proteins with catalytic activity between sensitive-1 and sensitive-2 strains are different in both sets. Fifteen of identified proteins belonged to the first set, while nine were found only in the second set. Although there were 20 and 24 differential proteins for set 1 and set 2 in resistant strain respectively, common proteins between two set have higher repeatability in different sets (Table 4.10).

a) Resistant



b) Sensitive-1



c) Sensitive-2

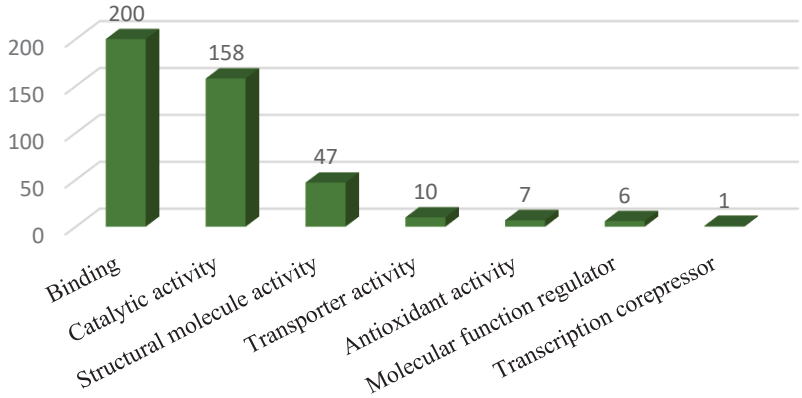
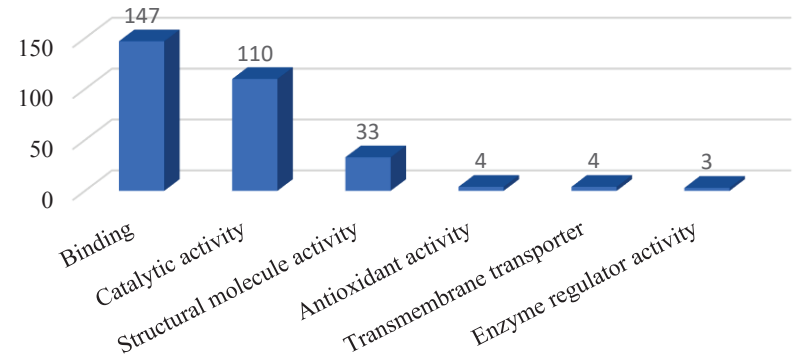
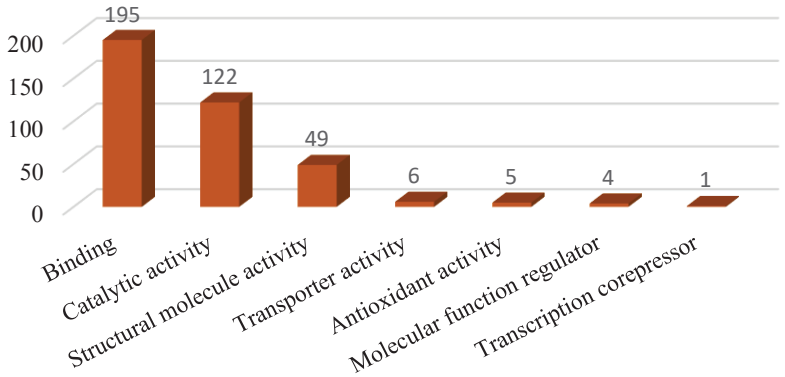


Figure 4.2. Molecular function classification of all detected proteins of *T. vaginalis* using Gene Ontology: a) Resistant, b) Sensitive-1, c) Sensitive-2 strain in set 1

a) Resistant



b) Sentisive-1



c) Sensitive-2

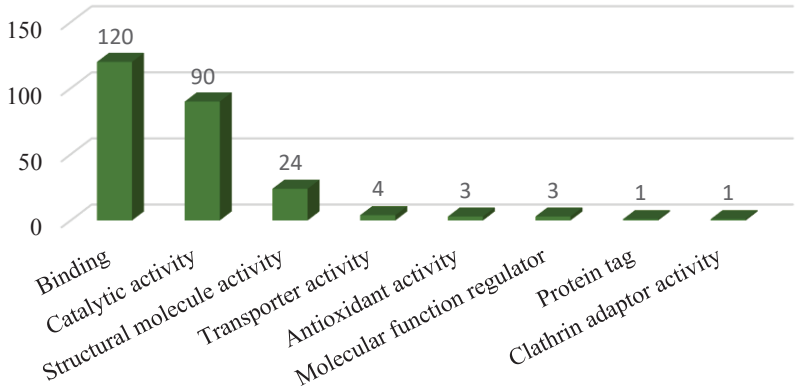
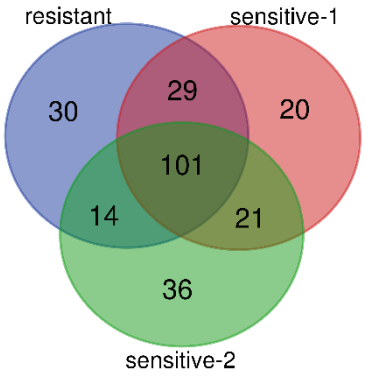


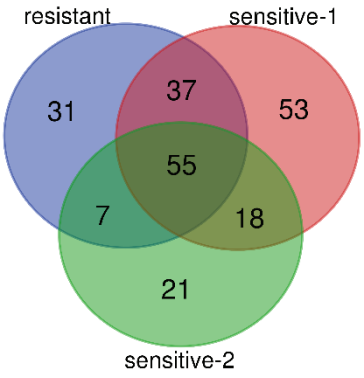
Figure 4.3. Molecular function classification of all detected proteins of *T. vaginalis* using Gene Ontology: a) Resistant, b) Sensitive-1, c) Sensitive-2 strain in set 2

a) Binding activity

Set 1

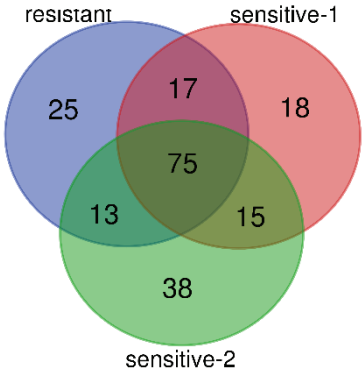


Set 2

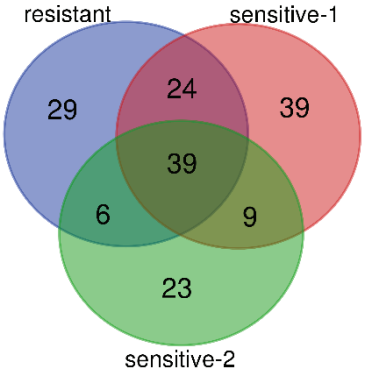


b) Catalytic activity

Set 1



Set 2



c) Structural molecular activity

Set 1

Set 2

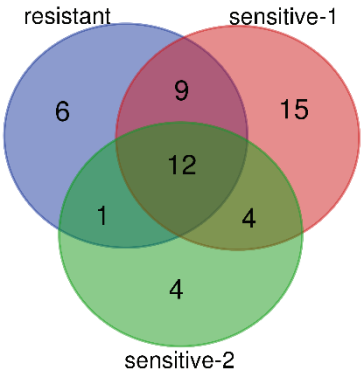


Figure 4.4. Venn diagrams showing the numbers of unique and overlapping proteins for activities: a) Binding activity b) Catalytic activity c) structural molecular activity in set 1 and set 2

Five proteins were differentially identified as chain A methionine gamma-lyase, glyceraldehyde-3-phosphate dehydrogenase (fragment), hydroxylamine reductase (putative), aminopeptidase and clan SB, family S8, subtilisin-like serine peptidase with catalytic activity for resistant strain in both sets. These five proteins may have a role in the formation of the resistance mechanism.

As for other activity types, half of the proteins in molecular activity for sensitive strains were found in the first set and the other half in the second set according to Table 4.11. There was no common protein observed in both sets, so the reproducibility of these proteins cannot be mentioned. On the other hand, there was only one differential protein, clathrin light chain, for resistant strain in both sets (Table 4.12). Since it is replicable in both sets, it may be mentioned the effect of protein on the resistance mechanism. Besides, other proteins identified in only one set might be responsible for resistance. However, any protein with transporter activity in both sets of resistant, sensitive-1 and, sensitive-2 strains was not considered as a difference in the resistance mechanism.

Differentially identified ferredoxin 5 and hydroxyl amine reductase (putative) proteins for only resistant strain were focused throughout the study. Because, these two proteins have iron-sulfur cluster binding activity and electron transfer activity. Iron-sulfur clusters (FeS) are best known for their role in the oxidation-reduction reactions. Also, iron concentration is always fluctuating in the vaginal environment, so profile of metronidazole resistance could be altered in *T. vaginalis*. It is known that the parasite requires high iron levels for optimal growth which depends on activities of FeS cluster containing proteins. In the recent study, metronidazole resistant strains of *T. vaginalis* were developed and evaluated if iron availability is important for the drug activation[67]. It was indicated that the iron content has a significant role in metronidazole action and likely to be related to FeS proteins involved in metronidazole activation. Thus, the increased number of iron-sulfur proteins exist in the resistant parasite.

PFOR, superoxide dismutase (SOD), thioredoxin reductase (TrxR), alcohol dehydrogenase (ADH 1) are studied proteins in the literature and identified as common proteins in sensitive-1, sensitive-2, and resistant strains in this study. Hydrogenosomal enzyme PFOR metabolizes pyruvate as acetate using acetyl CoA for ATP production and responsible for activating metronidazole for toxic nitro-radical formation.

Laboratory-induced metronidazole resistant T. vaginalis cell line was studied and demonstrated that enzyme activity and expression of PFOR was absent or very low[66].

Table 4.7. List of common proteins with binding activity between sensitive-1 and sensitive-2 strains

Protein Name	Set 1	Set 2
ATPase, AAA family protein	√	–
Clathrin and VPS domain-containing protein (Fragment)	√	–
RNA-binding protein	√	–
Enolase 3, putative	√	–
Polyadenylate-binding protein, putative	√	–
Cysteine synthases family protein	√	–
V-type proton ATPase subunit a	√	–
Ribosomal protein L8, putative	√	–
Ribosomal protein L15	√	–
Lysine--tRNA ligase	√	–
Small GTP-binding protein, putative	√	–
Serine palmitoyl transferase subunit, putative	√	–
Alcohol dehydrogenase	√	–
Guanine nucleotide regulatory protein, putative	√	–
Glucose-6-phosphate 1-dehydrogenase	√	–
Fructose-1,6-bisphosphate aldolase	√	–
Ras-related protein Rab11C, putative	√	–
Galactokinase family protein	√	–
Hydrogenosomal malic enzyme subunit B proprotein	√	–
40S ribosomal protein S26	√	–
40S ribosomal protein S2, putative	√	–
Alpha actinin-3	–	√
Aspartate aminotransferase	–	√
Ribosomal protein S14, putative	–	√
Dynein light chain	–	√
AP complex subunit beta	–	√
Chaperonin 60, putative	–	√
DnaK protein (Fragment)	–	√
Hsp90 protein	–	√
DALR anticodon binding domain containing protein	–	√
40S ribosomal protein S16, putative	–	√
Glutaminyl-tRNA synthetase family protein	–	√
Enolase 4, putative	–	√
60S ribosomal protein L22-1, putative	–	√
Cytoplasmic heat shock protein 70	–	√
GTP-binding protein SAR2, putative	–	√
Cofilin/tropomyosin-type actin-binding protein	–	√
T-complex protein 1 subunit delta	–	√
QXW lectin repeat family protein	–	√

Table 4.8. List of differential proteins with binding activity in resistant strain

Protein Name	Set 1	Set 2
Eukaryotic release factor 1, putative	√	√
Chain A, Methionine Gamma-Lyase	√	√
Hydroxylamine reductase, putative	√	√
Clathrin light chain	√	√
Aminopeptidase	√	√
ATP-dependent RNA helicase p47, putative	√	–
DEAD/DEAH box helicase family protein	√	–
Phosphofructokinase family protein	√	–
Metalloenzyme superfamily protein	√	–
Ferredoxin 5	√	–
CK1 family protein kinase	√	–
Fibrillarin, putative	√	–
NifU-like domain containing protein	√	–
Putative NADPH-dependent butanol dehydrogenase	√	–
TvhydB protein	√	–
Multidrug resistance protein, putative	√	–
G-protein alpha subunit, putative	√	–
RNA-binding protein, putative	√	–
60S ribosomal protein L7a, putative	√	–
Methionine aminopeptidase 2	√	–
CMGC family protein kinase	√	–
Putative tryptophanyl-tRNA synthetase	√	–
Succinate--CoA ligase [ADP-forming] subunit beta, hydrogenosomal	√	–
ZPR1-related zinc finger protein	√	–
Ribosomal protein L18ae, putative	√	–
H/ACA ribonucleoprotein complex subunit	√	–
Small GTP-binding protein, putative	√	–
60S ribosomal protein L22-1, putative	√	–
GTP-binding protein YPTM2, putative	√	–
Eukaryotic release factor 3 GTPase subunit	–	√
Proteasome endopeptidase complex, putative	–	√
Ribosomal protein L5	–	√
GTP-binding nuclear protein	–	√
SNF2 family N-terminal domain containing protein	–	√
Putative actin depolymerizing factor (Fragment)	–	√
Clan ME, family M16, insulinase-like metallopeptidase	–	√
Clathrin and VPS domain-containing protein (Fragment)	–	√
Enolase 2, putative	–	√
SnoRNA binding domain containing protein	–	√
Phenylalanyl-tRNA synthetase, alpha subunit family protein	–	√
Lysine--tRNA ligase	–	√

(cont. on next page)

Table 4.8. (cont.)

Protein Name	Set 1	Set 2
Translation initiation factor SUI1 family protein	–	√
6-phosphogluconate dehydrogenase, decarboxylating	–	√
Calmodulin, putative	–	√
Adenylate kinase family protein	–	√
KE2 family protein	–	√
Phenylalanyl-tRNA synthetase, beta subunit family protein	–	√
Ornithine carbamoyltransferase family protein	–	√
Ras-related protein Rab11C, putative	–	√
Protein kinase, putative	–	√
T-complex protein 1 subunit epsilon	–	√
Chaperonin 60 (Fragment)	–	√
Pyruvate:ferredoxin oxidoreductase A	–	√

Although another study demonstrated not changed mRNA level of PFOR gene in clinical isolates, enzymatic activity was reduced in resistant isolate[62]. In this study, PFOR proprotein and PFOR A protein were differentially identified for sensitive-2 and resistant strain, respectively. Also, PFOR A-like protein was identified differentially for the sensitive-1 strain. On the other hand, PFOR BII and PFOR A proteins were found as common proteins for all strains in set1. All kind of PFOR proteins which have 90% identity and same function but their name and codes are varied.

Reactive oxygen species such as superoxide radical anion (O_2^-) and hydrogen peroxide (H_2O_2) are highly toxic so cells evolved antioxidant enzymes like SOD and, peroxiredoxins which remove H_2O_2 and reduces to water via the cysteinyl-dependent mechanism. SOD activity of resistant clinical isolates was identified higher than drug susceptible organisms[68] and expression was upregulated to tolerate metronidazole in resistant laboratory strains[66]. In this study SOD was identified as a standard protein between resistant and sensitive-2 strains instead of being a differential protein for the resistant strain.

Also, TrxR is a redox regulator and reduces peroxiredoxin with thioredoxin. Peroxiredoxin is a thioredoxin-dependent peroxidase which keeps the cell from oxidative stress of H_2O_2. In laboratory induced resistant *T. vaginalis* strains, peroxiredoxins were found as upregulated[66]. It has been shown that peroxiredoxins reduce H_2O_2, so represents cellular protection system against oxidative damage. Therefore, peroxiredoxins has an

importance for survival of the parasite in the host, and loss of its function would terminate the infection[69].

Table 4.9. List of common proteins with catalytic activity between sensitive-1 and sensitive-2 strains

Protein Name	Set 1	Set 2
Tryparedoxin peroxidase, putative	√	–
Ras-related protein Rab11C, putative	√	–
Enolase 3, putative	√	–
Cysteine synthases family protein	√	–
V-type proton ATPase subunit a	√	–
Serine palmitoyl transferase subunit, putative	√	–
Alcohol dehydrogenase	√	–
Transketolase family protein	√	–
Glucose-6-phosphate 1-dehydrogenase	√	–
Guanine nucleotide regulatory protein, putative	√	–
V-type proton ATPase subunit C	√	–
Fructose-1,6-bisphosphate aldolase	√	–
Galactokinase family protein	√	–
Lysine--tRNA ligase	√	–
Hydrogenosomal malic enzyme subunit B proprotein (Malate dehydrogenase:SUBUNIT=B)	√	–
Dynein light chain	–	√
Thioredoxin family protein	–	√
Aspartate aminotransferase	–	√
DALR anticodon binding domain containing protein	–	√
Phosphoglycerate mutase	–	√
Proteasome subunit beta type	–	√
Glutaminyl-tRNA synthetase family protein	–	√
Enolase 4, putative	–	√
Glucose-6-phosphate isomerase	–	√

It was hypothesized that TrxR serves as both an activator and a target of metronidazole and the drug disrupts redox balance by reducing the activity of TrxR. By contrast, thioredoxin reductase activity was absent due to lack of its FAD cofactor. Reduced activity of TrxR leads to diminished activity of thioredoxin required processes like redox control of chaperones, sulfate reduction[66]. However, according to another study which deals with metronidazole susceptible and clinically resistant isolate, TrxR activity

was equal for all strains[59]. In this study, TrxR was found as another common protein for all strains in only set 1.

Alcohol dehydrogenase (ADH1) was proposed as a detoxifying enzyme to reduce acetaldehyde to ethanol as a minor product and the source of acetaldehyde was probably PFOR. Comparative analysis between metronidazole susceptible and resistant clinical isolates resulted with down-regulation of ADH1 in resistant isolate[59]. The situation could be resulted from lesser production of ethanol in resistant isolate compared to susceptible one. Down-regulation of ADH1 could be an adaptation to increased oxygen levels. However, the relation between ADH1 and metronidazole resistance is not apparent; therefore, further investigations are necessary to comment on its function for resistance. In this study, ADH1 was identified as a standard protein with high score values and match numbers for resistant, sensitive-1 and, sensitive-2 strains in both sets.

Flavin reductase, NADPH oxidase, is a part of antioxidant defense mechanism and reduces molecular oxygen to hydrogen peroxide with the flavin domain. Previously, flavin reductase activity was identified as reduced or absent for both laboratory-induced metronidazole resistant cell line[66] and clinical metronidazole-resistant isolates[59]. Reduction of the enzyme activity results with an increase of molecular oxygen level, impaired oxygen scavenging and futile cycling of metronidazole so directly related to clinical metronidazole resistance. However, in this study protein with flavin reductase activity could not be detected.

An alternative pathway for activation of the drug is that NADH is generated by hydrogenosomal decarboxylating malate dehydrogenase (MDH) and used by NAD: ferredoxin oxidoreductase to reduce ferredoxin. Expression of MDH was minimal and enzyme activity was absent or very low with laboratory-induced increasing metronidazole resistance cell line. Despite minimal expression of MDH in iron-depleted *T. vaginalis* cells, they remained susceptible to metronidazole so it was understood that MDH is not essential for metronidazole activation and toxicity[66]. In accordance with the literature, MDH putative which have the same molecular function but different accession code with MDH was identified as a standard protein for all strains in this study.

Table 4.10. List of differential proteins with catalytic activity in resistant strain

Protein Name	Set 1	Set 2
Chain A, Methionine Gamma-Lyase	√	√
Glyceraldehyde-3-phosphate dehydrogenase (Fragment)	√	√
Hydroxylamine reductase, putative	√	√
Aminopeptidase	√	√
Clan SB, family S8, subtilisin-like serine peptidase	√	√
ATP-dependent RNA helicase p47, putative	√	–
DEAD/DEAH box helicase family protein	√	–
Metalloenzyme superfamily protein	√	–
Methionine aminopeptidase 2	√	–
GTP-binding protein YPTM2, putative	√	–
Ferredoxin 5	√	–
CK1 family protein kinase	√	–
Fibrillarin, putative	√	–
Putative NADPH-dependent butanol dehydrogenase	√	–
Succinate--CoA ligase [ADP-forming] subunit beta, hydrogenosomal	√	–
TvhydB protein	√	–
Multidrug resistance protein, putative	√	–
Hypothetical hydrolase-like protein, putative	√	–
G-protein alpha subunit, putative	√	–
Cysteine protease, putative	√	–
CMGC family protein kinase	√	–
Putative tryptophanyl-tRNA synthetase	√	–
Dolichyl-diphosphooligosaccharide--protein glycosyltransferase subunit 1	√	–
Aldehyde dehydrogenase	√	–
Maf protein	√	–
Proteasome endopeptidase complex, putative	–	√
GTP-binding nuclear protein	–	√
6-phosphogluconate dehydrogenase, decarboxylating	–	√
Glucose-6-phosphate isomerase(Fragment)	–	√
Phosphofructokinase family protein	–	√
Clan ME, family M16, insulinase-like metallopeptidase	–	√
Succinate--CoA ligase	–	√
Eukaryotic release factor 3 GTPase subunit	–	√
Clan CA, family C19, ubiquitin hydrolase-like cysteine peptidase	–	√
L-lactate dehydrogenase, putative	–	√
Ras-related protein Rab11C, putative	–	√
Enolase 2, putative	–	√
Pyruvate:ferredoxin oxidoreductase A	–	√
Transketolase family protein	–	√

(cont. on next page)

Table 4.10. (cont.)

Protein Name	Set 1	Set 2
Phenylalanyl-tRNA synthetase, beta subunit family protein	–	√
Ornithine carbamoyltransferase family protein	–	√
SNF2 family N-terminal domain containing protein	–	√
Protein kinase, putative	–	√
Lysine--tRNA ligase	–	√
Lipase family protein	–	√
Glutamine synthetase, catalytic domain containing protein	–	√
Adenylate kinase family protein	–	√

Table 4.11. List of common proteins with structural molecule activity between sensitive-1 and sensitive-2 strains

Protein Name	Set 1	Set 2
Ribosomal protein L8, putative	√	–
40S ribosomal protein S26	√	–
Ribosomal protein L15	√	–
40S ribosomal protein S2, putative	√	–
Ribosomal protein S10p/S20e, putative	–	√
60S ribosomal protein L22-1, putative	–	√
Ribosomal protein S14, putative	–	√
40S ribosomal protein S16, putative	–	√

Table 4.12. List of differential proteins with structural molecule activity in resistant strain

Protein Name	Set 1	Set 2
Clathrin light chain	√	√
60S ribosomal protein L22-1, putative	√	–
Ribosomal protein L18ae, putative	√	–
60S ribosomal protein L32, putative	√	–
Ubiquitin, putative	–	√
Ribosomal protein L13e, putative	–	√
PCI domain containing protein	–	√
Ribosomal protein L5	–	√
Ribosomal protein L8	–	√

CHAPTER 5

CONCLUSION

Metronidazole, a kind of nitroimidazole drugs, is preferred for treatment of trichomoniasis. However, sometimes clinical isolates from patients are estimated to have some resistance to metronidazole. Treatment failure can be reasoned from inadequate delivery of drug to vaginal area, inactivation of drug by vaginal flora, resistance to metronidazole and reinfection. However, the resistance mechanism is still unknown. Metronidazole resistance has been studied so far to suggest a relation between metronidazole resistance and redox pathways for laboratory induced resistance and clinical resistance. This study was investigated qualitative comparative proteomic analysis of metronidazole sensitive and resistant strains directly from patients. Molecular activities of identified proteins in all strains were investigated for both sets to comment about resistance effect of protein.

Proteins with binding and catalytic activities are high in number contrary to other activities according to molecular activity bar graphs. Therefore, proteins with these two activities may have a role in the resistance mechanism. Eukaryotic release factor (putative), chain A methionine gamma-lyase, hydroxylamine reductase (putative), clathrin light chain and aminopeptidase proteins with binding activity were identified for only resistant strain in Set 1 and Set 2 separately. Furthermore, chain A methionine gamma-lyase, glyceraldehyde-3-phosphate dehydrogenase (fragment), hydroxylamine reductase (putative), aminopeptidase and clan SB, family S8, subtilisin-like serine peptidase with catalytic activity were found in the resistant strain of both sets. Also, clathrin light chain which has both binding and structural molecular activity was observed in two sets of the resistant strain. All these proteins with different activities may effect on the resistance against the drug because they were determined in both sets of the resistant strain.

It has published that FeS proteins are involved in metronidazole activation. In this study, it has shown that differentially identified proteins ferredoxin and hydroxylamine reductase in resistant strain might have a significant role in metronidazole

resistance. Because, increased number of iron-sulfur proteins exists in metronidazole resistant strain.

Due to anaerobic electron transport property of these proteins, there are enough electron transfer for both activations of drug and progression of hydrogenase cycle pathways. Also, thioredoxin reductase, pyruvate: ferredoxin oxidoreductase, superoxide dismutase, alcohol dehydrogenase activities which are related with metronidazole resistance mechanism in the literature were identified as common proteins in all strains.

Differences between two sets might be the reason for the sample preparation which is the key of both mass spectrometry and proteomic studies. Because any mistake caused by the human hand may result differently. Besides, sets should examine under the same time interval and conditions, so this might be another reason for protein difference. Also, every patient has a unique immune system, and protein levels can change under any stress such as drug or infection. As a result, protein differences could observe for the sensitive strains from two different patients. The findings should be corroborated with more samples of *T. vaginalis* for further investigations.

The loss of peroxiredoxin activity would terminate the infection. If TrxR activity can be inhibited, the peroxidase may not be in the active reduced form and the infection might be ended. Although, any inhibitor has not been discovered yet, detailed study in this field can contribute to understanding the resistance mechanism. It was shown in the literature that chlorinated metronidazole is a promising alternative pathway for treating trichomoniasis due to its greater activation than metronidazole[70]. As a further study, metronidazole can be re-synthesized with an attached group which might have increased potential for drug reduction.

On the other hand, single nucleotide polymorphisms in one of the nitro-reductase genes can cause metronidazole resistance in *T. vaginalis*[65]. Also, genetic mutations are believed to cause drug resistance. All in all, drug resistance is a complex study and research with an increased number of samples can help to catch a biomarker for drug resistance.

REFERENCES

1. Falconer, I. J., J J Thomson and the Discovery of the Electron. *Physics Education* **1999,** *32*, 226-231.

2. Han, X.; Aslanian, A.; Yates, J. R., Mass spectrometry for proteomics. *Current Opinion in Chemical Biology* **2008,** *12* (5), 483-490.

3. Banerjee, S.; Mazumdar, S., Electrospray Ionization Mass Spectrometry: A Technique to Access the Information beyond the Molecular Weight of the Analyte. *International Journal of Analytical Chemistry* **2012,** *2012*, 1-40.

4. Hoffman, E. d.; Stroobant, V., *Mass Spectrometry Principles and Applicaitons*. 3 ed.; John Wiley & Sons: England, 2007.

5. Ribeiro, D., A Short Overview of the Components in Mass Spectrometry Instrumentation for Proteomics Analyses. **2013.**

6. Kebarle, P., *Reactive Intermediates: MS Investigations in Solution*. WILEY: Weinheim, Udo H. Verkerk; p 1.

7. Yamashita, M.; Fenn, J. B., Electrospray Ion Source. Another Variation on the Free-Jet Theme. *J. Phys. Chem.* **1984,** *88*, 4451-4459.

8. Gabelica, V.; Pauw, E. D., Internal energy and fragmentation of ions produced in electrospray sources. *Mass Spectrometry Reviews* **2005,** *24* (4), 566-587.

9. Lane, C. S., Mass spectrometry-based proteomics in the life sciences. *CMLS Cellular and Molecular Life Sciences* **2005,** *62* (7-8), 848-869.

10. Liebler, D. C., *Introduction to Proteomics, Tools for the New Biology* Humana Press: Totowa, New Jersey, 2002.

11. Karas, M.; Hillenkamp, F., Laser Desorption Ionization of Proteins with Molecular Masses Exceeding 10 000 Daltons. *Anal. Chem.* **1988,** *60*, 2299-2301.

12. El-Aneed, A.; Cohen, A.; Banoub, J., Mass Spectrometry, Review of the Basics: Electrospray, MALDI, and Commonly Used Mass Analyzers. *Applied Spectroscopy Reviews* **2009,** *44* (3), 210-230.

13. Jr., G. C. S.; Kelly, P. E.; Syka, J. E. P.; Reynolds, W. E.; Todd, J. F. J., Recent improvements in and analytical applications of advanced ion trap technology. *International Journal of Mass Spectrometry and Ion Process* **1984,** *60*, 85-98.

14. Neetu, K.; Ankit, G.; Ruchi, T.; Ajay, B.; Prashant, B., A Review on Mass Spectrometry Detectors. *International Research Journal of Pharmacy* **2012,** *3* (10), 33-42.

15. Horgan, R. P.; Kenny, L. C., 'Omic' technologies: genomics, transcriptomics, proteomics and metabolomics. *The Obstetrician & Gynaecologist* **2011,** *13* (3), 189-195.

16. Yates, J. R.; Ruse, C. I.; Nakorchevsky, A., Proteomics by Mass Spectrometry: Approaches, Advances, and Applications. *Annual Review of Biomedical Engineering* **2009,** *11* (1), 49-79.

17. Graves, P. R.; Haystead, T. A. J., Molecular Biologist's Guide to Proteomics. *Microbiology and Molecular Biology Reviews* **2002,** *66* (1), 39-63.

18. Chandrasekhar, K.; Dileep, A.; Lebonah, D. E.; Kumari, J. P., A Short Review on Proteomics and its Applications. *International Letters of Natural Sciences* **2014,** *17*, 77-84.

19. Lehrer, S.; Roboz, J.; Ding, H.; Zhao, S.; Diamond, E. J.; Holland, J. F.; Stone, N. N.; Droller, M. J.; Stock, R. G., Putative protein markers in the sera of men with prostatic neoplasms. *BJU International* **2003,** *92* (3), 223-225.

20. Choi, J.; Malakowsky, C. A.; Talent, J. M.; Conrad, C. C.; Grac, R. W., Identification of oxidized plasma proteins in Alzheimer's diseas. *Biochemical and Biophysical Research Communications* **2002,** *293*, 1566–1570.

21. Greber, S.; Lubec, G.; Cairns, N.; Fountoulakis, M., Decreased levels of synaptosomal associated protein 25 in the brain of patients with Down Syndrome and Alzheimer's disease. *Electrophoresis* **1999,** *20*, 928-934.

22. Berkelman, T.; Stenstedt, T., *2-D electrophoresis : principles and methods*. Amersham Biosciences: Upsalla, Sweden, 2002.

23. Boyer, R. F., *Modern Experimental Biochemistry*. 3 ed.; Addison Wesley Longman: 2000.

24. Bendixen, E., The use of proteomics in meat science. *Meat Science* **2005,** *71* (1), 138-149.

25. Shi, Y.; Xiang, R.; Horváth, C.; Wilkins, J. A., The role of liquid chromatography in proteomics. *Journal of Chromatography A* **2004,** *1053* (1-2), 27-36.

26. Fournier, M. L.; Gilmore, J. M.; Martin-Brown, S. A.; Washburn, M. P., Multidimensional Separations-Based Shotgun Proteomics. *Chem. Rev.* **2007,** *107* (8), 3654-3686.

27. Opiteck, G. J.; Lewis, K. C.; Jorgenson, J. W., Comprehensive On-Line LC/LC/MS of Proteins. *Anal. Chem.* **1997,** *69*, 1518-1524.

28. Wolters, D. A.; Washburn, M. P.; Yates, J. R., An Automated Multidimensional Protein Identification Technology for Shotgun Proteomics. *Anal. Chem.* **2001,** *73*, 5683-5690.

29. McLafferty, F. W.; Breuker, K.; Jin, M.; Han, X.; Infusini, G.; Jiang, H.; Kong, X.; Begley, T. P., Top-down MS, a powerful complement to the high capabilities of proteolysis proteomics. *FEBS Journal* **2007,** *274* (24), 6256-6268.

30. Switzar, L.; Giera, M.; Niessen, W. M. A., Protein Digestion: An Overview of the Available Techniques and Recent Developments. *Journal of Proteome Research* **2013,** *12* (3), 1067-1077.

31. McDonald, W. H.; III, J. R. Y., Shotgun proteomics and biomarker discovery. *Disease Markers* **2002,** *18*, 99-105.

32. Miller, D. M. H. C. A. M. B. D., The Effect of MS/MS Fragment Ion Mass Accuracy on Peptide Identification in Shotgun Proteomics. *Agilent Technologies Application Note*.

33. Nesvizhskii, A. I.; Aebersold, R., Interpretation of Shotgun Proteomic Data. *Molecular & Cellular Proteomics* **2005,** *4*, 1419–1440.

34. Matallana-Surget, S.; Leroy, B.; Wattiez, R., Shotgun proteomics: concept, key points and data mining. *Expert Review of Proteomics* **2014,** *7* (1), 5-7.

35. Yang, F.; Shen, Y.; Camp, D. G.; Smith, R. D., High-pH reversed-phase chromatography with fraction concatenation for 2D proteomic analysis. *Expert Review of Proteomics* **2014,** *9* (2), 129-134.

36. Wang, Y.; Yang, F.; Gritsenko, M. A.; Wang, Y.; Clauss, T.; Liu, T.; Shen, Y.; Monroe, M. E.; Lopez-Ferrer, D.; Reno, T.; Moore, R. J.; Klemke, R. L.; Camp, D. G.; Smith, R. D., Reversed-phase chromatography with multiple fraction concatenation strategy for proteome profiling of human MCF10A cells. *Proteomics* **2011,** *11* (10), 2019-2026.

37. Smith, J. B., Peptide Sequencing by Edman Degradation. *Encyclopedia of Life Sciences* **2001**.

38. Towbin, H.; Staehelin, T.; Gordon, J., Electrophoretic transfer of proteins from polyacrylamide gels to nitrocellulose sheets: procedure and some applications. *Proceedings of the National Academy of Sciences* **1979,** *76* (9), 4350-4354.

39. Elias, J. E.; Gygi, S. P., Target-Decoy Search Strategy for Mass Spectrometry-Based Proteomics. *Methods Mol Biol.* **2010,** *604*, 55-71.

40. Mascot search parameters | Protein identification software for mass spec http://www.matrixscience.com/help/search_field_help.html (accessed Oct 31, 2018).

41. Brosch, M.; Yu, L.; Hubbard, T.; Choudhary, J., Accurate and Sensitive Peptide Identification with Mascot Percolator. *Journal of Proteome Research* **2008,** *8*, 3176–3181.

42. Kulda, J., Trichomonads hydrogenosomes and drug resistance. *International Journal for Parasitology* **1999,** *29*, 199-212.

43. Upcroft, P.; Upcroft, J. A., Drug Targets and Mechanisms of Resistance in the Anaerobic Protozoa. *Clinical Microbiology Reviews* **2001,** *14* (1), 150-164.

44. Ellis, J. E.; Cole, D.; Lloyd, D., Influence of oxygen on the fermentative metabolism of metronidazole-sensitive and resistant strains of Trichomonas vaginalis. *Mol. Biochem. Parasitol.* **1992,** *56*, 79-88.

45. Hrdý I.; Müller M., Primary structure of the hydrogenosomal malic enzyme of Trichomonas vaginalis and its relationship to homologous enzymes. *J Eukaryot Microbiol* **1995,** *42*, 593-603.

46. Jenkins T.M.; Gorrell T.E.; Müller M.; Weitzman P.D., Hydrogenosomal succinate thiokinase in Tritrichomonas foetus and Trichomonas vaginalis. *Biochem Biophys Res Commun* **1991,** *179*, 892-896.

47. Lahti C.J.; Bradley P.J.; Johnson P.J., Molecular characterization of the alpha-subunit of Trichomonas vaginalis hydrogenosomal succinyl CoA synthetase. *Mol Biochem Parasitol* **1994,** *66*, 309-318.

48. Johnson P.J.; D'oliveira C.E.; Gorrell T.E.; Müller M., Molecular analysis of the hydrogenosomal ferredoxin of the anaerobic protist Trichomonas vaginalis. *Proc Natl Acad Sci USA* **1990,** *87*, 6097-6101.

49.Bui E.; Johnson P.J., Identification and characterization of [Fe]-hydrogenase in the hydrogenosome of Trichomonas vaginalis. *Mol Biochem Parasitol* **1996,** *76*, 305-310.

50. De Jesus, J. B.; Cuervo, P.; Junqueira, M.; Britto, C.; Silva-Filho, F. C.; Sabóia-Vahia, L.; González, L. J.; Barbosa Domont, G., Application of two-dimensional electrophoresis and matrix-assisted laser desorption/ionization time-of-flight mass spectrometry for proteomic analysis of the sexually transmitted parasiteTrichomonas vaginalis. *Journal of Mass Spectrometry* **2007,** *42* (11), 1463-1473.

51. Calderaro, A.; Piergianni, M.; Montecchini, S.; Buttrini, M.; Piccolo, G.; Rossi, S.; Arcangeletti, M. C.; Medici, M. C.; Chezzi, C.; De Conto, F., MALDI-TOF mass spectrometry as a potential tool for Trichomonas vaginalis identification. *BMC Infectious Diseases* **2016,** *16* (1).

52. GIDEON Informatics, I.; Berger, S., *Trichomoniasis: Global Status: 2017 edition*. GIDEON Informatics, Incorporated: 2017.

53. Arroyo, R.; González-Robles, A.; Martínez-Palomo, A.; Alderete, J. F., Signalling of Trichomonas vaginalis for amoeboid transformation and adhesin synthesis follows cytoadherence. *Molecular Microbiology* **1993,** *7* (2), 299-309.

54. De Miguel N.; Riestra A.; Johnson PJ., Reversible association of tetraspanin with Trichomonas vaginalis flagella upon adherence to host cells. *Cellular Microbiology* **2012,** *14* (12), 1797 – 1807.

55. Ardalan, S.; Craig Lee, B.; Garber, G. E., Trichomonas vaginalis: The adhesins AP51 and AP65 bind heme and hemoglobin. *Experimental Parasitology* **2009,** *121* (4), 300-306.

56. Addis, M. F.; Rappelli, P.; Delogu, G.; Carta, F.; Cappuccinelli, P.; Fiori, P. L., Cloning and Molecular Characterization of a cDNA Clone Coding for Trichomonas vaginalis Alpha-Actinin and Intracellular Localization of the Protein. *Infection and Immunity* **1998,** *66* (49), 24-31.

57. Bandea C.I.; Joseph K.; Secor E.W.; Jones L.A.; Igietseme J.I.; Sautter R.L.; Hammerschlag M.R.; Fajman N.N.; Girardet R.G.; Black C.M.; Development of PCR assays for detection of Trichomonas vaginalis in urine specimens. **2013,** *51* (1), 298-300.

58. Cudmore, S. L.; Delgaty, K. L.; Hayward-McClelland, S. F.; Petrin, D. P.; Garber, G. E., Treatment of Infections Caused by Metronidazole-Resistant Trichomonas vaginalis. *Clinical Microbiology Reviews* **2004,** *17* (4), 783-793.

59. Leitsch, D.; Drinić, M.; Kolarich, D.; Duchêne, M., Down-regulation of flavin reductase and alcohol dehydrogenase-1 (ADH1) in metronidazole-resistant isolates of Trichomonas vaginalis. *Molecular and Biochemical Parasitology* **2012,** *183* (2), 177-183.

60. Yarlett, N.; Yarlett, N. C.; LLoyd, D., Metronidazole-resistant clinical isolates of Trichomonas vaginalis have lowered oxygen affinities. *Molecular and Biochemical Parasitology* **1986,** *19*, 111-116.

61. Quon, D. V. K.; D'oliveira, C. E.; Johnson, P. J., Reduced transcription of the ferredoxin gene in metronidazoleresistant. *Proc. Nati. Acad. Sci. USA* **1992,** *89*, 4402-4406.

62. Mead, J. R.; Fernadez, M.; Romagnoli, P. A.; Secor, W. E., Use of Trichomonas vaginalis Clinical Isolates to Evaluate Correlation of Gene Expression and Metronidazole Resistance. *Journal of Parasitology* **2006,** *92* (1), 196-199.

63. Carlton, J. M., Hirt, R. P., Silva, J. C., Delcher, A. L., Schatz, M., Zhao, Q., Wortman, J. R., Bidwell, S. L., Alsmark, U. C. M., Besteiro, S., Sicheritz-Ponten, T., Noel, C. J., Dacks, J. B., Foster, P. G., Simillion, C., Van De Peer, Y., Miranda-Saavedra, D., Barton, G. J., Westrop, G. D., Müller, S., Dessi, D., Fiori, P. L., Ren, Q., Paulsen, I., Zhang, H., Bastida-Corcuera, F. D., Simoes-Barbosa, A., Brown, M. T., Hayes, R. D., Mükherjee, M., Okumura, C. Y., Schneider, R., Smith, A. J., Vanacova, S., Villalvazo, M., Haas, B. J., Pertea, M., Feldblyum, T. V., Utterback, T. R., Shu, C., Osoegava, K., De Jong, P. J., Hrdy, I., Horvathova, L., Zubacova, Z., Dolezal, P., Malik, S., Logsdon, J. M., Henze, K., Gupta, A., Wang, C. C., Dunne, R. L., Upcroft, J. A., Upcroft, P., White, O., Salzberg, S. L., Tang, P., Chiu, C., Lee, Y., Embley, T. M., Cooms, G. H., Mottram, J. C., Tachezy, J., Fraser-Lidgett, C. M.; Johnson, P. J., Draft Genome Sequence of the Sexually Transmitted Pathogen Trichomonas vaginalis. *Science* **2007,** *315* (5809), 207–212.

64. Kulda, J., Trichomonads, hydrogenosomes and drug resistance. *International Journal for Parasitology* **1999,** *29*, 199-212.

65. Paulish-Miller, T. E.; Augostini, P.; Schuyler, J. A.; Smith, W. L.; Mordechai, E.; Adelson, M. E.; Gygax, S. E.; Secor, W. E.; Hilbert, D. W., Trichomonas vaginalis Metronidazole Resistance Is Associated with Single Nucleotide Polymorphisms in the Nitroreductase Genes ntr4Tv and ntr6Tv. *Antimicrobial Agents and Chemotherapy* **2014,** *58* (5), 2938-2943.

66. Leitsch, D.; Kolarich, D.; Binder, M.; Stadlmann, J.; Altmann, F.; Duchêne, M., Trichomonas vaginalis: metronidazole and other nitroimidazole drugs are reduced by the flavin enzyme thioredoxin reductase and disrupt the cellular redox system. Implications for nitroimidazole toxicity and resistance. *Molecular Microbiology* **2009,** *72* (2), 518-536.

67. Argáez-Correa, W.; Alvarez-Sánchez, M. E.; Arana-Argáez, V. E.; Ramírez-Camacho, M. A.; Novelo-Castilla, J. S.; Coral-Martínez, T. I.; Torres-Romero, J. C., The Role of Iron Status in the Early Progression of Metronidazole Resistance in Trichomonas vaginalis Under Microaerophilic Conditions. *Journal of Eukaryotic Microbiology* **2018**.

68. Rasoloson, D.; TomkovÁ, E.; Cammack, R.; Kulda, J.; Tachezy, J., Metronidazole-resistant strains of Trichomonas vaginalis display increased susceptibility to oxygen. *Parasitology* **2001,** *123* (01).

69. Leitsch, D.; Williams, C. F.; Hrdý, I., Redox Pathways as Drug Targets in Microaerophilic Parasites. *Trends in Parasitology* **2018,** *34* (7), 576-589.

70. Chacon, M. O.; Fonseca, T. H. S.; Oliveira, S. B. V.; Alacoque, M. A.; Franco, L. L.; Tagliati, C. A.; Cassali, G. D.; Campos-Mota, G. P.; Alves, R. J.; Capettini, L. S. A.; Gomes, M. A., Chlorinated metronidazole as a promising alternative for treating trichomoniasis. *Parasitology Research* **2018,** *117* (5), 1333-1340.

APPENDIX A

ACCESSION CODES OF UNKNOWN PROTEINS

Table A.1. Accession codes of resistant strain in set 1

A2HXI6	A2DYD5	A2G7A4	A2EA48	A2FWR0	A2GCR8
A2ECS2	A2DYM3	A2DLE4	A2DDY8	A2DQX1	A2DBA3
A2EE85	A2E4V9	A2FQT4	A2DZR3	A2EXJ9	A2DPG8
A2ED49	A2E335	A2GAI6	A2DXM1	A2FC80	A2DML2
A2EZH8	A2ELE4	A2DIB4	A2FL90	A2DAJ2	A2DIW4
A2F1E4	A2DZD8	A2DQR9	A2ER01	A2EUY4	A2E4F0
A2DQN8	A2EGR6	A2EKG5	A2FH19	A2FS41	A2EUM1
A2FCY2	A2DXP9	A2F9G0	A2DTB9	A2E452	A2DGW8
A2DGK2	A2DYS0	A2ESR0	A2FC78	A2DWQ4	A2EEW4
A2FXW4	A2FLC1	A2DM81	A2FZ87	A2D7Q9	A2F0X4
A2GNU7	A2GBF6	A2EWD1	A2DVU5	A2EX43	A2EI69
A2FMM4	A2EN70	A2E974	A2E977	A2EL63	A2FZV1
A2FSA6	A2EV87	A2DAX8	A2EDJ1	A2FGB8	A2G9K7
A2DXZ9	A2FHH0	A2DZ32	A2F397	A2EHA7	A2EU15
A2DJS4	A2FMX3	A2EEF1	A2FA23	A2EUD1	

Table A.2. Accession codes of sensitive-1 strain in set 1

A2FGR2	A2DHU8	A2FIQ8	A2DXT4	A2DDV8
A2DU63	A2EER9	A2DZV7	A2DHG9	A2F9Q9
A2DXQ6	A2G053	A2FU50	A2ENF3	A2E6G3
A2F8Y3	A2DRU5	A2F6G3	A2FQF2	A2GGG5
A2F0Q8	A2F3S5	A2DXB2	A2EAZ2	A2D7R8
A2EQQ6	A2H7Y7	A2DBS0	A2FN45	A2EPZ7
A2EXK6	A2ES02	A2E4W7	A2FJF3	A2DNY3
A2ENP9	A2EDY8	A2FTT2	A2FUZ0	A2D7V7
A2D7E3	A2FV62	A2D7I1	A2FWW4	

Table A.3. Accession codes of sensitive-2 strain in set 1

A2E927	A2E6W7	A2EIU7	A2DZY8	A2E2L3	A2DD97	A2EC52
A2FFS0	A2FME4	A2FU08	A2DQC8	A2FM16	A2DMK3	A2EDP7
A2E7L9	A2EPG2	A2FNT3	A2E803	A2EQ46	A2FCI2	A2DS53
A2EEU6	A2FQS4	A2EFJ6	A2ELD1	A2FF22	A2FD06	A2DDG8
A2EKA0	A2G6T3	A2EM76	A2G6M0	A2FNA4	A2FDL4	A2FA72
A2EYU8	A2DM52	A2G048	A2E5J8	A2FGE6	A2E0X3	A2DBV8
A2FZ58	A2EQF7	A2DGG7	A2DYD6	A2EP71	A2F637	A2DFI5

Table A.4. Accession codes of resistant strain in set 2

A2EWD1	A2FA23	A2ESR0	A2EDA0	A2DA58
A2DU71	A2E5L4	A2F2D9	A2G9X1	A2FFV6
A2EI69	A2EA48	A2E9X4	A2DYS0	A2DS38
A2FQT4	A2EWU2	A2FC02	A2EKG5	A2E599
A2D7Q9	A2DIB4	A2FLC1	A2E4W7	A2FXW4
A2F381	A2D8J7	A2DY38	A2F3D1	A2DS53

Table A.5. Accession codes of sensitive-1 strain in set 2

A2DQL6	A2G9K7	A2EGI2	A2GBF6	A2EPW4	A2DLM6	A2FIG9
A2G216	A2DHU8	A2EFJ6	A2E5J8	A2GD34	A2E964	A2EDY8
A2EKH0	A2G2L3	A2F6G3	A2DDV8	A2ENP9	A2E8Q5	A2DGI6
A2EEU6	A2EER9	A2FQF2	A2DDL6	A2E452	A2D7V7	A2DES6
A2DJI0	A2DJE2	A2EP62	A2HV84	A2DDV0	A2DZ32	A2F637
A2G0V7	A2H7Y7	A2DBA3	A2EXK6	A2DD97	A2FWR0	A2EC52
A2EQQ6	A2DYM3	A2FHQ5	A2ES02	A2DAJ2	A2DHN9	A2EV95
A2E844	A2E296	A2GNU7	A2DNY3	A2FH19	A2D7E3	A2DBV8
						A2DFI5

Table A.6. Accession codes of sensitive-2 strain in set 2

A2E927	A2FIQ8	A2DIC3	A2E0A8	A2EHZ6	A2DZV7	A2G0P7
A2DZR3	A2FKH4	A2ETD9	A2DKS5	A2DYD5	A2FFR9	A2G273
A2FWP4	A2G048	A2F2B2	A2F9G0	A2DM81	A2F033	A2E977
					A2DFK9	A2G338

APPENDIX B

COMMON PROTEIN LISTS

Table B.1. Common proteins between resistant strain and sensitive-1 strain in set 1

Accession Code	Protein Name
A2DKH3	Alpha-actinin, putative
A2G5E4	Ribosomal protein L32, putative
A2DCS2	Ribosomal protein L24, putative
A2E888	LOC443698 protein, putative
Q7Z007	Rac1, putative
A2FR77	Phenylalanyl-tRNA synthetase, beta subunit family protein
Q5S7T0	Hydrogenase chain, putative
A2EIJ3	Coronin
A2DYA8	40S ribosomal protein S29, putative
A2DAL5	Putative translation initiation inhibitor
A2DGL5	Ubiquitin, putative
A2ENX0	Flavodoxin family protein
A2DTR8	Ribosomal protein S13p/S18e, putative
A2E4S6	40s ribosomal protein S5-B, putative
A2EDA1	Ribosomal protein L34e, putative
A2DYG1	UTP--glucose-1-phosphate uridylyltransferase family protein
A2FGB9	Transcription factor BTF3, putative
A2EZW4	6-phosphogluconate dehydrogenase, decarboxylating
A2FV69	TPR Domain containing protein
A2DB09	60S ribosomal protein L12, putative
A2ES64	Starch branching enzyme, putative
A2FKE2	DnaK protein (Fragment)
A2ECC6	40S ribosomal protein S4
A2DDV9	Ribosomal protein S19e, putative
A2G3F8	Ribulose-phosphate 3-epimerase
A2FT29	Triosephosphate isomerase
A2DP20	Alpha-adaptin, putative
A2FQ69	Adenylate kinase OS
A2ELL1	Heat shock factor binding protein, putative
A2FDT5	Clan ME, family M16, insulinase-like metallopeptidase
A2ELB9	Enolase family protein OS
A2F7I0	Proteasome endopeptidase complex, putative
O15565	Methionine gamma-lyase
A2GYE6	40S ribosomal protein S21

(cont. on next page)

Table B.1. (cont.)

Accession Code	Protein Name
A2FIP8	L-lactate dehydrogenase, putative
A2EDY9	Glycerol-3-phosphate dehydrogenase [NAD(+)]
A2EU51	14-3-3 protein
Q65ZG5	Succinate--CoA ligase [ADP-forming] subunit alpha, mitochondrial
A2EU62	Purine nucleoside phosphorylase, putative
Q2PCA5	A-type flavoprotein (Fragment)
A2EKU1	Phosphoenol pyruvate carboxykinase, putative
A2EES5	60S acidic ribosomal protein P0
A2DUA7	QXW lectin repeat family protein
Q8IEV2	Thioredoxin peroxidase
A2DUV8	Ribosomal protein L23, putative
A2EVS2	Alanyl-tRNA synthetase family protein
O61068	Pyrophosphate--fructose 6-phosphate 1-phosphotransferase 1

Table B.2. Common proteins between resistant strain and sensitive-2 strain in set 1

Accession Code	Protein Name
A2FJV7	Proteasome endopeptidase complex
A2D7H9	Superoxide dismutase
A2G3U5	Ornithine carbamoyltransferase family protein
O15556	Glyceraldehyde-3-phosphate dehydrogenase (Fragment)
A2EVF8	Glycyl-tRNA synthetase family protein
A2ETD8	Chaperonin 60, putative
A2G2I8	Mitotic apparatus protein, putative
A2E157	Transducin alpha subunit, putative
A2FD35	Clan CA, family C1, cathepsin L-like cysteine peptidase
A2DDE2	Ribosomal protein L29, putative
A2DJE6	Phenylalanyl-tRNA synthetase, alpha subunit family protein
A2EYC4	Probable beta-eliminating lyase
A2EIG8	Tctex-1 family protein
A2E436	Protein kinase, putative
A2DUS1	KE2 family protein
A2FE95	Vacuolar protein sorting-associated protein 35
A2DZB2	Preprotein translocase, SecY subunit, putative
A2ERB4	Seryl-tRNA synthetase family protein
A2DYJ9	PCI domain containing protein
A2E7L8	60S ribosomal protein L11, putative
A2FAF3	Fructose-1,6-bisphosphate aldolase, putative
A2FYA1	BTB/POZ domain containing protein

Table B.3. Common proteins between sensitive-1 strain and sensitive-2 strain in set 1

Accession Code	Protein Name
A2FSH1	60S ribosomal protein L18a
A2EM29	Glyceraldehyde-3-phosphate dehydrogenase
A2FAK9	40S ribosomal protein S2, putative
A2E9D9	Chaperonin subunit alpha1 CCTalpha, putative
A2DTN3	Proteasome subunit alpha type
A2FCW8	V-type proton ATPase subunit C
A2D9P8	Serine palmitoyl transferase subunit, putative
Q4G2C9	Ras-related protein Rab11C, putative
A2ECR7	Ribosomal protein L15
A2DYG8	Calcium-transporting ATPase
A2FNN7	Cytosolic repetitive antigen, putative
A2F6S0	Fructose-1,6-bisphosphate aldolase
A2DST7	40S ribosomal protein S26
A2EIU6	Aminotransferase, classes I and II family protein
A2FE30	F-actin capping protein alpha subunit, putative
A2DGQ4	Ribosomal protein S10p/S20e, putative
A2FB71	Nascent polypeptide-associated complex subunit beta
Q4G2D2	Small GTP-binding protein, putative
A2EBX0	Succinate--CoA ligase [ADP-forming] subunit beta, mitochondrial
A2FRL9	Alcohol dehydrogenase
Q5EFD8	Enolase
A2EIF8	Transketolase family protein
A2DLW4	Ribosomal protein S15a
Q9UAB6	RNA-binding protein
Q27090	Hydrogenosomal malic enzyme subunit B proprotein
A2GL34	Clathrin and VPS domain-containing protein (Fragment)
A2F719	Lysine--tRNA ligase
A2DJT4	Guanine nucleotide regulatory protein, putative
A2DA25	ATPase, AAA family protein
A2E6Z8	Ribosomal protein L8, putative
A2HG75	Tubulin beta chain (Fragment)
A2EC78	Glucose-6-phosphate 1-dehydrogenase
A2FED9	V-type proton ATPase subunit a
A2D919	40S ribosomal protein S16, putative
A2FYE4	Polyadenylate-binding protein, putative
A2GHK3	Ribosomal protein L36e, putative

Table B.4. Common proteins between resistant, sensitive-1 and sensitive-2 strain in set 1

Accession Code	Protein Name
A2EFS6	Guanine nucleotide-binding protein beta subunit, putative
A2DZL5	Fimbrin, putative
A2ED60	Dynein light chain 1, cytoplasmic, putative
A2FDS0	Ribosomal protein S3Ae, putative
A2DEH5	HEAT repeat family protein
A2EN28	Arp2/3 complex 34 kDa subunit
A2E942	Glutaminyl-Trna synthetase family protein
A2F8R8	WD repeat protein, putative
A2ENN9	Profilin
A2FCF5	Clan MG, familly M24, aminopeptidase P-like metallopeptidase
X5FVA1	Alpha actinin-3
A2F106	UDP-glucose 6-dehydrogenase
A2F422	Adenylyl cyclase-associated protein
A2F4T2	Trna binding domain containing protein
O00827	Cytoplasmic heat shock protein 70
A2DI07	60S acidic ribosomal protein P1, putative
Q27088	Pyruvate:ferredoxin oxidoreductase A
A2FDU9	Chaperonin, putative
A2DD35	Phosphatidylinositol transfer protein
A2F0H0	Transaldolase
A2EW36	Alcohol dehydrogenase, iron-containing family protein
A2DTN2	Dtdp-4-dehydrorhamnose 3,5-epimerase family protein
A2DAE1	Clan MA, family M3, oligopeptidase A-like metallopeptidase
A2DLD6	Malate dehydrogenase:SUBUNIT
A2F7C5	Endoplasmic reticulum heat shock protein 70, putative
A2DJ40	Alcohol dehydrogenase 1, putative
A2ES57	Adenosinetriphosphatase, putative
A2DJ26	Histone H2B
A2FPV1	60s Acidic ribosomal protein
A2FPE2	Oxidoreductase, aldo/keto reductase family protein
A2FSV6	Chaperonin subunit eta CCTeta, putative
A2F716	Proteasome subunit beta type
A2E926	Glycosyl transferase, group 1 family protein
A2FH22	Actin family protein
A2D8A9	60S ribosomal protein L30, putative
A2FR45	Mu adaptin, putative
A2E253	Calreticulin family protein
A2DIG6	PEP-utilizing enzyme, TIM barrel domain containing protein

(cont. on next page)

Table B.4. (cont.)

Accession Code	Protein Name
A2FW50	Phosphoglucomutase/phosphomannomutase, alpha/beta/alpha domain I family protein
A2FB52	Rho GDP exchange inhibitor, putative
A2E1L5	40S ribosomal protein S8
A2DDT1	Ribosomal protein L6, putative
A2ECR2	1,4-alpha-glucan branching enzyme IIB, chloroplast, putative
A2EIB0	Leucyl-tRNA synthetase family protein
A2EKE3	Adhesin AP65-1
A2E9H3	Pyrophosphate--fructose 6-phosphate 1-phosphotransferase 2
A2FKC7	Lactate dehydrogenase family protein
A2FRJ4	Vesicle transport v-SNARE protein
A2FRU0	Actinin, putative
A2E9K2	GTP-binding protein SAR2, putative
A2DXL4	TolA, putative
A2FKG4	Alcohol dehydrogensae, putative
A2EMI2	60S ribosomal protein L7-2, putative
A2F763	Clan CA, family C19, ubiquitin hydrolase-like cysteine peptidase
A2DLF1	Aspartate aminotransferase
Q27120	Actin (Fragment)
A2E8B1	Tubulin alpha chain
A2DSX4	Alpha-1,4 glucan phosphorylase
A2EC21	Peptidyl-prolyl cis-trans isomerase
A2FQZ6	DnaJ domain containing protein
Q95WG9	IscS/NifS-like protein
A2E3T7	Thioredoxin family protein
A2ER05	Obg-like ATPase 1
A2F259	4-alpha-glucanotransferase family protein
Q9NFT3	Coronin (Fragment)
A2DEK5	Eukaryotic translation initiation factor 5A
Q5KTX1	Cytosolic-type hsp90 (Fragment)
A2DWC1	DnaK protein
A2EAY1	EF hand family protein
A2G8P4	DALR anticodon binding domain containing protein
Q6XKE2	Hydrogenosomal oxygen reductase
A2E0V9	Actin-like protein 3, putative
A2DVT9	Adenosylhomocysteinase
A2ENN5	Histone H4
A2FF46	Isoleucyl-tRNA synthetase family protein
A2DZ49	Ribosomal protein L14, putative
A2F347	Threonyl-tRNA synthetase family protein
A2F0H5	Ribosomal protein S24e, putative

(cont. on next page)

Table B.4. (cont.)

Accession Code	Protein Name
A2F0H5	Ribosomal protein S24e, putative
A2FNW2	Actin-binding protein, putative
A2EP65	Spermatogenesis associated factor, putative
A2DPX5	Translation elongation factor 1 beta, putative
A2F8Y2	Clan MG aminopeptidase P-like metallopeptidase
A2EPR1	Heat shock protein, putative
A2EAJ8	Malic enzyme
A2DIM3	Ribosomal protein S3, putative
A2ESK3	Ras family protein
A2F855	Pyruvate:ferredoxin oxidoreductase BII
A2F520	T-complex protein 1 subunit gamma
A2DUP2	Thiol peroxidase, putative
A2FJ30	Malate dehydrogenase, putative
A2E2P0	ATP-dependent Zn protease, putative
A2FL92	TCP-1/cpn60 chaperonin family protein
Q8IEV3	Thioredoxin reductase
A2DS85	T-complex protein 1 subunit delta
A2E7V8	Glucose-6-phosphate isomerase
A2E269	Enolase 4, putative
A2FCA3	Cytidine deaminase
A2DA37	Peptidylprolyl isomerase
A2FZ56	AT5g28840/F7P1_20, putative
A2DI13	Ribonucleotide reductase, all-alpha domain containing protein
A2FMK8	Clan MH, family M20, peptidase T-like metallopeptidase
A2EXE5	Ubiquitin-conjugating enzyme family protein
A2EV07	Clathrin and VPS domain-containing protein
A2DR42	Chaperonin subunit zeta CCTzeta
A2DVD8	Ribosomal protein S19, putative
A2D8B6	C2 domain containing protein
A2FQW2	ADP-ribosylation factor 1, putative
A2EEC2	Pyruvate, phosphate dikinase
A2F093	Phosphoglycerate kinase
A2G9W5	Rab GDP dissociation inhibitor

Table B.5. Common proteins between resistant strain and sensitive-1 strain in set 2

Accession Code	Protein Name
A2FQR3	Pyruvate kinase
A2HR91	Tubulin alpha chain (Fragment)
A2EM29	Glyceraldehyde-3-phosphate dehydrogenase
A2DCS2	Ribosomal protein L24, putative
A2E888	LOC443698 protein, putative
A2F422	Adenylyl cyclase-associated protein
Q5S7T0	Hydrogenase chain, putative
A2EIJ3	Coronin
A2DAL5	Putative translation initiation inhibitor
A2DTN2	dTDP-4-dehydrorhamnose 3,5-epimerase family protein
O15556	Glyceraldehyde-3-phosphate dehydrogenase (Fragment)
A2EVF8	Glycyl-tRNA synthetase family protein
A2D8A9	60S ribosomal protein L30, putative
A2FR45	Mu adaptin, putative
A2EDA1	Ribosomal protein L34e, putative
A2DLJ8	Dynamin central region family protein
A2G7V0	ThiF family protein
A2ECR2	1,4-alpha-glucan branching enzyme IIB, chloroplast, putative
A2DDE2	Ribosomal protein L29, putative
A2EIB0	Leucyl-tRNA synthetase family protein
A2FV69	TPR Domain containing protein
A2EKE3	Adhesin AP65-1
A2FE30	F-actin capping protein alpha subunit, putative
A2DFX7	Cytosolic heat shock protein 70, putative
A2FRJ4	Vesicle transport v-SNARE protein
A2DXL4	TolA, putative
A2ECC6	40S ribosomal protein S4
A2DDV9	Ribosomal protein S19e, putative
A2DFZ5	40S ribosomal protein S5, putative
Q95WG9	IscS/NifS-like protein
A2F1M8	Malate dehydrogenase:SUBUNIT
A2DP20	Alpha-adaptin, putative
A2F259	4-alpha-glucanotransferase family protein
Q9NFT3	Coronin (Fragment)
A2DEK5	Eukaryotic translation initiation factor 5A
Q6XKE2	Hydrogenosomal oxygen reductase
A2E0V9	Actin-like protein 3, putative
Q9UAB6	RNA-binding protein
A2F0H5	Ribosomal protein S24e, putative

(cont. on next page)

Table B.5. (cont.)

Accession Code	Protein Name
O15565	Methionine gamma-lyase
A2DYJ7	Aminotransferase, classes I and II family protein
A2ESJ6	Chaperonin subunit alpha1 CCTalpha, putative
Q65ZG5	Succinate--CoA ligase [ADP-forming] subunit alpha, mitochondrial
A2ESK3	Ras family protein
A2FJ30	Malate dehydrogenase, putative
A2DLV9	Rac1, putative
A2DG97	Centrin, putative
A2FMK8	Clan MH, family M20, peptidase T-like metallopeptidase
A2EXE5	Ubiquitin-conjugating enzyme family protein
Q2PCA5	A-type flavoprotein (Fragment)
A2G9W5	Rab GDP dissociation inhibitor

Table B.6. Common proteins between resistant strain and sensitive-2 strain in set 2

Accession Code	Protein Name
A2DEH5	HEAT repeat family protein OS
A2DI07	60S acidic ribosomal protein P1, putative OS
A2F7C5	Endoplasmic reticulum heat shock protein 70, putative OS
A2FPV1	60s Acidic ribosomal protein OS
A2E926	Glycosyl transferase, group 1 family protein OS
A2FCM7	Proteasome subunit alpha type OS
A2FVK7	Succinate--CoA ligase [ADP-forming] subunit beta, mitochondrial OS
A2EFN8	Enolase family protein OS
Q8IEV4	Thioredoxin OS
A2ECU5	UTP--glucose-1-phosphate uridylyltransferase family protein OS
A2FYE4	Polyadenylate-binding protein, putative OS
A2DD10	Actin, putative OS
A2DTU7	Nucleoside diphosphate kinase OS
P49983	Adenylate kinase OS
Q6LC42	Succinate-CoA ligase subunit beta (Fragment) OS

Table B.7. Common proteins between sensitive-1 and sensitive-2 strain in set 2

Accession Code	Protein Name
A2ED60	Dynein light chain 1, cytoplasmic, putative
A2E942	Glutaminyl-tRNA synthetase family protein
X5FVA1	Alpha actinin-3
A2DD35	Phosphatidylinositol transfer protein
A2E264	Ribosomal protein S14, putative
A2FY07	Hsp90 protein
A2F716	Proteasome subunit beta type
A2FR65	Basic proline-rich protein, putative
A2FRU0	Actinin, putative
A2FKE2	DnaK protein (Fragment)
A2E9K2	GTP-binding protein SAR2, putative
A2DAA4	Ribosomal protein, putative
A2DLF1	Aspartate aminotransferase
A2E3T7	Thioredoxin family protein
A2DLW4	Ribosomal protein S15a
A2EAY1	EF hand family protein
A2G8P4	DALR anticodon binding domain containing protein
A2DVT9	Adenosylhomocysteinase
A2ENN5	Histone H4
A2EAJ8	Malic enzyme
A2EU51	14-3-3 protein
A2DKE2	60S ribosomal protein L22-1, putative
A2DS85	T-complex protein 1 subunit delta
A2E7V8	Glucose-6-phosphate isomerase
A2E269	Enolase 4, putative
A2DGX6	Cofilin/tropomyosin-type actin-binding protein
A2D8B6	C2 domain containing protein

Table B.8. Common proteins between resistant, sensitive-1 and sensitive-2 strain in set 2

Accession Code	Protein Name
A2DKH3	Alpha-actinin, putative
A2DZL5	Fimbrin, putative
A2EN28	Arp2/3 complex 34 kDa subunit
A2FAK9	40S ribosomal protein S2, putative
A2F8R8	WD repeat protein, putative
F8QX91	Elongation factor 1-alpha (Fragment)
A2FCF5	Clan MG, familly M24, aminopeptidase P-like metallopeptidase
A2F4T2	tRNA binding domain containing protein
A2EW36	Alcohol dehydrogenase, iron-containing family protein
A2DJ40	Alcohol dehydrogenase 1, putative
A2ES57	Adenosinetriphosphatase, putative
A2DJ26	Histone H2B
A2ENX0	Flavodoxin family protein
A2FH22	Actin family protein
A2E253	Calreticulin family protein
A2FD35	Clan CA, family C1, cathepsin L-like cysteine peptidase
A2FW50	Phosphoglucomutase/phosphomannomutase, alpha/beta/alpha domain I family protein
A2FB52	Rho GDP exchange inhibitor, putative
A2E1L5	40S ribosomal protein S8
A2FKC7	Lactate dehydrogenase family protein
A2FKG4	Alcohol dehydrogensae, putative
A2EMI2	60S ribosomal protein L7-2, putative
Q27120	Actin (Fragment)
A2DSX4	Alpha-1,4 glucan phosphorylase
A2EC21	Peptidyl-prolyl cis-trans isomerase
A2FQZ6	DnaJ domain containing protein
A2FT29	Triosephosphate isomerase
A2FF46	Isoleucyl-tRNA synthetase family protein
A2DZ49	Ribosomal protein L14, putative
A2F347	Threonyl-tRNA synthetase family protein
A2FNW2	Actin-binding protein, putative
A2EP65	Spermatogenesis associated factor, putative
A2DPX5	Translation elongation factor 1 beta, putative
A2EPR1	Heat shock protein, putative
A2EKU1	Phosphoenol pyruvate carboxykinase, putative
A2DIM3	Ribosomal protein S3, putative
A2EES5	60S acidic ribosomal protein P0
A2F855	Pyruvate:ferredoxin oxidoreductase BII

(cont. on next page)

Table B.8. (cont.)

Accession Code	Protein Name
A2F520	T-complex protein 1 subunit gamma
A2DUP2	Thiol peroxidase, putative
A2E2P0	ATP-dependent Zn protease, putative
A2FL92	TCP-1/cpn60 chaperonin family protein
Q8IEV3	Thioredoxin reductase
A2FZ56	AT5g28840/F7P1_20, putative OS
A2DI13	Ribonucleotide reductase, all-alpha domain containing protein
A2EV07	Clathrin and VPS domain-containing protein
A2EEC2	Pyruvate, phosphate dikinase
A2F093	Phosphoglycerate kinase
O61068	Pyrophosphate--fructose 6-phosphate 1-phosphotransferase 1