Dudu Boyvat

A Master's Thesis

AGU 2022

IDENTIFICATION OF SURFACE PROTEOME OF B CELL ACUTE LYMPHOBLASTIC LEUKEMIA CELL LINE

A THESIS
SUBMITTED TO THE DEPARTMENT OF BIOENGINEERING
AND THE GRADUATE SCHOOL OF ENGINEERING AND SCIENCE
OF ABDULLAH GUL UNIVERSITY
IN PARTIAL FULFILLMENT OF THE REQUIREMENTS
FOR THE DEGREE OF
MASTER OF SCIENCE

By
Dudu Boyvat
January 2022

IDENTIFICATION OF SURFACE PROTEOME OF B CELL ACUTE LYMPHOBLASTIC LEUKEMIA CELL LINE

A THESIS
SUBMITTED TO THE DEPARTMENT OF BIOENGINEERING
AND THE GRADUATE SCHOOL OF ENGINEERING AND SCIENCE OF
ABDULLAH GUL UNIVERSITY
IN PARTIAL FULFILLMENT OF THE REQUIREMENTS
FOR THE DEGREE OF
MASTER OF SCIENCE

By
Dudu Boyvat
January 2022

SCIENTIFIC ETHICS COMPLIANCE

I hereby declare that all information in this document has been obtained in accordance with academic rules and ethical conduct. I also declare that, as required by these rules and conduct, I have fully cited and referenced all materials and results that are not original to this work.

Name-Surname: Dudu Boyvat
Signature :

REGULATORY COMPLIANCE

M.Sc. thesis titled Identification of Surface Proteome of B cell Acute Lymphoblastic Leukemia Cell Line has been prepared in accordance with the Thesis Writing Guidelines of the Abdullah Gül University, Graduate School of Engineering & Science.

Prepared By	Advisor
Dudu Boyvat	Assist. Prof. Dr. Şerife Ayaz-Güner
Signature	Signature

Head of the Bioengineering Program

Prof. Dr. Sevil Dinçer İşoğlu

Signature

ACCEPTANCE AND APPROVAL

M.Sc. thesis titled Identification of Surface Proteome of B cell Acute Lymphoblastic Leulemia Cell Line and prepared by Dudu Boyvat has been accepted by the jury in the Bioengineering Graduate Program at Abdullah Gül University, Graduate School of Engineering & Science.

17 /01 / 2022

JURY:

Advisor : Assist. Prof. Dr. Şerife Ayaz Güner…….....

Member : Prof. Dr Servet Özcan…….....

Member : Assist. Prof. Dr. Emel Başak Gencer Akçok…….

APPROVAL:

The acceptance of this M.Sc. thesis has been approved by the decision of the Abdullah Gül University, Graduate School of Engineering & Science, Executive Board dated ….. /….. / …….. and numbered ……..……. .

/02 /2022

Graduate School Dean
Prof. Dr. İrfan ALAN

Acknowledgement

Firstly, I would like to express my appreciation to my advisor Assist. Prof. Dr. Şerife Ayaz Güner. I could not have completed this study without her support, encouragement, patience, understanding, and kindness. She did every possible help in every condition and supported me during every moment of my thesis studies.

I would also like to express my thanks to my jury committee members Assist. Prof. Dr. Emel Başak Gencer Akçok and Prof. Dr. Servet Özcan for their valuable suggestions and comments.

I would like to express my sincere thanks to Dr. Mustafa Burak Acar and Özcan lab members for their support and kindness.

I would like to thank AGÜ-Proteomik lab members for their help, cooperation, and support, and I express my appreciation to Nihan Aktaş Pepe, Merve Şansaçar, Büşra Acar, and Helin Sağır for sharing experiences, supportive friendship, and countless coffee and chocolate meetings.

Finally, I would like to thank my siblings, Nurcan Boyvat and Furkan Boyvat for their support and motivation, above all, I would like to express my grateful thanks to my parents Elif Boyvat and Kazım Boyvat for their endless love, support, encouragement, and belief.

Dudu Boyvat

ABSTRACT

IDENTIFICATION OF SURFACE PROTEOME OF B CELL ACUTE LYMPHOBLASTIC LEUKEMIA CELL LINE

Dudu Boyvat
MSc in Bioengineering
Advisor: Assist. Prof. Şerife Ayaz-Güner

January 2022

B-cell acute lymphoblastic leukemia is characterized by over and uncontrolled expression of B lymphocytes. B-ALL may occur as a result of aberrant cytosolic signal transduction and molecular abnormalities such as gene mutations, abnormal protein interactions, and an un-arrested cell cycle. Due to these abnormalities, surface proteins that compromised one-third of the proteome show different expressions compared to the healthy cells. These differences are currently in use for diagnostic and treatment approaches. Here, we aimed to isolate and identify the surface proteins of the CCRF-SB cell line to identify new, additional possible target antigens with the mass spectrometry-based proteomics approach using two different surface protein isolation strategies. The surface proteins of CCRF-SB cells were isolated with the surface biotinylation method and N-linked glycoprotein enrichment methods. With the biotinylation method, we isolated 782 proteins with 1% FDR. Gene Ontology Cellular Compartment analysis showed that 467 of these isolated proteins are annotated as 'Membrane'. 263 of those proteins are annotated as 'Extracellular Space'. These isolated cell surface proteins include HLA protein complexes and well-known CD19 surface markers. With the N-linked glycosylation enrichment method 229 protein identified with 1% FDR rate. Gene Ontology Cellular Compartment analysis showed that 155 of these isolated proteins are annotated as 'Membrane', 132 of those proteins are annotated as 'Extracellular Space'. Both methods identified different proteins from each other. This result showed that to map the surfaceome of CCRF-SB cell line, it is required to combine these two enrichment methods.

Keywords: B-ALL, Proteomics, Surfaceome, Biomarker

ÖZET

B HÜCRELİ AKUT LENFOBLASTİK LÖSEMİ HÜCRE HATTINDA YÜZEY PROTEOMUNUN BELİRLENMESİ

Dudu Boyvat
Biyomühendislik Anabilim Dalı Yüksek Lisans
Tez Yöneticisi: Dr. Öğr. Üyesi Şerife Ayaz Güner
Ocak 2022

B hücreli akut lenfoblastik lösemi, B lenfositlerinin aşırı ve kontrolsüz ifadesi ile karakterizedir. B-ALL, anormal sitozolik sinyal iletimi ve gen mutasyonları, anormal protein etkileşimleri ve durdurulmamış hücre döngüsü gibi moleküler anormalliklerin bir sonucu olarak ortaya çıkabilir. Bu anormallikler nedeniyle, proteomun üçte birini oluşturan yüzey proteinleri, sağlıklı hücrelere kıyasla farklı ifadeler gösterir. Bu farklılıklar günümüzde tanı ve tedavi yaklaşımlarında kullanılmaktadır. Bu çalışmada, iki farklı yüzey protein izolasyon stratejisinin karşılaştırılması ile kütle spektrometrisi tabanlı proteomik yaklaşımı ile yeni, ek olası hedef antijenleri belirlemek için CCRF-SB hücre hattının yüzey proteinlerini izole etmeyi ve tanımlamayı amaçladık. CCRF-SB hücrelerinin yüzey proteinleri, biyotinilasyon yöntemi ve N-bağlı glikoprotein zenginleştirme yöntemleri ile izole edildi. Biyotinilasyon yöntemi ile % 1 FDR oranı ile 782 protein izole ettik. Gene Ontology Cellular Component analizi, bu izole edilmiş proteinlerin 467'sinin 'Membran' ile ilişkili, 263'ünün "Hücre Dışı Boşluk" ile ilişkili olarak tanımlamıştır. Bu izole edilmiş hücre yüzeyi proteinlerinin, HLA protein komplekslerini ve iyi bilinen CD19 yüzey işaretleyicilerini içerdiği gösterilmiştir. N-bağlı glikozillenmiş protein zenginleştirme yöntemi ile %1 FDR oranı ile tanımlanan 229 protein Gene Ontology Cellular Component 155'inin "Membran" olarak, bu proteinlerin 132'sinin "Hücre Dışı Boşluk" ile ilişkili olarak açıklandığını gösterdi. Her iki yöntem de birbirinden farklı proteinleri tanımlamıştır. Bu sonuç, hücre yüzeyini proteomunu haritalamak için bu iki zenginleştirme yöntemini birleştirmenin gerekli olduğunu gösterdi.

Anahtar kelimeler: B-ALL, Proteomiks, Surfaceome, Biyobelirteç

TABLE OF CONTENT

LIST OF FIGURES

LIST OF TABLES

LIST OF ABBREVIATIONS

ACN	Acetonitrile
ALL	Acute Lymphoblastic Leukemia
AML	Acute Myeloid Leukemia
B-ALL	B-Cell Acute Lymphoblastic Leukemia
BSA	Bovine Serum Albumin
C18	Octadecylsilane
CD	Cluster Of Differentiation
CE	Collision Energy
CES	Collision Energy Spread
CML	Chronic Myeloid Leukemia
CID	Collision Induced Dissociation
DDA	Data Dependent Acquisition
DIA	Data Independent Acquisition
DP	Declustring Potential
DTT	Dithiothreitol
ESI	Electrospray Ionization
ETD	Electron Transfer Dissociation
FBS	Fetal Bovine Serum
FTICR	Fourier Transform Ion Cyclotron Resonance
FT-MS	Fourier Transform Ion Cyclotron
HCD	Higher-Energy Collision Dissociation
HSC	Hematopoietic Stem Cells
Ig μ	Immunoglobulin μ
IHT	Interface Heater Temperature
ISVF	Ion Source Voltage Frequency
LC	Liquid Chromatography
MALDI	Matrix-Assisted Laser Desorption Ionization
PBS	Phosphate Buffered Saline
PTM	Post Translational Modification

Q-TOF	Quadrupole Time of Flight
SDS	Sodium Dodecyl Sulfate
TBS	Tris Buffered Saline
TCEP	Tris(2-carboxyethyl) phosphine
TM	Transmembrane
TOF	Time of Flight
VEGF	Vascular Endothelial Growth Factor
WHO	World Health Organization

To my family…

Chapter 1

Introduction

Genomic and transcriptomic studies have brought different perspectives to diseases at the DNA and RNA level. However, these pieces of information often lack the phenotypic diversity application of disease and processes in the cell [1]. Proteomic studies have contributed to the understanding of these phenotypic diversities. Proteins have a crucial role in almost every cellular process in healthy or diseased cells with their ability to be a receptor for intercellular or intracellular signals, enzymatic processes, ability to be a substrate, and their role in many regulatory processes [2].

Acute lymphoblastic leukemia (ALL) is one of the common worldwide fatal diseases [3]. Especially in childhood cancers, ALL is compromising 80% of the cases [4]. Because of their aberrant expression and accessible features current approaches, for the diagnosis and the treatment of the ALL, are using the surface proteins as a target [5-8]. However, cancer cells have the ability to resist the treatment, hide and change their receptors, which negatively affect the surface marker-based diagnostic approaches. Previous studies showed that additional biomarkers are necessary for surface marker-based approaches [9, 10].

1.1 Acute Lymphoblastic Leukemia

Acute lymphoblastic leukemia is the most common childhood malignancy and is characterized by means of malignant transformation and proliferation of T or B lymphoid progenitor cells. Hallmarks of the malignant cell are mutations, aneuploidies, and translocations in the genes that regulate cell growth, division, differentiation, and other cellular processes [11]. ALL phenotype is represented with bone marrow, peripheral blood, or extramedullary part accumulation of malignant and immature cells. For diagnosing ALL; immunotyping, morphologic analysis, the cytogenetic analysis are commonly used [12]. The most frequent symptoms are anemia, leukopenia, thrombocytopenia, fever, easy bleeding, and infection [4].

ALL, mainly divided into 2 major categories: B lineage and T lineage malignancies. Classification is achieved based on immunologic, cytogenetic, and chromosomal features [13,14]. Both types include chromosomal rearrangements and genetic alteration. In addition, collaborative genetic mutations are required for a complete leukemia phenotype [15]. Classification of ALL started with FAB group classification but it is replaced by the World Health Organization (WHO) classification [12-14] (Figure 1.1). Cytogenetic and molecular features show that these 2 major groups also have sub-classes (Figure 1.1).

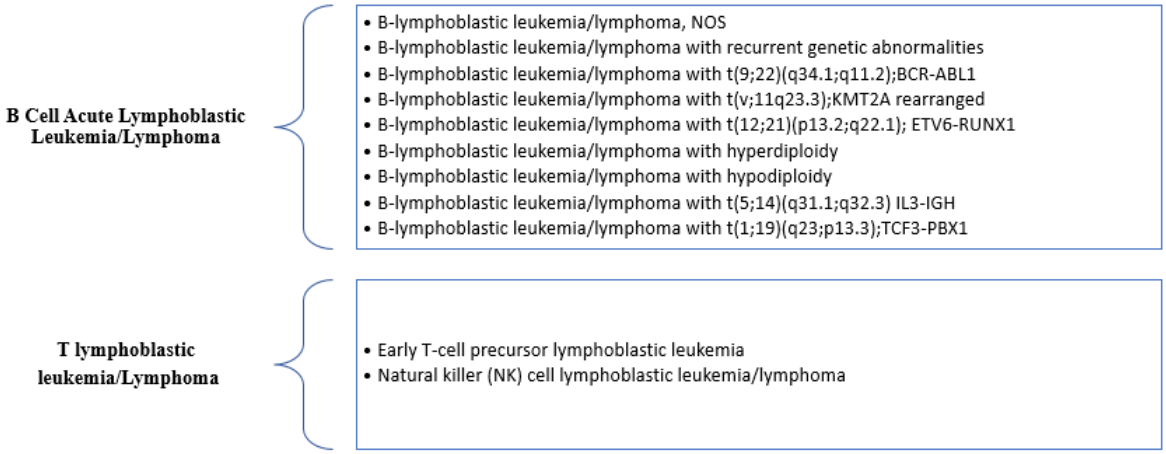


Figure 1.1 Classification of ALL. Adapted from [16]

B-ALL (B-cell Acute Lymphoblastic Leukemia) is a neoplasm derived from B-cell progenitor. B-ALL is especially observed in children and adolescents almost 4 times more common than T-ALL (T-cell Acute Lymphoblastic Leukemia). It can be classified under different titles. For instance, classification can be achieved depending on the genetic abnormalities (Table 1.1) or immunophenotypes (Table 1.2). These classifications are highly important for following a treatment path and these sub-classes represent the prognosis of the neoplasm (Table 1.1). For example, hyper diploidy can be considered as a good prognosis while hypodiploidy is a sign of poor prognosis. Additionally, modulation of the cluster of differentiation (CD) 34 and CD38 expression can be an example of a favorable prognosis [17].

T-ALL covers ~15% of the cases and like B-ALL, includes gene rearrangements, chromosomal deletions, and gene deletions [18]. Chromosomal abnormalities also lead to fusion genes such as *PICALM-MLLT10, NUP214-ABL1*. To give an example to gene mutations, *NOTCH1* gene mutation is a common abnormality and is observed in 60% of

the cases [19]. Older diagnosis age has a low influence on the T-ALL patient by contrast with B-ALL [20].

Table1.1 Genetic abnormalities in ALL and effect on prognosis.[13, 14]

Genetic Abnormality	Children		Adult	
	Prognosis	Frequency	Prognosis	Frequency
Hyperdiploidy	Good	23-30%	Good	7-8%
Hypodiploidy	Poor	6%	Poor	7-8%
t(12;21) (p13;q22) *(TEL-AML1)*	Good	22-26%	Good	0-4%
t(9;22) (q34;q11.2)/*BCR-ABL1*	Intermediate	1-3%	Intermediate	11-29%
MLL (11q23) rearrangements	Poor	1-2%	Poor	4-9%
t(1;19) (q23;p13.3)	Intermediate	1-6%	Intermediate	1-3%

Immunologic classification is important for the diagnosis of the sub-class of the ALL and also specified immune characteristics can lead to the antibody treatment due to the identification of the target antigens. The identified antigens can give information about the characteristics of diagnosed ALL such as relapse risk (Table 1.2) [21].

Table 1.2 ALL subgroups and characteristics. Adapted from [21].

Subgroup / Differentiation Marker	Characteristics (Relapse and localization)
-B lineage	
Common Precursor B-ALL/ CD10+, no cytoplasmic Ig	>90% Bone marrow localization >90% Bone marrow localization relapse: up to 5-7 years >10% CNS and extramedullary relapse: up to 1-1,5 years
Pro-B-ALL / CD10-, no cytoplasmic Ig	
Pre-B-ALL / intracytoplasmic IgM+	
Mature B-ALL / intracytoplasmic or surface κ or λ	
-T lineage	
Early-T-ALL/ CD5±, CD2- , surface CD3- , CD1a-	10% CNS 6% extramedullary localization relapse: up to 3–4 years
Thymic-T-ALL/ CD2+, CD5+, CD1a+, surface CD3±	
Mature T-ALL/ CD2+, CD5+, sCD3+, CD1a-	

1.2 B-Cell Acute Lymphoblastic Leukemia

ALL is the most common leukemia type for the childhood patient and B-ALL is the most common type of ALL with an 80% rate [4]. B-ALL is characterized by overexpression of B malignant cells. Malignancy starts in pluripotent hematopoietic stem cells (HSC). In a normal process, the differentiation of pluripotent HSCs are highly controlled by cytosolic signals, transcription factors, and selection. B-ALL occurs as a result of cytosolic signal transduction and molecular abnormalities such as gene mutations, abnormal protein interactions, and an un-arrested cell cycle [4]. Abnormalities inhibit lymphoid differentiation and cause the proliferation of damaged cells. Also, abnormalities cause the survival of the malignant cells [22]. Different genes are involved in ALL pathogenesis. For example, transcriptional regulator genes (PAX5, IKZF1), tumor suppressor genes (CDKN2A, RB1, TP53), lymphoid signaling genes (BTLA, CD200 TOX) [23, 24]. B-ALL may be subclassified based on immunologic features,

a) Common precursor B-ALL,

b) pro-B-ALL,

c) pre-B-ALL,

d) mature B-ALL (Table 3).

In the normal process, HSCs follow three stages, pro-B, pre-B, and immature B cells. Depending on the stages, cells express different surface markers. Early pro B cells express CD19, cytoplasmic CD22, CD34, and TdT might be expressed and CD20 is negative at this stage. Pro B cells express; CD19, CD22, CD79a, CD45, TdT, CD20 might be expressed at this stage. For pre-B cells, almost all the surface markers are expressed and maturation of the cell is characterized with immunoglobulin μ (Igμ) presence [23].

Current treatment guidelines follow different paths depending on the patient's age and a number of specific genetic patterns. Older patients, especially with unfavorable genetic conditions for instance ETV6, IKZF1 deletions have a higher relapse risk. Some patients with mutations in NT5C2 have a lower response to chemotherapy [23, 25]. In addition to age, a higher leukocyte number and Ph chromosome positivity are prognostic factors and are a representative feature for poor prognosis [23, 26].

1.3 Cell Surface Proteins

The cell membrane is a highly functioning surface that includes receptors, transporters, carriers, and channels [2]. Cell surface proteins (membrane proteins) are fundamental molecules to interact with the other cells and the environment. These proteins have the ability to orchestrate communication with the surroundings which means they are involved in cell fate decisions through the exchange of information, ion-metabolite transportation, cellular proliferation, and cell-cell interactions (Figure 1.2) [27].

Cell surface proteins, which comprise approximately one-third of the proteome besides playing an important role in many essential cellular functions, their aberrant expression, and changed localization on the cell surface are linked to the molecular pathology of many diseases [28]. These proteins are exposed to extracellular environments, and they are easily accessible for therapeutic purposes, thus they can account for a significant proportion of drug targets. Besides being a direct drug target,

because of their altered number and configuration, those proteins could also serve as diagnostic biomarkers of disease progression [29].

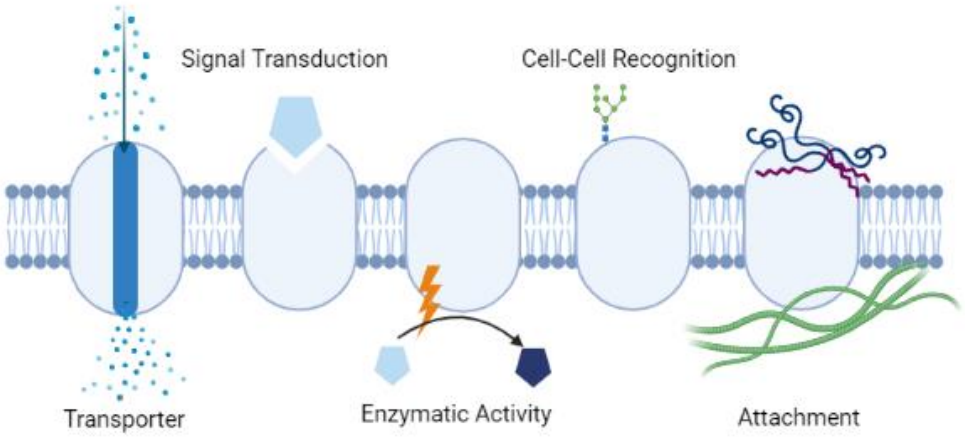


Figure 1.2 Cell surface proteins roles.

Membrane proteins can be classified into two main categories, peripheral and integral proteins. Peripheral proteins do not pass through the membrane and are usually attached to the membrane through indirect or direct interactions [30]. Integral proteins also known as transmembrane domain (TM) include different types. All the types have one or multiple regions/regions embedded in the phospholipid bilayer [31, 32].

Integral Proteins

Integral proteins can be classified under 3 main categories based on their interactions and localization on the membrane. These categories include Type I, Type II, and Multi-pass proteins (Figure1.3). Basically, Type I and Type II classes contain single-pass protein across the membrane. Type I proteins are oriented extracellular N-terminus and cytoplasmic C-terminus. Type II proteins are oriented oppositely, with an extracellular C-terminus and cytoplasmic N-terminus. Multi-pass proteins localized with cytoplasmic N and C-terminus part (Figure1.3) [33].

Localization of the proteins is crucially important for their role and function. Integral proteins are localized on the membrane and contain one or more TM, maybe a lipid anchor and extracellular regions. These regions give the cell potential to receive and respond to external stimuli. Due to their receiving capability cells may alter their behavior as a response to these signs. These cell surface receptors may include enzyme and G

protein-linked proteins, ligand and ion channel gated proteins, and cell adhesion-linked proteins [34].

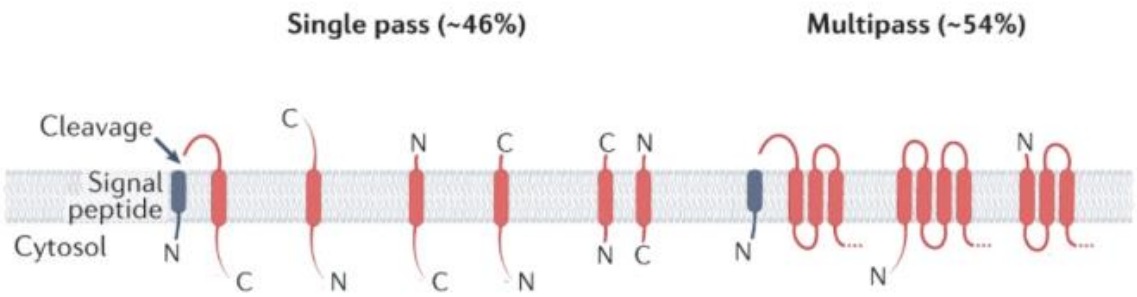


Figure 1.3 Cell surface protein types based on their localization. Adapted from [35].

Peripheral Proteins

Peripheral proteins are not able to interact with the hydrophobic site of the phospholipid bilayer. They interact with the membrane indirectly with integral proteins or direct bounds with the lipid bilayer's head groups. Peripheral proteins are also called extrinsic proteins and they are mostly found in the cytosolic part of the cell. In addition to that, some of the peripheral proteins are able to localize on the outer surface of the cell membrane (Figure1.4) [30, 32].

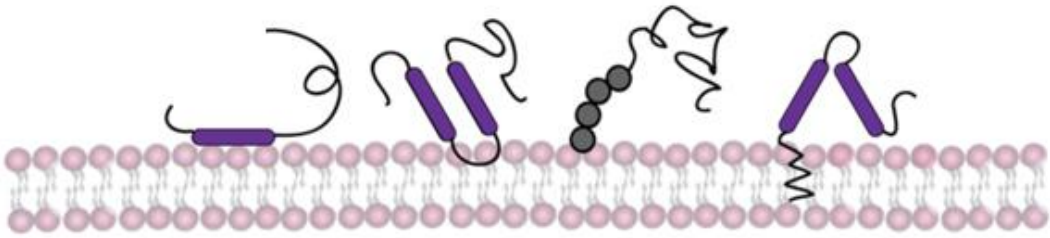


Figure 1.4 Illustration of peripheral protein [5].

1.3.1 Cell Surface Proteins in Cancer

Cell surface proteins are known for their important role in cellular proliferation, signal transduction, and other crucial mechanisms for the cellular life cycle in healthy or diseased cells. Healthy cells have the ability to control their proliferation, growth, and their lifespan, through controlling their surface proteins such as receptors and transporters [36]. In cancer cells, cell surface protein plays an important role in many mechanisms such as immortalization [6]. By arranging its surfaces in different ways, a cancer cell can change its own fate. One of the ways to support their own growth, they can produce and release growth-promotive molecules (growth-factor ligands). Production of growth-factor ligands such as Vascular Endothelial Growth Factor (VEGF) and IL1α play a role as autocrine signal molecules and help the cancer cells with their continuous proliferation [7, 37, 38]. Another way to support their growth, cells can alter their receptor amount [6]. Besides the production of growth-receptor ligands, cancer cells may increase their receptor number to receive higher amounts of growth factor ligands. In addition to increasing surface protein number, cells may change the structural configuration of these proteins to ensure constitutive proliferation [8].

Cancer cells also have a different metabolism than healthy cells. It is widely known that cancer cells choose to use glycolysis rather than oxidative phosphorylation [39, 40]. To support their constitutive proliferation and their abnormal life cycle they also need extreme energy levels. To fulfill this need, cancer cells are able to re-arrange their transporter amount.

1.3.2 Cell Surface Proteins in Diagnosis and Treatment

Allelic variants, alternative splicing of RNA transcripts, and numerous co- and post-translational modifications can all contribute to complexity at the protein level [41]. These events produce different protein molecules that influence a wide range of biological processes, including cell signaling within and between cells, gene regulation, and protein complex activation. Even a single gene can create many proteoforms due to multistage regulation. Proteoforms differ from one another in terms of subcellular localization, binding partners, structure, and kinetics. In addition to all these, cancer cells are known to have unstable genomes [6]. This genomic instability leads to overexpression of cancer supportive proteins and their proteoforms. While cancer cells are recognized for their

ability to evade the immune system and treatment, this proteome alteration allows them to be targeted for diagnosing and treatment [6].

Cell surface-based diagnostic and therapeutic techniques are increasingly favorable because of their facile accessibility and high specificity targeting abilities. In diagnostic, surface markers can be detected by different approaches for different purposes. For instance, surface biomarkers can be targeted with an isotopic label to image and detect cancerous tissues. This imaging may help to understand the cancer is benign or malign, the stages of cancer, metastatic condition, and invasiveness which all are important for following and deciding a treatment path for the patients. Another use of the biomarkers in cancer, detection and diagnosing of drug resistance.

In clinics, surface biomarker-based therapy approaches are commonly used. The procedure is preferred because of its specificity and efficiency. Surface biomarkers can function in a variety of ways. Their mechanism of function may show differences depending on the usage purpose. For instance, they may lead the targeted cell to death by apoptosis or may inhibit their proliferation [42]. Also, they can be combined for dual or even triple usage (Table 1.3). As examples of the usage of surface-biomarker based therapies,

a) Monoclonal Antibodies: Rituximab, it is targeting the CD20 surface marker which is known overexpressed in leukemia. It is shown that Rituximab has an effect on improving overall survival to 70% from the 38 % survival rate and decreasing the relapse incidence. In time, with changing a binding of the Rituximab a new monoclonal antibody called Ofatumumab generated and increased the overall survival rate by 82% [42].

b) Dual Specific Antibodies: AMG 330, it is targeting the CD3/CD33 surface markers with dual specificity. It is known that AMG 330 treatment increases the release of immune-stimuli factors [42].

c) Chimeric Antigen Receptors (CARs): The basic principle of this method is generating specific cells with specifically engineered receptors and targeting the surface marker of the cancer cells. For this treatment method, patients' own cells were used and engineered. With some drawbacks such as toxicity, these therapies are improved in time with the addition of co-stimulatory factors. CD19, CD20, CD123, and CLL/1 are some of the surface markers that are used for CAR generation [42].

Table 1.3 Surface-marker based therapy examples [42-47].

Target Molecule	Clinical Name	Disease	Results	Result Rate
CD20	Rituximab	ALL	Increased OS	70% vs 30%
			CR	75% vs 47%
CD52	Alemtuzumab	ATL	Antitumor Activity	71%
CD3/CD19	Blinatumomab	Adult ALL	CR	34% vs 16%
CD19	CD19 CART	B-ALL	CR	70-90%
CD19/CD20	LV20.19 CART	B Cell Non-Hodgkin Lymphoma	CR	70%

OS: Overall Survival CR: Complete Remission

1.3.3 Protein Glycosylation and Cancer

Glycosylation is one of the most common co- and post-translational modifications[48]. In this modification, oligosaccharide chains are bound to the proteins covalently. These oligosaccharide chains are composed of ten monosaccharide units which are xylose (Xyl), fucose (Fuc), galactose (Gal), glucose (Glc), N-acetylgalactosamine (GalNAc), N-acetylglucosamine (GlcNAc), glucuronic acid (GLCA), iduronic acid (IdoA), mannose (Man), sialic acid (SA) [49]. These groups are able to bind protein in different ways and a variety of combinations [49]. With the different combinations and different bindings, this modification constitutes diversities at a huge number.

Human cells include different types of glycosylation (Figure 1.5). The most known and common types of glycosylation are the N- and O- linked glycosylation. It is necessary to understand the biosynthesis of this alteration in order to appreciate its significance and ability to generate varieties. To summarize the biosynthesis of N-linked glycoprotein, eukaryotic glycans are first generated on the endoplasmic reticulum's cytoplasmic surface. The transfer of GlcNAc-P to Dolichol-Phosphate (Dol-P) produces Dolichol pyrophosphate N-acetylglucosamine (Dol-P-P-GlcNAc). The precursor molecule then makes its way into the ER lumen. The inclusion of mannose and glucose residues expands this precursor molecule. After this, the '14-sugar' group, Dol-P-carbohydrate, is transferred to the polypeptide chain's Asn-X-Ser/Thr (X: amino acid except for proline) region. At this point, some of the glucose residues are removed and after quality control, suitable proteins are transferred to the Golgi Apparatus. In the Golgi, some additional residue additions occur. For a different type of glycosylation synthesis,

different glycans may attach to the different sites for instance O-linked glycosylation occurs on the hydroxyl group's Ser and Thr residues [49-51].

Glycosylation changes can be evaluated under two titles. The reasons and the results of the glycosylation changes.

a) Reasons: Altered glycosylated protein synthesis can occur due to the metabolic changes of the cells (Warburg effect). As mentioned above, cancer cells prefer to use aerobic glycolysis and this glucose usage upregulates the production of UDP-GlcNAc through the upregulation of the hexosamine biosynthesis pathway [49, 52, 53]. Altered expression of the enzymes that play a role in glycoprotein synthesis and modified transport of the precursor molecule may play a role in the altered glycoprotein production [49]. Chaperones are the molecules that are responsible for protein folding. Dysregulation of these molecules may play role in altered glycoprotein synthesis [49, 54]. After the synthesis in the ER, the molecule is transferred into the Golgi like it is described above. In the Golgi apparatus, if the related enzymes are localized differently it will also affect the production of the glycoprotein [49, 55].

b) Results: Due to the Warburg effect, cancer cells start using glucose, and as described above its related to glycan motif changes. These changes also lead to continuing aerobic glycolysis [49]. Glycan motif changes on the glycoproteins modulate the neoplastic transformation, cellular growth, and differentiation [56, 57]. Also, altered expression of glycoproteins affects the invasiveness and migration ability of the cells due to alteration on the cellular adhesion proteins [58]. All these changes lead to heterogeneity and support the cancer cell progression [59].

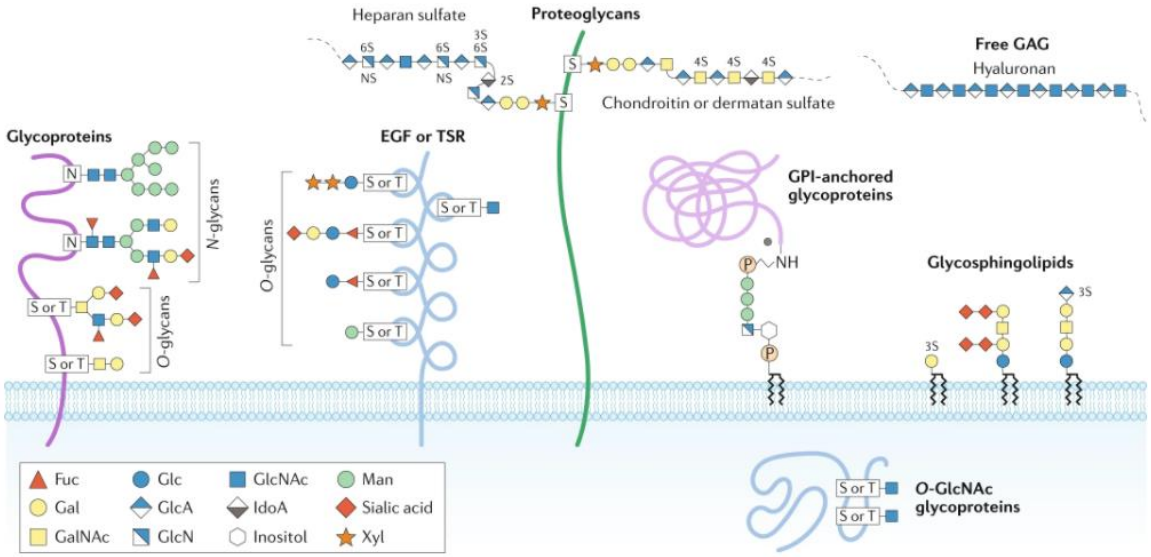


Figure 1.5 Glycosylation types for human cells [50].

1.4 Proteomics

Different and more extensive investigations in the field of molecular biology have been made possible by technological advancements and the advancement of instrumental analysis methods. One of the major milestones in this subject is the completion of the human genome project. Following the discovery of a way to examine the human genome, researchers moved their attention to transcripts derived from the genome and proteins, which are the functional states of genes. The age of "-omics" in molecular biology began with the start of RNA and protein study [60].

Proteomics is the study of proteins in living organisms, tissues, or cells of complex living things at a large scale in a certain condition or period [61]. The proteome of these species is vary depending on dynamic and environmental factors, despite the fact that they have a single and unique DNA. Proteomics research is complicated and unclear as a result of this. When we look at the human genome, for example, we find roughly 23000 genes. The number of splice variants made up of these genes, as well as the number of proteins made up of these variants. We come across a significant number of proteins when we consider their post-translational modifications. It's easy to see how complicated proteomics investigations are when we realize that around 1 million distinct proteins can be generated in a dynamic process [60].

1.4.1 Mass Spectrometry

Mass spectrometry is a highly sensitive and specific method of choice to analyze proteins or specify additional particulars such as post-translational modifications (PTM). Different strategies generated and different integrations have been achieved in time after the first setup of the mass spectrometry. The first-generation mass spectrometers were just able to weigh the molecules. After the generation of new strategies, mass spectrometers gain the ability to select the specific ions and fragment these selected ones. This multi-step approach is specifically called Tandem Mass Spectrometry (MS/MS). Basically, in MS/MS, the first analyzer (MS1) selects specific ions (precursor ion) and measures their masses and charges. After selection, selected ions fragmented again in the collision cell. Precursor ions are separated into smaller fragments and these fragments (product ion) are analyzed by a second analyzer (MS2). To undergo these multiple steps, one of the most commonly used fragmentation approaches is the Collision-Induced Dissociation (CID).

There are different types of MS analyzers for proteomics studies including ion-trap, time-of-flight (TOF) quadrupole, and Fourier transforms ion cyclotron (FT-MS). These differently designed analyzers are performing differently. The combination of these analyzers and MS creates different types of MS such as TOF-TOF, Triple Q, and Q-TOF (Figure1.5).

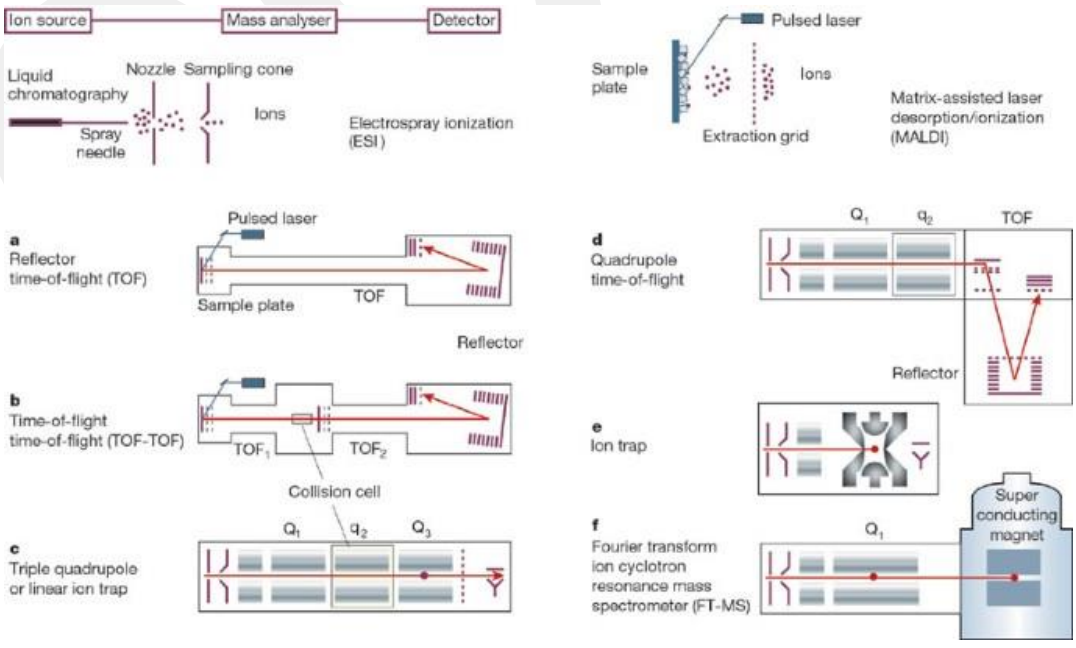


Figure 1.6 Mass spectrometry types used for proteomic studies [62].

One of the breakthrough developments of proteomics studies is the electrospray ionization's (ESI) integration to the MS. This complex system can be summarized as, an analyte sprayed by an ESI needle, with high voltage and temperature [63]. This spraying ends up with the charged droplets. Charged ion covered by solvent and this solvent evaporate by the heat on its way through the analyzer.

Different chromatography systems are integrated into MS systems to separate the complex samples. Liquid chromatography (LC) is one of the integrated systems and commonly used ones. These chromatographies can be classified depending on their separation mechanisms and these classes are named, affinity, reverse-phase, size-exclusion, and ion-exchange. Reverse phase chromatographies are the most common choice to combine with MS systems.

1.4.1.1 Separation Techniques

Separating the peptide is crucial before they get into ionization. This separation is reducing the complexity of the sample and helps to analyze peptides more easily. There are several separation techniques such as ion-exchange chromatography, size-exclusion chromatography, gas, and liquid chromatography. Reverse-phase liquid chromatography (RP-LC) is a highly used method to separate peptides for proteomics applications. RP-LC setup includes a non-polar stationary phase and a polar mobile phase. While silica and octadecylsilane (C18) are frequently preferred as non-polar phases in the columns, the most frequently used mobile phases are water and acetonitrile (ACN). When the sample enters the column, it moves through the chromatography column. This moving couples the peptide with the stationary phase. After coupling, peptides are eluted from the column with different percentages of the polar solvent. As the percentage of organic solvent increases, peptides with increasing hydrophobicity eluted from the column. This fractionation helps the system to analyze the peptide [64-66].

1.4.1.2 Ionization Techniques

The invention of soft ionization technologies, which allow proteins and peptides to be studied by MS, is one of the most significant advancements in instrumentation. Proteins and peptides are polar, nonvolatile, and thermally unstable species that necessitate an ionization approach that allows analytes to be transferred into the gas phase without severe degradation. electrospray ionization (ESI) and matrix-assisted laser desorption ionization (MALDI) are two techniques that paved the way for modern bench-top MS proteomics.

Electrospray Ionization

The most typical method of producing charged ions is electrospray ionization (ESI) (Figure1.6). At the end of the ionization, the sample results in droplets. Spraying of the droplets is achieved by a continuous flow of the sample with a capillary injection needle at a high electrical voltage (e.g., 2.5-6 kV). The charged droplets are surrounded by nebulizing gas, which helps the droplets flow, while they fly over the analyzer, the solvents surrounding the droplets evaporate with the help of the high temperature and reduce the size of the droplets [63].

Matrix-Assisted Laser Desorption Ionization

MALDI uses laser pulses to sublimate and ionize materials from a dry, crystalline state and is used for slightly simple samples when it's compared to the electrospray combined MS analysis. Mostly combined with 2D SDS PAGE [62].

Figure 1.7 Illustration of electrospray ionization.

1.4.1.3 Fragmentation Methods

For fragmentation of the ions, different mass spectrometry modes can be used such as CID, Higher-energy Collision Dissociation (HCD), and Electron Transfer Dissociation (ETD).

Collision-Induced Dissociation

Collison-Induced Dissociation is a technique used for peptide fragmentation with the use of inert gas such as nitrogen or argon. As a summary of this technique, precursor ions are selected by analyzers and flow through the collision cell. In collision cells, peptide ions and gas particles collide with each other. This collision breaks the peptide back bonds and peptide ions are fragmented into fragment ions [67- 69].

Higher-energy Collision Dissociation

Higher energy collision dissociation is an improved method of CID. HCD mode uses a higher voltage than CID to acquire a higher level of fragmented ion [68]. HCD mode has advantages and disadvantages depending on the aim of the studies. When the goal is obtaining the maximum level of the fragmented ions it may also cause a low-level trapping rate in the low range [70].

Electron Transfer Dissociation

Electron transfer dissociation is a useful method to analyze highly charged (z>2) molecules [71]. ETD has the ability to fragment long peptide/protein. This ability makes the method useful and preferable for Top-Down proteomics [72].

1.4.1.4 Mass Analyzers

Mass analyzers are separate the ionized or fragmented ions based on their m/*z* ratio and guide them to the detector. Different types of analyzers are used for proteomic analysis. Four main mass analyzers can be named as, TOF, Quadrupole, Ion Trap, and Orbitrap. These analyzers may also use in combination. Such as Quadrupole Ion Trap or Quadrupole-TOF.

Time of Flight Mass Analyzer

In the TOF analyzer, any electric or magnetic field is not used. The ion masses are fly into a chamber called a drift tube with has a certain length. The analyzer measures their flight time at this certain length. With the different masses, ions have different flight times. Depending on these differences analyzer separates the ion masses [73].

Quadrupole Mass Analyzer

The basic logic of the quadrupole analyzers, consists of four rods, and oscillating electricity is applied to these rods in a vacuumed field and ionized masses behave differently depending on their m/z ratios [74].

The TripleTOF® 5600+ contains four total quadrupoles in the MS that was used for this thesis. QJet, the first quadrupole, is positioned just ahead of Q0. Ions are not filtered by QJet; ins ad, they are guided through the Q0 quadrupole. This guidance

enhances the signal-to-noise ratio and increases sensitivity. The Q0 quadrupole refocuses them through the Q1 quadrupole. Q1 is capable of filtering ions before they enter the collision cell and can operate in two modes. To begin the TOF MS analysis, the Q1 quadrupole sorts all ions and passes them through the Collison cell with the specified m/z ratio. TOF systems analyze all ions that pass through. Second, Q1 sorts only one ion with the specified m/z ratio for TOF MS/MS analyses. Only this specific ion is analyzed by TOF systems. Finally, in the Q2 quadrupole, collisions with gas molecules increase the internal energy of the ions in the Q2 collision cell to the point where molecular bonds break, leading to product ions [75].

Ion Trap Mass Analyzer

Ion trap analyzers, as understood from the name, they trap the ions in a certain field with the use of an electromagnetic field. With the help of the gas phase, ions are move through the center. Ion trap analyzers produce both positive and negative ions. It is also employed as a GC-MS detector [76].

Orbitrap Mass Analyzers

Because there is no RF or magnetic field, this is simply an ion trap. In this experiment, moving ions are trapped around an electrode, and electrostatic attraction is controlled by centrifugal force, which is caused by tangential velocity [77].

1.4.2. Top-Down Proteomics

Targeted proteins are studied in depth using a top-down proteomics approach. Using a 2D-PAGE and MALDI-TOF combination or Ion trap MS, this method examines the sequence variants of target molecules. The instruments that are widely used for top-down proteomics are the Fourier transform ion cyclotron resonance (FTICR), Quadrupole trap, and Orbitrap MS. The top-down approach analyzes intact proteins, as opposed to protein digests evaluated using the bottom-up approach, resulting in lesser sample complexity. After MS analysis of all intact proteoforms in a sample, a specific proteoform of interest can be extracted and then fragmented in the mass spectrometer using tandem MS (MS/MS) methods to map both amino acid changes and PTMs [78, 79].

1.4.3 Bottom-Up Proteomics

Bottom-Up proteomics, also known as shotgun proteomics, is used to determine the amino acid sequence and posttranslational modifications of proteins and may be called peptide-based proteomics. In this approach, proteins are digested into peptides with the help of proteolytic enzymes such as trypsin which is lyse the protein from their C-terminal arginine (Arg) or lysine (Lys) residues. Then, digested peptides are separated with mostly the use of liquid chromatography based on their hydrophobicity levels. After separating the peptides, they are ionized and continue in the gas phase. These ionized peptides are called precursor ions. After all these steps, the ionized precursor ion flew into the mass spectrometer. With different methods, mass spectrometers guide and store these ions and then fragment them to achieve MS/MS data. These data are analyzed in the sequence database or using the peptide spectra library. This scanning process is usually done with special software for this process. MASCOT, SEQUEST, Protein Pilot can be given as examples [62]. In this method, where large-scale data are obtained, a large amount of protein can be identified from complex samples [80]. The bottom-up approach is also well suited for chemical modification of peptides, with the aim of peptide and protein quantification [80].

Data-dependent acquisition (DDA), selected or multiple reaction monitoring (S/MRM), and data-independent acquisition (DIA) are the three key bottom-up methodologies for assessing the proteome. In the DDA strategy, the precursor ions with the most intense value are selected depending on the MS1 m/z values and MS/MS analysis is continued with these peptides. DDA is a fast technique since MS/MS analysis of all peptides entering mass spectrometry in the sample is not conducted. This fast strategy is commonly utilized in biomarker investigations [81]. In DIA on the contrary of DDA, without listing any pre-knowledge all the precursor ions are fragmented. Full knowledge of fragment ions for each precursor created from a sample is generated using DIA methods (Figure1.7).

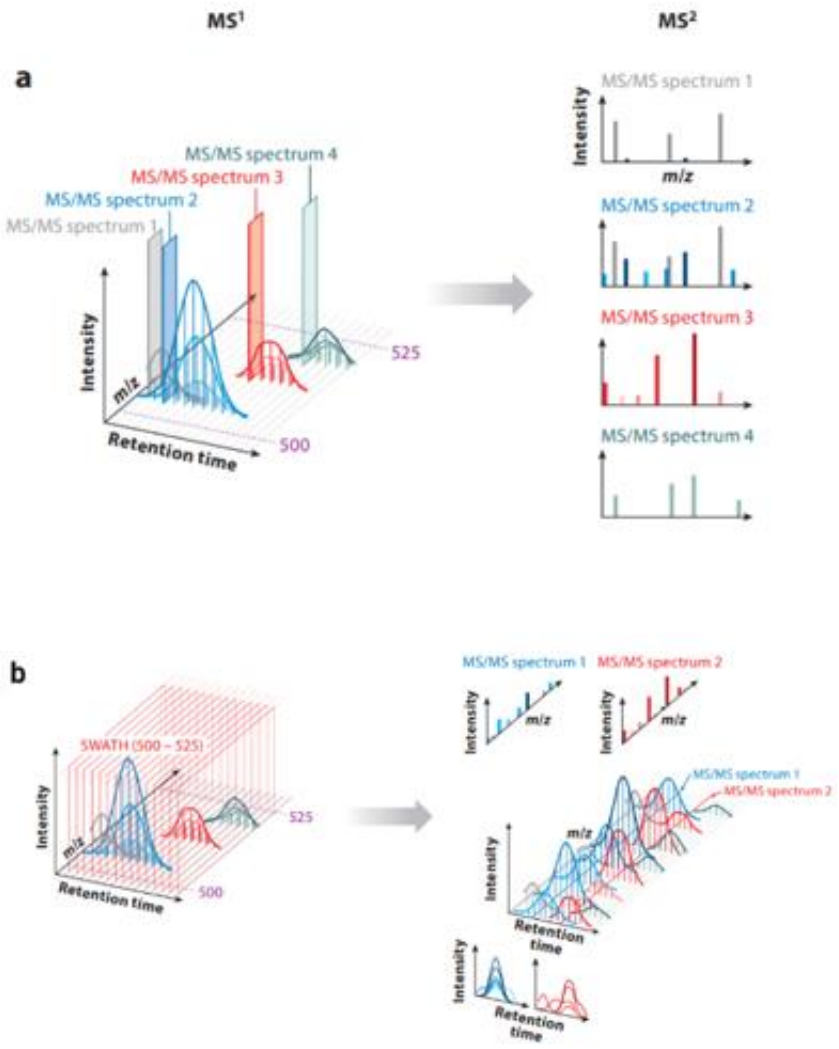


Figure 1.8 Illustration of DDA and DIA strategies. Colors represent ion signals. Adapted from Gillet et al., 2016 [82].

1.5 Cell Surface Protein Enrichment

Membrane proteins are essential for a cell's healthy development and perform a variety of roles in cellular growth, signaling, and other processes. Their expression accounts for about 30% of the total protein. Aside from their expression ratio, they have challenges to work with. These challenges arise from their low abundance, hydrophobicity, and highly heterogeneous post-translational modifications [83]. Membrane proteins are embedded on the cell lipid bilayer and this condition makes them non-soluble in the aqueous environment. Because of all these features, membrane proteins should be enriched and solubilized before mass spectrometry analysis.

1.5.1 Cell Surface Protein Isolation Strategies

Different methods are used in surfaceome study history which can be classified into methods that use Physico-chemical properties of membrane proteins, enzymatic shaving, and chemical enrichments which are used for this thesis [27].

Physico-chemical approaches include ultracentrifugation and surface property-based enrichment techniques. Ultracentrifugation is one of the most commonly used ones in physicochemical approaches. This old method separates the molecules depending on the differences between their size and density. As a summary of the workflow of this method, cells are disrupted in proper buffer without disrupting the nuclei and organelles. Then, various subcellular components are isolated using sedimentation rates that are an order of magnitude difference. This method is easily applicable and requires no chemical modifications. Besides these advantages this method also has limitations. First of all, for enriching the molecules with a low abundance this method requires a decent sample quantity, to begin with. Also, this method requires several centrifugation steps for hours. This long-time procedure may cause sensitive proteins to lose their structural integrity. Another limitation of this method is the organelle contamination. Because of their similar size and density, some organelles may be isolated with the surface proteins and cause low purity of enrichment [27, 31, 84]. Overall, the ultracentrifugation method is a low-throughput technique that could prevent the preparation of a large number of samples reproducibly.

Enzymatic enrichment methods are simply shaving the surface proteins. With the use of non-specific enzymes, surface proteins are directly collected from the cells. Detergents, acids, or any other chemicals are not necessary for this method which enables to observe protein's native topology. Besides being easily applicable, time-saving, and high specificity of the surface proteins, this method requires whole membrane integrity. The method is successfully applied for prokaryotes. Any disruption of the membrane will affect the intracellular proteins which are disrupting the purity of the results [27, 81, 85].

Chemical enrichment methods either label the proteins with chemical bonding or chemically modify a part of the proteins. One of the most widely used chemical enrichment is the biotinylation of surface proteins. In this method, one reactive group is attached to the surface protein and labeled proteins are purified with the help of avidin beads. The success of this method arises from a highly strong ($K_d \approx 10^{-14}$ M) non-covalent bond between the biotin and avidin beads [86]. Hydrazide chemistry is another chemical

enrichment method used for surface proteins. Briefly, glycan groups are covalently coupled on the solid support. When surface protein glycosylation rate is considered, glycoenrichment is a very advantageous method.

1.5.1.1 Biotinylation of The Cell Surface Proteins

Biotinylation is a widely use chemical-labeling enrichment method for cell surface proteins. The reagent that used for biotinylation is composed of three main portions. The first part is included biotin moiety which is essentially used for affinity purification of the biotinylated samples with the streptavidin-derived molecules. The valeric acid side chain of the biotin moieties helps create a bond between biotin and beads [83]. The second part is the spacer or linker group. This group act as a bridge between the captured protein by the ester group and the biotin molecule to purify the protein of interest. The Spacer group is important for blocking the biotin group move into the cellular space and labeling the intracellular proteins [87]. Also, the linker group with a strong interaction with the reactive and biotin group helps the labeled protein stay on captured streptavidin or derived beads under different conditions. The third part is the reactive group. Proteins contain a range of functional groups. These groups are enabling the protein for chemical modifications. The targeted functional group must be expressed in an accessible form on the cell surface for successful surface protein labeling. Therefore, reactive ester groups are interacts with the ε-amines in lysine side chains or N-terminus of the proteins (Figure 1.10). The reactive groups hydrophilicity may enhanced by the addition of Sulfo- groups to the ester moiety (Figure1.9). Also, addition of the Sulfo- group due to the molecules charge, prevents the cell permeabilization of the biotin label.

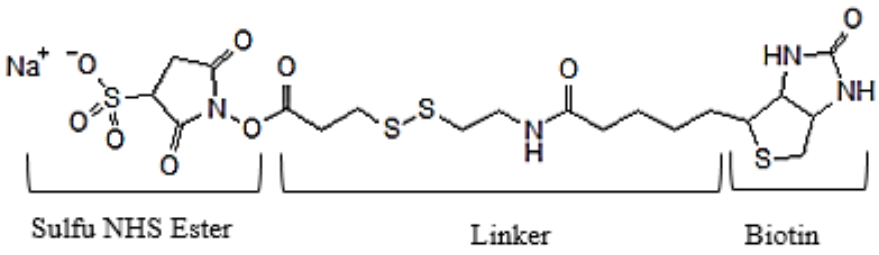


Figure 1.9 Sulfo-NHS-Ester Biotin molecular structure. Adapted from Thermo Fisher product site.

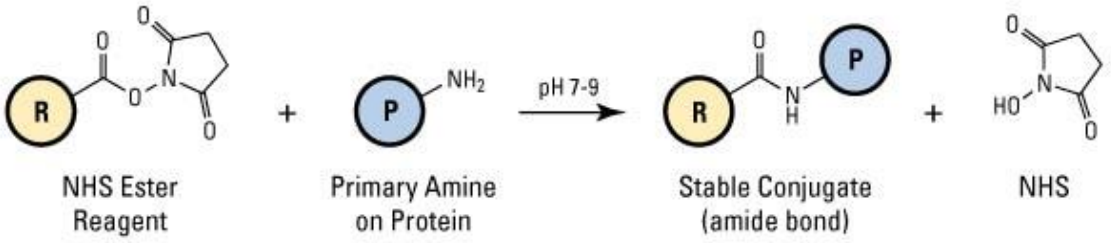


Figure 1.10 NHS Ester - primary amine conjugation. Adapted from Thermo Fisher product site.

1.5.1.2 Glycoprotein Enrichment of Cell Surface Proteins

Glycosylation is required for proper protein functioning because of its role in folding, intracellular motility, and ligand binding. It is also required for crucial cellular processes such as cell adhesion. Also, it is known that 90% of the surface proteins are glycosylated. This condition makes glycoprotein targeting a very effective strategy for surface protein enriching. In this method, cis-diol groups of the glycan moieties are oxidized to aldehydes with sodium periodate ($NaIO_4$) and occurred aldehydes groups attached to hydrazine solid support with covalent hydrazone bonds (Figure1.11). Because of this strong binding, the non-specific binding possibility can be eliminated by just several stringent washes [88]. For releasing the captured proteins or peptides, an amidase called PNGase F was used.

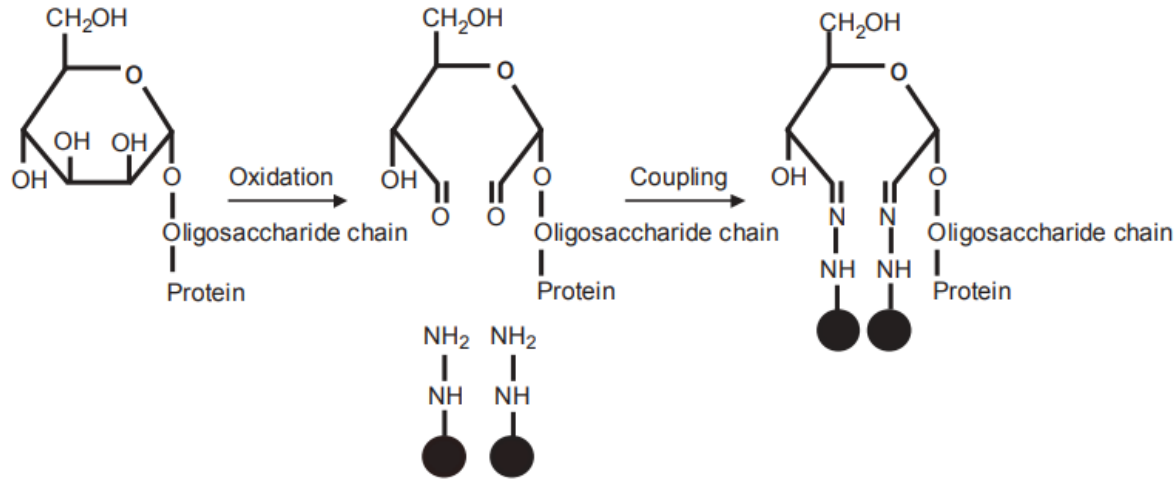


Figure 1.11 Oxidation of carbohydrate groups to an aldehyde for hydrazide coupling [89].

1.6 Aim of the Study

Cancer is one of the most death cause diseases all over the world. This common disease may affect different organs and systems. In childhood cancers, B cell leukemias compromise 80% of the diagnosed patient [4]. Because of their role in cancer, protein became more attractive to study with different approaches and for different purposes [6]. For instance, biomarker analysis in patient sera in chemotherapy-resistant small cell lung cancer was undertaken in the study conducted by Han et al. SELDI-TOF MS was used to perform the study, which included 60 patients and 48 control groups. They discovered two biomarkers as a result of their investigation, one of which is particular to resistance. Surface proteins were not used in the research indicated above [90]. In another study, Sun et al. used dimethyl labeling and SWATH-MS analysis to identify a glycoprotein biomarker in an Adriamycin-resistant chronic myeloid leukemia (CML) cell line. In this investigation, 15 glycoprotein biomarkers were found along with known MDR-related proteins by comparing resistant and non-resistant lines' glycoproteins. The experiment in question is focused on pancreatic cancer and exclusively compares glycosylated proteins [91].

This thesis study aimed to map the membrane of the CCRF-SB cells (B-cell acute lymphoblastic cell line) with two different surface enrichment approaches as summarized in Figure 1.9. Both approaches have advantages and drawbacks. Enrichment of glycosylated proteins is a very efficient strategy for analyzing surface proteins considering 80% of them are glycosylated [92]. But considering their less abundant level compared to an intracellular protein, the remaining ~20% non-glycosylated membrane proteins are a huge loss for surface proteomics studies. Taking consideration of this, another enrichment approach, biotinylation has been chosen. For the biotinylation strategy, un-accessible primary amines can be considered as drawbacks however, since their inaccessibility may be due to heavy N-linked modifications, the glycol-enrichment method will also be able to cover the disadvantages at this point. Identification of the enriched surface protein can result in a list of proteins that would include possible candidates for diagnostic or therapeutic approaches.

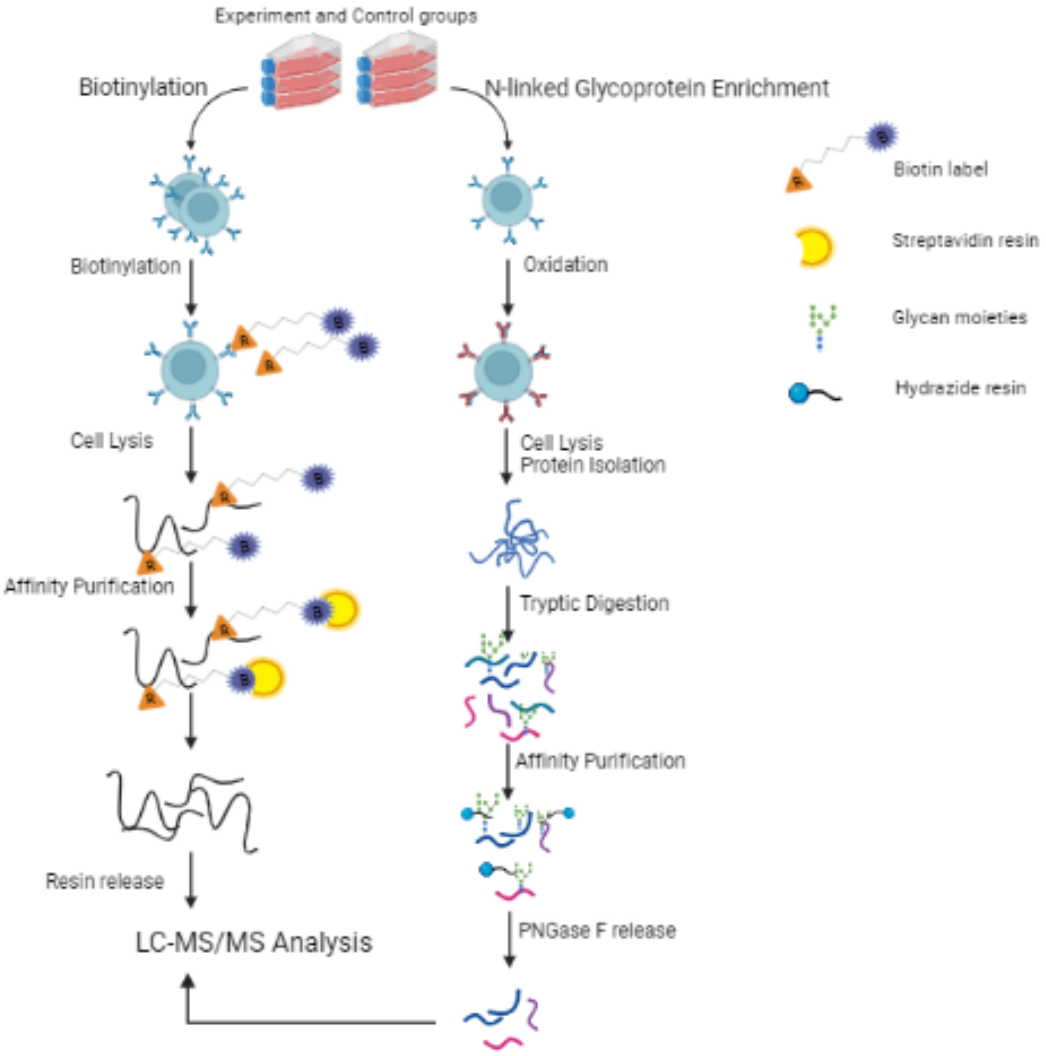


Figure 1.12 Summary of the followed protocols.

Chapter 2

Material and Methods

2.1 Cell Culture

CCRF-SB human B cell lymphoblastic leukemia cell line (ATCC, Rockville, MD) were cultured in RPMI 1640 medium supplemented with 10% Fetal Bovine Serum (FBS), 100 U/mL penicillin/streptomycin, 4mM L-Glutamine, and 1% D-Glucose at 37°C in 5% CO_2. CCRF-SB cells were seeded out into 75 cm^2 sterile low-bind flasks as 1.0 x 10^6 cells/ml and were split into 1:3 to 1:4 every 2-3 days. The cells were collected and centrifuged for 10 minutes at 125 x *g*. After discarding the supernatant, the pellet was dissolved with fresh complete media. Cell stocks were stored with freezing media which includes 70% FBS, 10% RPMI 1640, and 20% DMSO, and stored in a nitrogen tank at -196°C.

2.2 Isolation of Membrane Proteins

2.2.1 Biotinylation of the Membrane Proteins

For the biotinylation of the surface proteins, Pierce™ Cell Surface Protein Biotinylation and Isolation Kit (Catalog Numbers A44390) was used with the following protocol. 15 x 10^6 cells were collected into a conical tube and centrifuged at 125 x *g* for 10 minutes. After discarding the supernatant, the pellet was dissolved in 15 mL of phosphate-buffered-saline (PBS, 10 Mm, pH 7.4). 6 mg of Sulfo-NHS-SS-Biotin was dissolved in 1 mL of PBS to achieve a 24X stock biotin solution. 15 mL cell suspension was centrifuged again at 125 x *g* for 10 minutes. Following the supernatant discard, the cell was suspended in 11,5 mL PBS. 500 µL 24X Biotin solution was added to the cell suspension and incubated at room temperature for 10 minutes. For the control group except for the addition of biotin, the same procedures were followed. Then, the cell suspension was centrifuged at 125 x *g* for 10 minutes. The cell pellet was resuspended in

15 mL ice-cold tris-buffered-saline (TBS) to stop the labeling reaction and centrifuged at 300 x *g* for 5 minutes and the excess reactive biotin was removed from the solution. The cell pellet was resuspended in a 500 µL Lysis Buffer which is supplemented with a 10 µL protease inhibitor (Roche, Germany). To facilitate the lysis the solution was pipetted up and down several times. Afterward, cells were incubated on ice for 30 minutes and vortexed intermittently. After incubation, the lysate was centrifuged at 15 000 x *g* for 5 minutes at 4°C. 50 µL Streptavidin agarose beads (Pierce™, Product number: 53116) were washed with 200 µL ddH_2O and Lysis buffer with centrifugation at 300 x *g* for 5 minutes. Cell lysate and streptavidin beads were incubated for 1 hour at room temperature on a rotary mixer. Streptavidin beads were centrifuged 300 x *g* for 5 minutes and the supernatant was transferred to a fresh tube as a flow-through. To get rid of nonspecific binders, streptavidin beads were washed with 300 µL lysis buffer three times. Finally, the beads were washed with 50 mM Ammonium Bicarbonate (pH: 7.8) and the beads were stored at -80°C. The stored beads were thawed before being analyzed by mass spectrometry. 200 µL 50 mM Ammonium Bicarbonate added onto beads with the addition of 4 µg trypsin (BI, Trypsin B solution). Solution incubated overnight at 37°C. After overnight incubation beads are centrifuged at 2500 x *g* for 5 minutes and the supernatant collected. Beads are washed with 100 µL 50 mM Ammonium Bicarbonate with centrifugation and washes were combined with supernatant. Samples are cleared with C18 (Empore) In Stage tips with the following protocol. Three C18 disks were inserted into 200 µL pipette tips (Diamond, Gilson). To condition, the disks, 50% ACN was added onto tips and centrifuged at 10.000 x *g* for 10 minutes. For equilibration, disks are washed three times with a washing solution (1% TFA). The sample is acidified with TFA loaded onto equilibrated columns and centrifuged at 5000 x *g* for 5 min. The loaded sample is washed three times with a washing solution. After washing, the conjugated sample was eluted twice with 200 µL elution buffer (80% ACN). Eluted sample dried with SpeedVac (SpeedVac, Eppendorf). Dried samples are stored at -20°C until LC-MS/MS analysis.

2.2.2 N-Linked Glycoprotein Enrichment

For the surface glycoprotein enrichment protocol adapted from Tian et al, Zhang et al., and McDonald et al [88, 89, 93]. 30 x 10^6 cells were transferred into a conical tube and centrifuged at 125 x *g* for 10 minutes. The cell pellet was resuspended in labeling buffer (PBS pH: 6,5). Then, cells were oxidized in 2 mM sodium periodate ($NaIO_4$). Incubated at RT, dark, and at 4°C. After incubation, cells were washed with DPBS (Ph: 7,4) at 300 x g f or 5 minutes. The washing step was repeated 3 times. At the end of these steps, the cell pellet was dissolved in 2 mL denaturing buffer (8M Urea, 0,4M Ammonium Bicarbonate, 0.1% SDS). The Sample was sonicated for 6 minutes (10 seconds on 10 seconds off) at 4°C. Sonicated samples protein concentration determined with BCA Assay (Biorad, Protein Assay Reagent, 660 nm) and continued with 1 mg protein. 10 mM Tris(2-carboxyethyl) phosphine (TCEP) and 40 mM Chloroacetamide (CAA) (Sigma, CO267-500G) were added into 1 mg protein and incubated at 60 °C for 1 hour with gentle shaking. After denaturing the proteins urea concentration was reduced to less than 2M to ensure the trypsin worked. 20 µg trypsin (BI, Trypsin B solution) was added into denatured protein samples and incubated overnight at 37 °C. After incubation, digestion efficiency was checked on SDS PAGE. The digested sample was cleared with centrifugation at 12 000 x *g* for 10 minutes.

Digested peptides were cleared with C18 columns with the following protocol (Finisterre SPE Columns C18/17% 100mg/1ml, Cat. No: TR-F034000). Sample acidified with 1% TFA before loading to columns. Columns are conditioned with 1 mL conditioning buffer twice which includes 0.1% TFA and 50% ACN. After conditioning it, columns are equilibrated with 1 mL 0.1% TFA twice. Sample loaded onto columns and columns are washed three times with washing buffer (0.1% TFA). Coupled samples are eluted with 400 µL elution buffer (0.1% TFA in 80% ACN). Eluted samples pH adjusted into 5.5 with PBS. For the oxidation of glycan groups, eluted peptides were split into two as control and sample groups and sodium periodate ($NaIO_4$) was added onto sample peptides with the 10 mM final concentration and incubated at 4 °C for 1 hour at dark. After this step for the sample and control, the same steps are followed. The oxidized sample was acidified again with the addition of 1% TFA. The acidified sample was loaded onto a column that is conditioned and equilibrated like described above. Loaded sample eluted with elution buffer. 75 µL hydrazide resin (Biorad, catalog number: 1536047) was washed with 1 mL of ddH_2O to remove the storage solution and then washed with 500

µL elution buffer to equilibrate with the sample. Washing steps were done at 300 x *g* for 3 minutes. Eluted samples and washed hydrazide beads got together in low-bind tubes and incubated overnight at 37°C with gentle mixing. After overnight conjugation, the sample was centrifuged at 300 x g for 3 minutes and unbound peptides (upper part) discarded. Beads were washed three times each with 1.5 M NaCl, ddH_2O, and 50mM ammonium bicarbonate. Then, beads are suspended in 40 µL ammonium bicarbonate and 2000 Unit PNGase F (Serva, 36405.01) was added. Samples were incubated overnight at 37 °C. Beads were centrifuged at 2500 x g for 5 minutes. The eluent was transferred into low-bind tubes, beads were washed twice with rapigest to collect the eluted peptides.

The combined elutions were cleaned with C18 (Empore) In Stage tips with the following protocol to get rid of salts and detergents. Three C18 disks were inserted into 200 µL pipette tips (Diamond, Gilson). To condition, the disks, 50% ACN was added onto tips and centrifuged at 10.000 x *g* for 10 minutes. For equilibration, disks were washed three times with a washing solution (1% TFA). The sample was acidified with TFA loaded onto equilibrated columns and centrifuged at 5000 x *g* for 5 min. The loaded sample was washed three times with a washing solution. After washing, the conjugated sample was eluted twice with 200 µL elution buffer (80% ACN). Eluted sample dried with SpeedVac (SpeedVac, Eppendorf). Dried samples were stored at -20 °C until LC-MS/MS analysis.

2.3 Sample Preparation for Mass Analysis

Dried samples were dissolved in the sample solution (5% ACN, 0.1% Formic Acid). Each sample was dissolved in a 12 µL sample solution. Solutions were vortexed for 30 seconds and sonicated in the ultrasonic water bath for 1 minute. After sonication, samples were vortexed for 30 seconds, spin downed, and this process is repeated three times. Afterward, samples were centrifuged at 10 000 x *g* for 5 minutes. Centrifuged samples were transferred into thin-glasses ad MS tubes. Transferred samples were placed in an autosampler.

2.4 LC-MS/MS and Data Analysis

2.4.1 LC Method and MS Analysis

Analysis was carried out by taking 4 µL of the sample loaded into the device. Samples passed through the attached trap column were directed to the analysis column. NanoACQUITY UPLC® Column (1.8 µM HSS T3 75 µM x 250 mm) (Waters) was used as the analysis column. The 180-minute method was used during the separation of the complex peptide mixture. During the analysis, the flow rate was determined as 250 nL/min. In the first 5 minutes of the analysis, 96% H2O and 4% ACN were used. From the 5th minute to the 155th minute, the ACN ratio was increased to 45.6%. The ACN ratio was increased sharply to 85% in the 162nd minute and continued in this manner until the 172nd minute. Starting from the 172nd minute, the ACN ratio was reduced to 4% at the 177th minute and the analysis was continued in this way until the 180th minute.

Peptides eluted by LC were ionized by applying 15 units of Gas1, 25 units of Curtain Gas, using 2400 V ISVF (ion source voltage frequency) and 75°C IHT (interface heater temperature) conditions.

ABSciex TripleTOF® 5600+ brand and model Hybrid Quadrupole-TOF mass spectroscopy was used for the analysis of peptides after ESI. DDA (Data Dependent Acquisition) method was used for the analysis of peptides. The process was performed by selecting the best 35 candidate molecules (Top35) in terms of mass and charge in each analysis cycle (Switch criteria: >350.000 m/z, <1250.000 m/z, Ion tolerance 150.000 mDA). The analysis was performed at 180 minutes and an average of 5250 cycles.

The selected molecules were analyzed in two different experiments within a single period. First, precursor ion (Precursor-lon) analysis was performed using Positive TOF MS scanning type (Accumulation time: 0.250 sec, Collision Energy (CE). 10, Declustring Potential (DP): 100, TOF Masses 350.0000-1250.0000 Da). In the second experiment, product ions obtained by fragmentation of precursor peptides were analyzed using Positive Product Ion scanning type. (Accumulation time: 0.0500 sec, Collision Energy (CE): 35 Collision Energy Spread (CES). 10, Declustring Potential (DP): 100, TOF Masses 230.0000-1500.0000 Da) Calibration of the device was performed once after each injection. A peptide mixture of B-Galactosidase protein was used during calibration. During the 50-minute analysis, a calibrant reference table was created by using the

retention times of 14 peptides in the peptide mixture. In addition, the products of the peptide with a mass of 729 3650 Da were used during the calibration.

2.4.2 Data Analysis

Analysis of raw data generated by mass spectroscopic reporting and measurement of multiple analytes in a single sample was performed with Analyst® TF v 1.6 (ABSciex) Evaluation of peptides and product-ions was performed with Peak View (1.2) Mass fingerprint ProteinPilot 4.5 Beta (ABSciex) generated by the reporting of the device) were analyzed with the program The current protein session downloaded from UniProtKB was used to identify proteins, identify modifications, isoforms, and protein subsets.
For the identification of proteins, Carbamidomethyl was chosen as a fixed modification. For the analysis of biotinylated samples, Thioacyl modification was chosen as a modification. The removal of N-linked sugar causes releasing the entire sugar chain which corresponds to the mass of Asn to Asp substitution and causes a characteristic mass shift.

To investigate the cellular component annotation of the data PANTHER™ GO slim was used. To illustrate the Gene Ontology Cellular Component data gProfiler was also used.

2.4.3 Flow Cytometry Analysis

CCRF-SB cells,~10^6, cells are collected and centrifuged at 300 x *g* for 5 minutes. Then, washed with DPBS. Cell pellet dissolved in 50 µL Cell Wash Buffer (2 mM EDTA, 1X PBS pH: 7,4) to minimize the non-specific bindings. 5 µL Anti CD-10 (BioLegend, cat. no: 312210) and 2 µL Anti-CD19 (Miltenyi Biotec Ord. No: 130-113-171). Samples are analyzed by flow cytometry (BD FACS Aria III). Evaluations were performed with FACS Diva 8.0.1 software.

Chapter 3

Results

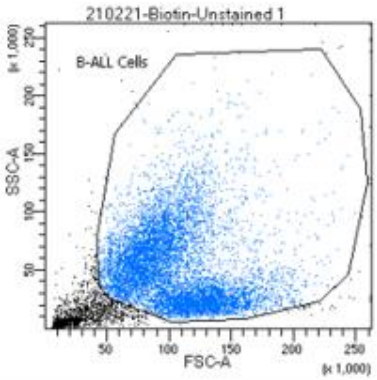
3.1 Surface Biotinylation Results

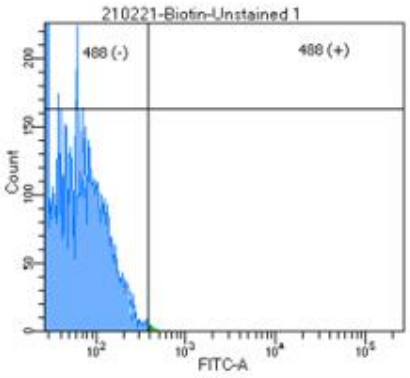
3.1.1 Biotinylation and Isolation of the Whole Membrane of the CCRF-SB B-Cell Acute Lymphoblastic Leukemia Cells.

For confirmation of biotin-labeled proteins, 2 µL of Streptavidin Alexa-488 (1/50 dilution ratio) was added to 100 µL of cell solution. For staining of biotin-labeled proteins, they were incubated at room temperature for 30 minutes in a cyclic mixer. After incubation, the samples were washed 3 times with 1 ml of PBS. Prepared samples were analyzed by fluorescence microscopy and flow cytometry (BD FACS Aria III). Evaluations were performed with FACS Diva 8.0.1 software (Figure 3.1).

To determine the autofluorescence of the cells, unstained, non-biotinylated cells were used, and the threshold value was determined based on these cells. Based on the flow cytometry data only 0.3% of the unlabeled and unstained (double negative control) cells showed autofluorescence. 99.7% of the cells showed up at 488- gate. This set is analyzed with flow cytometry to determine the autofluorescence level of the cells to create a threshold. Streptavidin-488 staining was performed on non-biotinylated cells to detect non-specific Streptavidin binding (Figure 3.1B). Three control groups showed 10.9%, 10.3%, and 9.2% non-specific binding of streptavidin, respectively. In Figure 3.1C, cells are labeled with biotin and stained with streptavidin. Flow data showed that 93,4% (average) of the cells are labeled with biotin and specifically stained with streptavidin. These flow data showed that our sample group cells were successfully labeled with biotin and stained with streptavidin-488.

a.

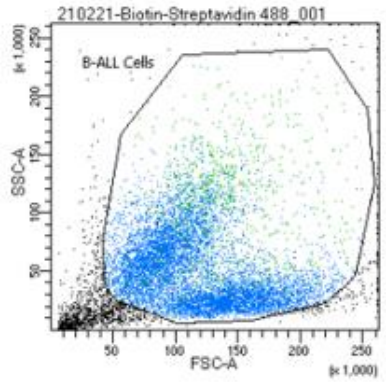


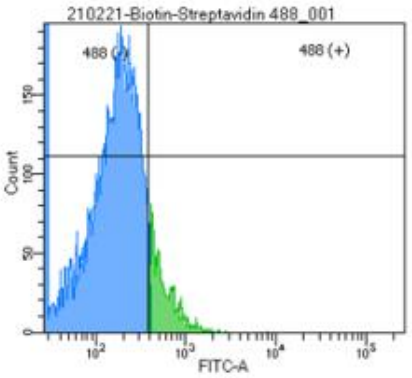


Tube: Unstained 1

Population	#Events	%Parent	%Total
All Events	10,000	####	100.0
B-ALL Cells	8,468	84.7	84.7
488 (-)	8,441	99.7	84.4
488 (+)	27	0.3	0.3

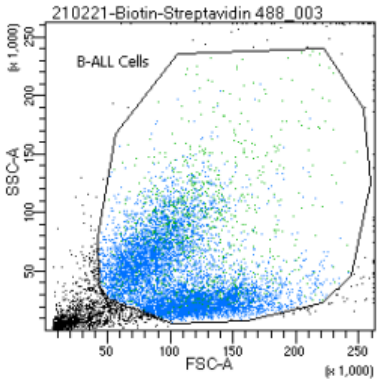
b.

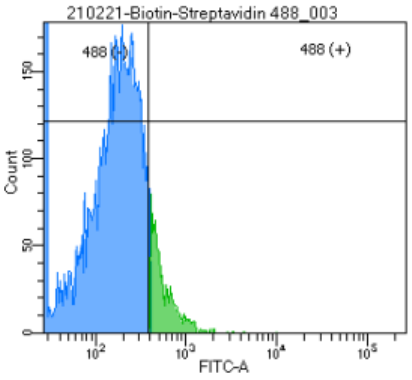




Tube: Streptavidin 488_001

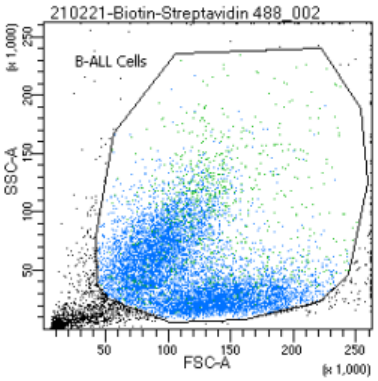
Population	#Events	%Parent	%Total
All Events	10,000	####	100.0
B-ALL Cells	8,366	83.7	83.7
488 (-)	7,456	89.1	74.6
488 (+)	910	10.9	9.1

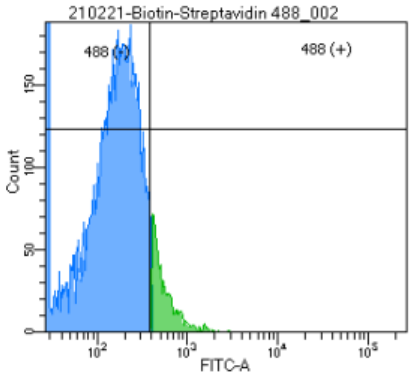




Tube: Streptavidin 488_003

Population	#Events	%Parent	%Total
All Events	10,000	####	100.0
B-ALL Cells	8,274	82.7	82.7
488 (-)	7,423	89.7	74.2
488 (+)	851	10.3	8.5

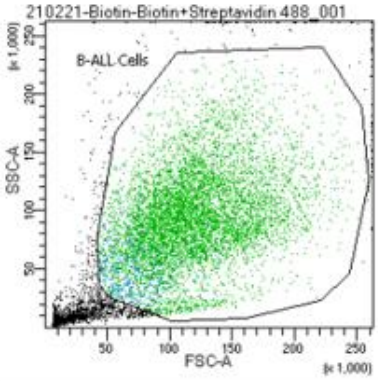


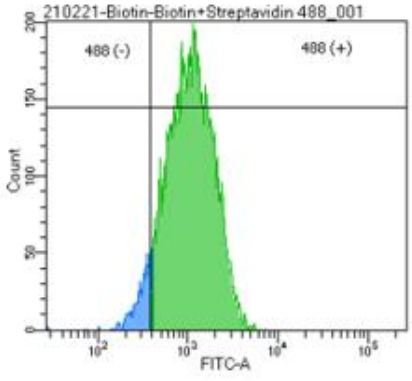


Tube: Streptavidin 488_002

Population	#Events	%Parent	%Total
All Events	10,000	####	100.0
B-ALL Cells	8,414	84.1	84.1
488 (-)	7,639	90.8	76.4
488 (+)	775	9.2	7.8

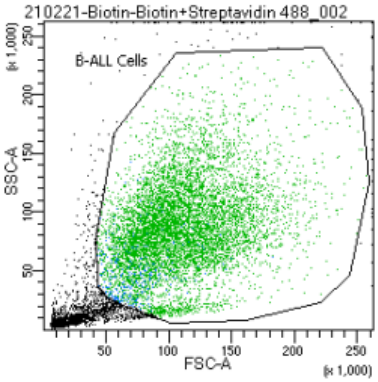
c.

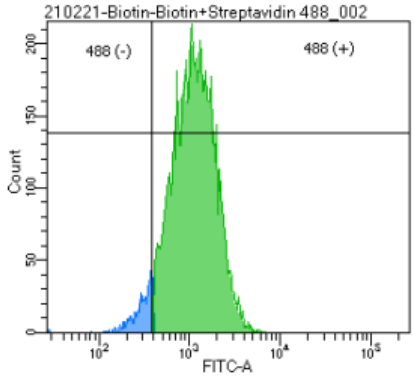




Tube: Biotin+Streptavidin 488_001

Population	#Events	%Parent	%Total
All Events	10,000	####	100.0
B-ALL Cells	7,829	78.3	78.3
488 (-)	599	7.7	6.0
488 (+)	7,230	92.3	72.3





Tube: Biotin+Streptavidin 488_002

Population	#Events	%Parent	%Total
All Events	10,000	####	100.0
B-ALL Cells	7,759	77.6	77.6
488 (-)	491	6.3	4.9
488 (+)	7,268	93.7	72.7

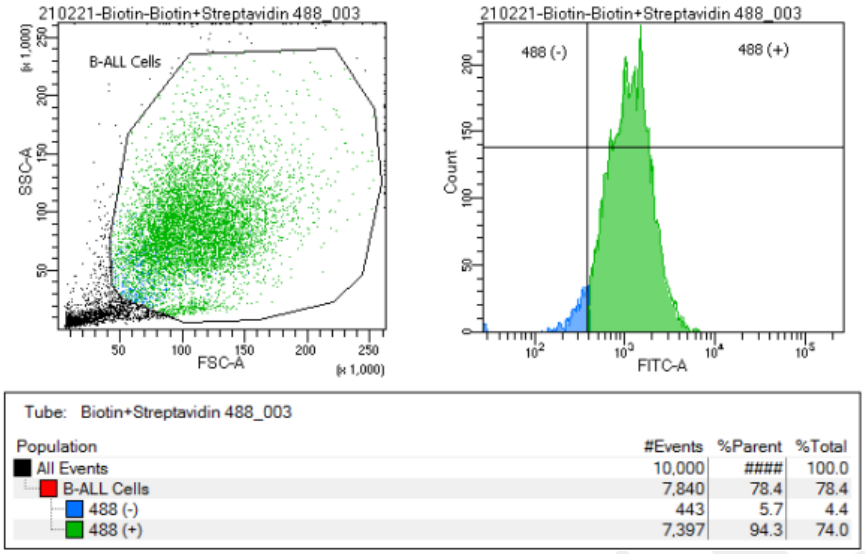


Figure 3.1 Verification of the surface biotinylation with flow cytometry analysis (a) Unstained control cells (b) Streptavidin-488 staining of cells without surface biotinylation (c) Streptavidin-488 staining of cells after surface biotinylation

In addition to flow cytometry analysis, validation of the biotinylation process was performed by observing the biotin-bound Streptavidin-488 molecule in the FITC filter under a fluorescent microscope. Images were taken on Zeiss Axio Vert.A1 with 100x magnification (Figure 3.2). Fluorescence microscope images showed that cells are labeled with biotin and stained with streptavidin.

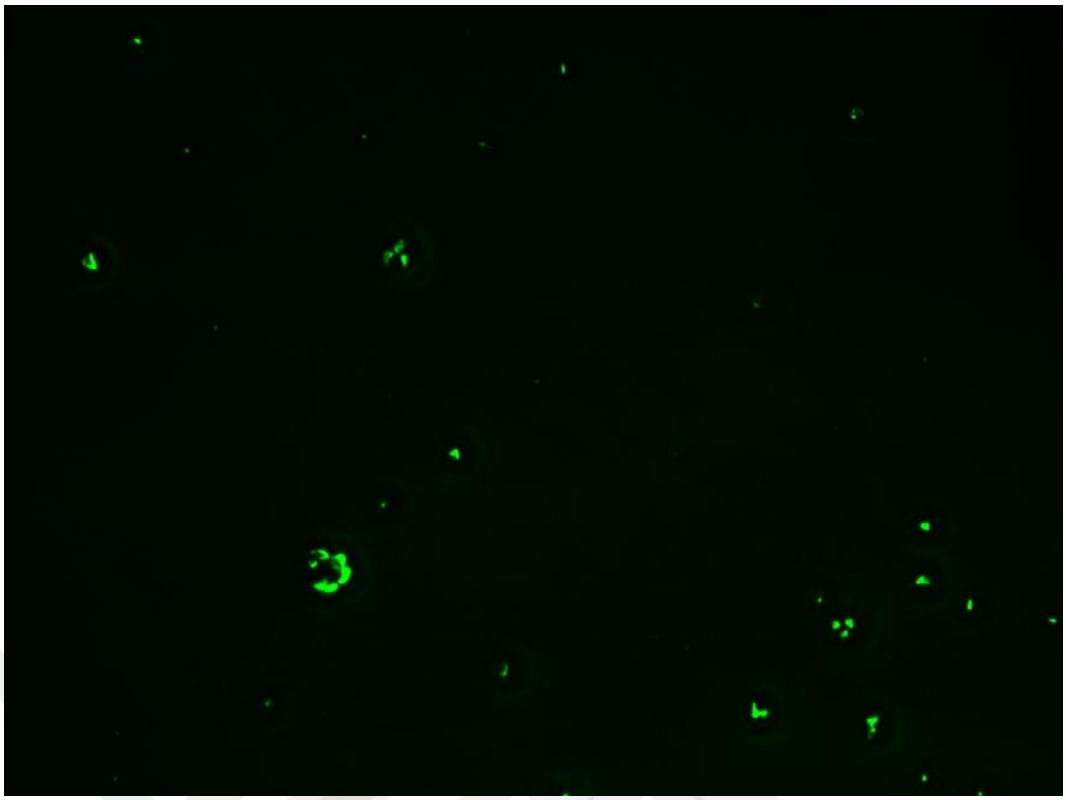


Figure 3.2 Fluorescent microscopy image of the biotin-labeled cells, Zeiss Axio Vert.A1 with 100x magnification.

3.1.2 Mass Spectrometry Data Analysis of Cell Surface Biotinylation

Mass spectrometry data are analyzed with Max Quant software. For the control groups, proteins were identified with the 1% False Discovery Rate (FDR). Once the common contaminants and reverse identified proteins were subtracted from the list of identified proteins in the control group, only 3 proteins were left and none of them was membrane protein. This result showed us there was nonspecific enrichment with streptavidin beads.

From the biotinylated group, 782 proteins were identified with the 1% False Discovery Rate. After the deduction of common contaminants and reverse proteins, a total of 748 proteins was further analyzed using bioinformatics tools (Table 3.2).

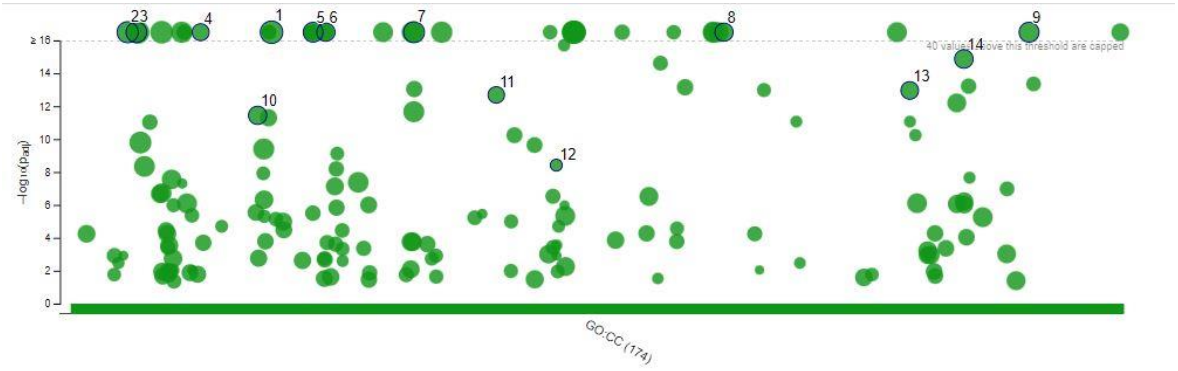
Later on, detected proteins were analyzed with Gene Ontology Cellular Component analysis. The analysis showed that 467 of the detected peptides were annotated with ‘Membrane’, 235 of those annotated with ‘Plasma membrane’, 77 of those annotated with ‘membrane protein complex’, and 66 of those annotated with ‘cell surface’ (Table 3.1). Cellular Component analysis showed that the presence of mitochondrial and

ribosomal proteins could be result be of the rapture of the cell membrane during biotinylation processes or not sufficient quenching.

Table 3.1 Gene Ontology Cellular Component analysis of Identified proteins from the MS data with Max-Quant

#termID	Term description	Observed gene count	Strength	FDR
GO:00106020	Membrane	467	0,19	1,07E-37
GO:0005615	Extracellular Space	263	0,59	2,82E-90
GO:0031982	Vesicle	369	0,46	4,98E-90
GO:0071944	Cell periphery	245	0,14	6,08E-07
GO:0005886	Plasma membrane	235	0,13	6,71E06
GO:0098796	Membrane protein complex	77	0,31	2,18E-07
GO:0009986	Cell surface	66	0,42	2,32E-11
GO:0005759	Mitochondrial matrix	30	0,48	6,46E-06
GO:0005840	Ribosome	55	1,09	9.73E-37

Gene Ontology analysis illustrated by gProfiler functional enrichment analysis. Analysis showed statistically significantly enriched terms that represent our data. Data annotated with extracellular space and associated terms.



ID	Source	Term ID	Term Name	p_{adj} (query_1)
1	GO:CC	GO:0016020	membrane	1.412×10^{-22}
2	GO:CC	GO:0005576	extracellular region	6.282×10^{-59}
3	GO:CC	GO:0005615	extracellular space	8.514×10^{-76}
4	GO:CC	GO:0005925	focal adhesion	5.082×10^{-58}
5	GO:CC	GO:0030054	cell junction	2.237×10^{-27}
6	GO:CC	GO:0030141	secretory granule	5.502×10^{-19}
7	GO:CC	GO:0031982	vesicle	1.478×10^{-87}
8	GO:CC	GO:0070161	anchoring junction	3.101×10^{-41}
9	GO:CC	GO:1903561	extracellular vesicle	1.929×10^{-122}
10	GO:CC	GO:0009986	cell surface	3.581×10^{-12}
11	GO:CC	GO:0034774	secretory granule lumen	2.049×10^{-13}
12	GO:CC	GO:0042611	MHC protein complex	3.815×10^{-9}
13	GO:CC	GO:0098552	side of membrane	1.096×10^{-13}
14	GO:CC	GO:0099503	secretory vesicle	1.346×10^{-15}

Figure 3. 3 Illustration of Gene Ontology Cellular Component analysis with gProfiler. Data annotated with extracellular space and associated networks.

Table 3.2 List of the identified proteins with surface protein biotinylation method.

Identified Proteins
ACTN4 RPS11 DBNL CD44 RPS17 YBX1 EIF2B5 ITGAM MAP4 ACTR3 FH CCDC183 PDE6A PTPRCAP APOL2 NCOA5 TKT HSPB1 CNPY2 PRKCSH RTN4 VAMP8 PTMA QARS DECR1 RPL26 COX5A SLC25A3 RPS18 AIFM1 MCCC1 GGT1 ZNF572 CNN2 HSD17B10 NT5DC1 LMO7 RAB1B RAC1 NCF1C HNRNPH1 UMODL1-AS1 ATP13A1 TUBA4A MTA2 SLTM ME2 RPL19 TLN1 EEF2 CHCHD3 HIST1H1B KRT9 MATR3 SERPINB8 UGGT1 BTLA DDX5 ACOT1 SMC2 HNRNPD CARM1 MS4A1 ATP5A1 SMARCC1 CHD9 ATP5J2 FGR SEMA4D C7orf55-LUC7L2 SQRDL THRAP3 TUBB2A CENPE RPL36 MARCKS HLA-DRB5 SPN LMNB1 MAX CD276 MSN RPS5 ALPK3 LSP1 SRP14 ACADM CD59 DERA PLCG2 CAPZA1 NSF RPL7A ERP29 SCARA5 UBA1 PGRMC1 RAB8B UVRAG CR1L ECH1 NAMPT GNAI2 FASN MRPS14 BST2 ELOVL5 ESAM TPM4 SFPQ ILF2 SLC39A14 RPS6 SFXN1 FARSB ATP5O FCGR2B PSMA1 TIGAR CAND1 CDO1 C5AR1 COX6B1 CES3 CEP290 RPL29 RPS20 KRT1 SLC44A2 ACACA RPL35 LILRB1 PDLIM1 G3BP2 HNRNPA2B1 CUL5 RPL21 DDRGK1 KPNA2 RAC2 RPS23 PDIA3 SNRNP70 RAP1B ACAT1 KPNA3 SLC7A5 EIF5 HMOX2 IGKC RBM14 SELL CAPRIN1 SF3A1 TAGLN2 SLC16A7 TAP2 LY75 EFHD2 HNRNPF UTS2 DNAH17 MDH1 LIFR DPP4 HLA-Cw SNAP23 RPL5 SEC24C MND1 HSPA9 PGRMC2 TNPO1 PSMC2 EEF1D FYCO1 MYL12A RAB6A RPS15A CAPZA2 GSTK1 TIAL1 RAPH1 MYO15A FSCN1 SRSF10 IL3RA HIST1H1E CYFIP1 ALDH18A1 LGALS1 TPI1 PCMT1 MBNL1 SLC25A5 IL2 MYO1C ATP1B1 WDR1 CDK8 RPL15 FAS SRSF6 ZNF844 TPM3 HNRNPDL ARNT HSPD1 RPS3A RPS28 MZB1 REXO4 IQGAP1 HSPA5 HYOU1

TUBB CFL1 CD82 TUBB6 COPA HNRNPR CAD RPLP1 COPG1 RPL34 ALCAM HLA-DPB1 KIF5B RPS12 CCT2 ZNF117 MYH9 CLPTM1 RPL28 HNRNPM SEC61G SRSF1 ATP5B IDH2 HSD17B4 IARS NPM1 CD74 SUPT5H RPL18A XRCC6 LAMTOR1 ZDHHC5 CORO1A MPP6 HLA-DRA EVI2B NOP2 PSMD11 SLC25A11 EHD4 HNRNPAB LARS UQCRC1 SYNCRIP RCSD1 ZBTB11 CCT8 LETM1 RPS4X SPTBN1 SND1 NIPSNAP1 ATP8A1 GNB2L1 EIF3M SF3B1 ICAM1 BCAP31 RPS3 RPL3 RAB5B ACSL1 CD40 CD19 HVCN1 COL6A5 LGALS9 SLAMF1 ARHGEF2 TAP1 ACOT7 ERO1L CS PRDX1 H3F3B STIP1 ELMO1 PLEC SRPRB HLA-G CYBB INTS8 CRYZ DYNC1H1 HSP90AB1 CCT3 HACE1 SLC3A2 RAD50 RPL18 SPTAN1 HSPE1 RPL32 MTHFD1L HSP90B1 NCL DPP9 ARPC1B NCEH1 DSG2 PMPCB CD300A EIF3C HSPA4 RHOA VARS HSPA8 PSG3 SMC1A SLC1A5 PRKCD COG3 ACAD10 AGPS RPS19 RPL27A LRRC59 HP1BP3 LRMP ARPC4 RPS13 RPL9 KHSRP ARPC3 MYL6 RPSA CYP51A1 TAPBP GSTP1 FBL TARDBP TRIM28 DDX21 RCC2 NOP58 HLA-DQB1 ADAR HADHB PTPN6 PLXNB2 SF3A3 SLC12A2 CD58 CD48 HCFC1 FUBP1 CD70 TMEM8A RAB21 SRC RPS8 CYB5R3 ZW10 TUBA1B PSMC6 FN1 RPL4 CDC37 SLAMF7 EIF4A3 RPS21 ACO2 ATP6V1G2-DDX39B POLR2A COX5B ATP1A1 ALYREF RPL13 IGSF8 RRP12 INPP5D ITGB7 YWHAZ DNAJC7 MAP9 HSPH1 HUWE1 DDX17 RPL10 ACTB TMPO TOMM70A MT-ATP8 ACIN1 ZC3HAV1 RPL23 SRSF2 PCCA GOT2 MTOR ENO1 MCM3 RPS2 RNPS1 CCT4 SLC43A3 DNAH8 MTHFD1 GLS ICAM3 MTHFD2 SCAMP3 TRPV2 TOR1AIP1 ANXA6 FAU RPL13a EWSR1 IFNGR1 RNH1 CR2 CSK VAPB LIMA1 IGHM SRSF9 ENTPD1 PHB2 RPL7 RBMX MX1 CHL1 HIST1H2BN MLEC C1QBP SUB1 GANAB PCBP3 PRPF6 HNRNPA0 HIST2H3PS2 PGAM1 UQCRFS1 PCBP1 RPS27L PDIA6 HLA-DRB1 TCP1 MARCKSL1 AGO2 DLD PAICS ITGB1 CAT LSR LARP1 ICE2 RAB11A HLA-B RPN1 PKM VIM CALR WARS CD84 SNRPD1 FCER2 NONO GPD2 NPM3 TPM2 PCNA RBBP4 HNRNPA3 ELAVL1 FAM120A HLA-A IARS2 NRCAM AP1M1 ATP5H CANX GPANK1 RPL22 LDHA EIF2S1 LCP1 EPRS HSPA6 SCAMP2 SRGAP2 CORO1C PABPC1 ROBO1 MPZL1 CD22 ENKUR RPS26 CD47 SNRPD3 LAP3 TSR1 MYO1G RPL11 NMT1 KRT10 TUBB4B HADHA CCT5 FDXR SEMA4A RPL8 NDUFA4 TALDO1 RBM39 RAB1A CAP1 B2M NT5DC3 NNT STX7 ADAM10 C11orf98 MIF ATP6V1A PFN1 SGOL2 HNRNPK TUBB8 NKX2-6 TRMT44 MARS SLIRP ACTA1 LRPPRC GNB1 CD97 ATP5C1 P4HB FLNB SNX2 WDFY4 PRKDC PDCD1LG2 PDCD6IP COX4I1 SEMA7A EIF3E IPO5 DDX39B CHDH ACLY F11R HMGCS1 HSPA1A RAB10 SCCPDH RPS25 MARS2 MCAM GIGYF2 EXOC7 CD79B PSMA6 TAF15 PDCD2L MCM5 CPT1A KCNAB1 PPIB ACTBL2 RPL17 SCFD1 ACOT9 HIST1H1D VDAC3 PPIA VDAC1 LMNA HNRNPH2 CSE1L TXNDC5 PHB HLA-C CEPT1 KPNB1 GNG2 ETNK1 KHDRBS1 ITGB2 FTSJ3 GAPDH SIT1 MDH2 ERH PSMB8 SEC61B PFKM SLC4A7 LCN1P1 WNK1 SCARB1 SEC22B HIST1H1C FAM120B APEX1 TPR TFRC SCARF2 MAP1S RPL6 TUFM SSBP1 OCIAD1 KPNA6 GCN1L1 CAMK2N1 PDIA4 HIST1H4A AK2 SERBP1 HLA-DRB3 MFI2 RANBP2 SERPINI2 SUPT16H ILF3 TNFRSF8 RPS14 ATP2A2 MTCH2 APBB1IP ATIC GOSR1 GNB2 VTA1 STRN RPL30 CLPP RUVBL2 ACTG1 ATP1B3 TRIM25 DDX3X RPS24 HNRNPL HNRNPU ADRM1 ATXN2L PGD NME1 RPL27 RPL24 FLII RPL38 VCP CYP2J2 CPPED1 RPS16 ALDOA RPS9 CD86 SRSF3 DDX56 FLNA SRRM2 NOTCH1 CD274 RPL14 HNRNPA1 GPS2 EEA1 HGS CCT6A SPATA16 SLC16A1 SFXN3 TNS2 ADA CPZ STAT1 CLTC HLA-DQA1 SRP9 MAT2A HIST2H2AC ARID1A ACSL4 TRAP1 ITGAL CAPZB HIST2H2BE EPHB1 SHMT2 SNW1 NCLN RAN VDAC2 SLC2A1 EEF1G EEF1A1 DACT1 EIF4G1 AGO3 UQCRC2 HNRNPUL2-BSCL2 HNRNPC BSG SF1 FXR1 UBC SAFB HDLBP IDH3A HBA2 HSP90AA1 HCLS1 RAB7A SRRM1 ACTR2 PABPN1

IGHV3-66 ZNF429 ZNF33B YWHAB LBR PARP1 RAB8A PGK1 DHX9 PTPRC CCT7 SRY MYO1E KRT2 H2AFY ITGA

3.2 N-linked Glycoprotein Enrichment Results

3.2.1 Mass Spectrometry Data Analysis of Surface Protein of the CCRF-SB Cell Line using N-Linked Glycopeptide enrichment.

For the surface glycoprotein enrichment cells were collected and were incubated with 2 mM sodium periodate (NaIO4) in PBS for 30 minutes to be able to oxidize glycoproteins. Cells were lysed and proteins coupled to hydrazide agarose beads. For the control group, glycoproteins were not oxidized to check the nonspecific coupling or adsorption of proteins to hydrazide beads. Glycoenriched and control samples were eluted from the beads with PNGase treatment and analyzed with mass spectrometry.

For the control groups, 57 proteins were identified, after removal of common contaminants proteins added by accident through dust or physical contact no protein was left in the list. From the N-linked glycol-enriched samples 218 proteins were identified and listed in Table 3.5.

Later on, detected proteins were analyzed with Gene Ontology Cellular Component analysis. The analysis showed that 155 of the detected proteins were annotated with ‘Membrane’, 132 of those annotated with ‘Extracellular Space’ (Table 3.4). Cellular Component analysis showed that the presence of mitochondrial and ribosomal proteins could be the result of nonspecific enrichment through protein-protein interactions.

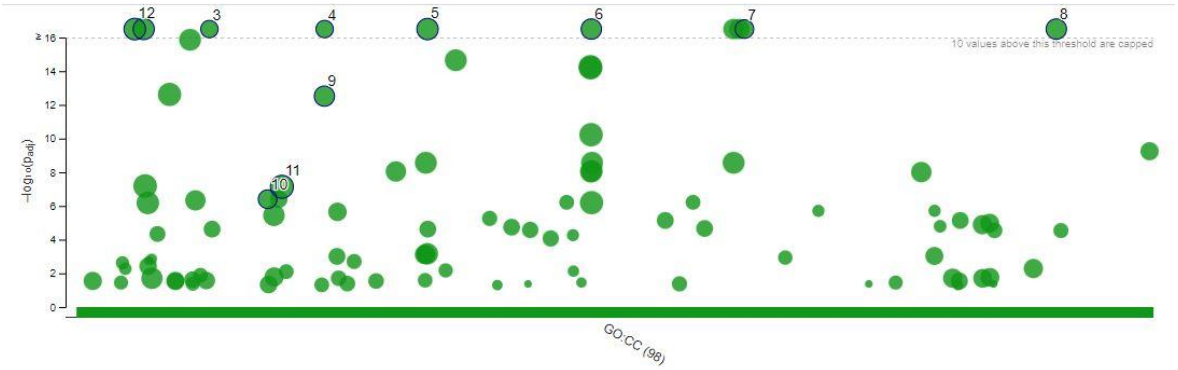
Table 3.3 Gene Ontology Cellular Component analysis of identified proteins from the MS data with Max-Quant.

#termID	Term description	Observed gene count	Strength	FDR
GO:0016020	Membrane	155	0.19	5.65e-12
GO:0005615	Extracellular Space	132	0.57	4.35e-46
	Vesicle	143	0.52	5.08e-46
GO:0071944	Cell periphery	81	0.13	0.0214
GO:0005886	Plasma membrane	78	0.12	0.0389
GO:0098796	Membrane protein complex	32	0.4	6.14e-05
GO:0009986	Cell surface	33	0.56	2.35e-08
GO:0005759	Mitochondrial protein complex	15	0.67	0.0077
GO:0005840	Ribosome	11	0.45	0.0011

Table 3.4 List of the identified proteins with glycoprotein enrichment.

Identified Proteins
ACTB HSPD1 CD44 ANXA2 PSMA2 TXN RPS28 MAP4 ENO1 CTSH TIPRL FABP5 MCM3 HYOU1 TUBB CFL1 TKT YWHAQ RTN4 PRKCSH LAMTOR4 CD82 PTMA IQCD EIF6 AIFM1 SYNPO RPLP1 MTHFD1 ICAM3 PRDX5 LMNA PPIA CSE1L HNRNPH1 PHB HLA-DPB1 RPS12 ANXA6 AGFG1 MYH9 CLTB TLN1 ITGB2 EEF2 OTUB1 HNRNPM GAPDH KRT9 HIST1H1B ATP5B PHB2 MDH2 RBMX H2AFJ MX1 XRCC6 NPM1 CORO1A EEF1B2 HLA-DRA SUB1 RPLP0 HIST1H1C UQCRFS1 PCBP1 ARCN1 TMSB4X U2AF2 MS4A1 DLST ATP5A1 PDIA6 TUFM HLA-DRB1 GCN1L1 PDIA4 SEMA4D HIST1H4A BAK1 FAM134C SUN2 ICAM1 MARCKSL1 LMNB1 ATP5D CD40 TOMM22 PARK7 MSN IPO9 ATIC PSMC4 LSP1 H2AFX HLA-B ACADM DAG1 PKM VIM CALR WARS GORASP2 ACTG1 CACYBP UBA1 RPLP2 CPSF7 PRDX1 NUP62 STIP1 PUF60 SETMAR BST2 HNRNPU HNRNPL ELMO1 CHCHD2 CYFIP2 PCNA TPM4 HLA-A ALDOA CANX UBAP2L LDHA LCP1 SRSF3 CALM2 FLNA LAMP1 NCL CAND1 PEBP1 ARPC1B BAX U2AF1L4 FDPS NPTN PPP4R4 KRT1 GOLGB1 PDLIM1 PSMA5 HNRNPA2B1 CD47 LAP3 CLTA LPXN PDIA3 EPB41L2 NHP2L1 CLTC RAP1B ARHGAP27 KHSRP HADHA ARPC3 MYL6 RPSA NUP93 CAP1 RALGAPA1 PTBP1 HLA-DQB1 PTPN6 CHMP5 STOML2 CER1 EEF1G PFN1 EEF1A1 CD48 FUBP1 CD70 HNRNPK MAST1 PCBP2 TUBA1B UQCRC2 PSMC2 HNRNPC LSM5 BSG SOD1 CCDC129 CDC37 ATP5C1 FNBP1 FSCN1 RPS21 P4HB LGALS1

PRKDC TPI1 NOC2L PARP1 YWHAZ ITGB7 HIST2H2BF SEMA7A PGK1 SLC25A5 DHX9 PTPRC IPO5 PRRC2C CCT7 F11R SNRPD2 DCD NACA



ID	Source	Term ID	Term Name	p_{adj} (query_1)
1	GO:CC	GO:0005576	extracellular region	4.190×10^{-29}
2	GO:CC	GO:0005615	extracellular space	1.259×10^{-37}
3	GO:CC	GO:0005925	focal adhesion	2.213×10^{-24}
4	GO:CC	GO:0030055	cell-substrate junction	4.177×10^{-24}
5	GO:CC	GO:0031982	vesicle	2.760×10^{-41}
6	GO:CC	GO:0043230	extracellular organelle	1.273×10^{-55}
7	GO:CC	GO:0070161	anchoring junction	6.950×10^{-21}
8	GO:CC	GO:1903561	extracellular vesicle	1.154×10^{-55}
9	GO:CC	GO:0030054	cell junction	3.034×10^{-13}
10	GO:CC	GO:0009986	cell surface	3.971×10^{-7}
11	GO:CC	GO:0016020	membrane	7.275×10^{-8}

Figure 3.4 Illustration of Gene Ontology Cellular Component analysis with gProfiler. Data annotated with extracellular space and associated networks.

3.3 Comparison of Surface Biotinylation vs N-linked Glycoprotein Enrichment

We next compared the result of mass spectrometry analysis of surface biotinylation and N-linked glycoprotein enrichment approaches. Among these, the biotin-labeling approach yielded better protein enrichment. Although a much lower number of proteins was detected with the N-linked glycoprotein enrichment method, 80 of those proteins were detected specifically with glycoprotein enrichment. This indicates that the glycoprotein enrichment approach could potentially enrich different sets of proteins than biotinylation. Once we compared the cellular component analysis of these two methods, an increase in the plasma membrane enrichment with glycoprotein enrichment is observed. This indicates that the glycoprotein enrichment method is much more specific compared to biotinylation. However, biotin labeling was much more successful for the identification of proteins in our setup.

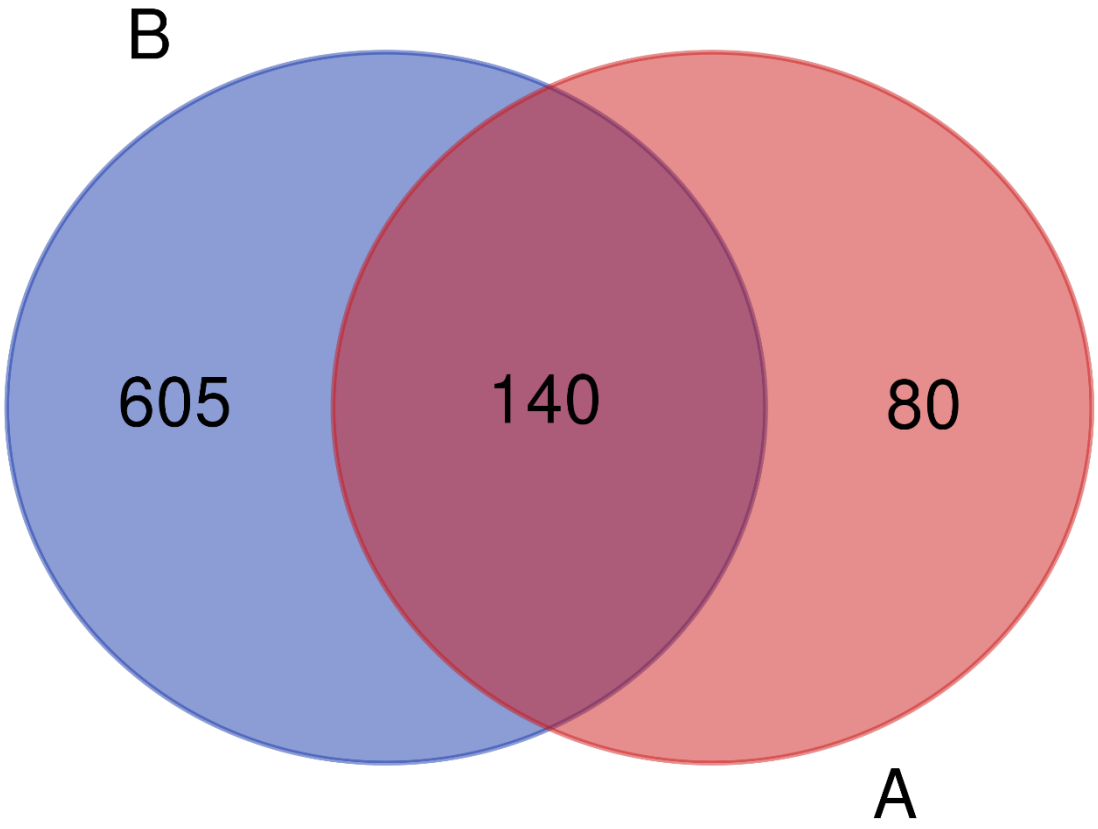
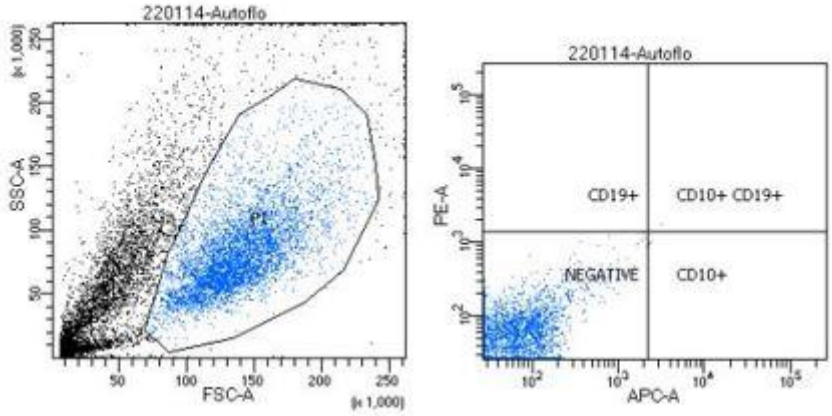


Figure 3. 5 Venn Diagram comparison of the biotinylation and N-linked Glycosylated. Group A represents protein numbers identified with glycoprotein enrichment. Group B represents the protein number identified with the biotinylation method.

3.4 Flow Cytometry Verification of Proteomics Data

Surface CD19 and CD10 together with some other antigens used for discrimination between B-ALL subtypes. CD19 was present in our proteomics data, while CD10 was absent. Since the surface markers of CCRF-SB cell lines were not characterized previously, to verify our proteomics data, CCRF-SB cells were labeled with CD10 and CD19 antibodies. Unlabeled cells were analyzed to determine the autofluorescence of the cells (Figure 3.10).

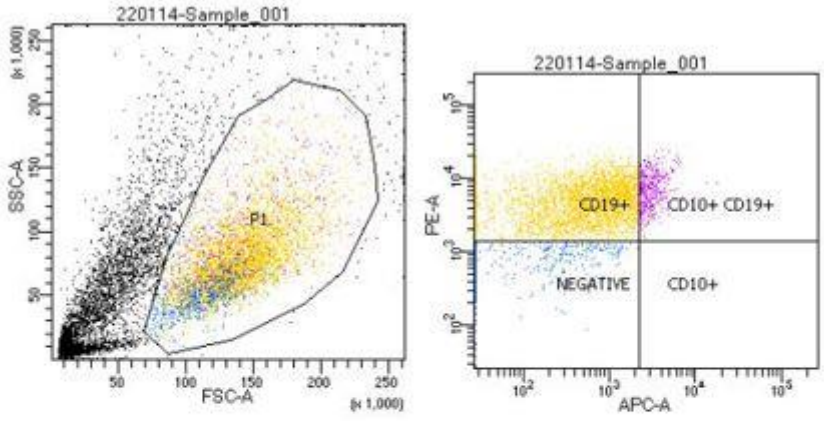
Figure 3.11 showed that CCRF-SB cells are expressing CD19 surface marker with an 81.3% rate. Also, flow cytometry analysis showed that 6.9% of the cells are double-positive CD10 and CD19. However, CD10 expressing cells was shown as 0 %. In the whole population, the percentage of CD10+ cells was low, this explains why we cannot identify CD10 on the surface of CCRF-SB cells.



Tube: Autoflo

Population	#Events	%Parent	%Total
All Events	10,000	####	100.0
P1	4,784	47.8	47.8
CD19+	0	0.0	0.0
CD10+ CD19+	1	0.0	0.0
NEGATIVE	4,781	99.9	47.8
CD10+	2	0.0	0.0

Figure 3. 6 Flow cytometry analysis of the unlabeled CCRF-SB cells.



Tube: Sample_001

Population	#Events	%Parent	%Total
All Events	10,000	####	100.0
P1	5,350	53.5	53.5
CD19+	4,352	81.3	43.5
CD10+ CD19+	367	6.9	3.7
NEGATIVE	631	11.8	6.3
CD10+	0	0.0	0.0

Figure 3. 7 Flow Cytometry analysis of the immunofluorescent stained CCRF-SB cells

Chapter 4

Discussion

This thesis study aimed to map the membrane of the CCRF-SB cells (B-cell acute lymphoblastic cell line) using mass spectrometry-based proteomics. Since the surface proteins are less represented in whole-cell extracts, cell surface proteins are enriched using two different approaches: surface biotinylation and N-linked glycoenrichment.

Cell surface biotinylation technique depends on the labeling of surface proteins using membrane-impermeable reactive biotin (NSH-SS-biotin). The surface biotinylation method was optimized using a fluorescent microscope and flow cytometry analysis. Both methods confirmed biotinylation of cell surface proteins using Streptavidin-488 staining. Biotinylated proteins were captured with streptavidin agarose beads and after washing, proteins were prepared for mass spectrometry analysis using on bead digestion protocol. After LC-MS/MS analysis, biotin captured proteins were identified. In our data, proteins that are achieved from mass spectrometry are analyzed with Max-Quant, and 229 proteins were identified with a 1% FDR rate. Gene Ontology Cellular Component analysis showed that 467 of the identified proteins are annotated as 'Membrane', 263 of those annotated as 'Extracellular Space', 235 of those annotated as 'Plasma membrane' 77 of those annotated as 'Membrane protein complex', and 66 of those annotated as 'Cell surface' (Table 3.1). Surface biotinylation was previously applied to enrich the surface of the different types of cells. Küçük and colleagues biotinylated the carboxyl groups of the HeLa S3 cell line. With this approach, they identified almost on average 450 proteins. But identified proteins are not fully annotated with membrane, extracellular region, or cell surface. Identified protein list also contains cytoplasmic and organelle-based contamination [94]. In another study Karhemo et al., they applied surface biotinylation on MDA-MB-435 cell lines and they identified 86 proteins and 60% of the isolated proteins annotated with cell surface [95]. Compared to previous studies on the number of identified proteins, our approach was successful. However not all identified proteins belong to the surface proteins and the number was less when compared to the mentioned studies. One of the reasons for this result could be cytoplasmic and organelle-based

contamination. These contaminants were thought to originate from dead cells that were disrupted during the washing steps. For the attached cells, it is easier to clean dead cells and remaining proteins with simple washing thereby the possibility of intracellular contamination is low. For the suspended cells, several centrifugation steps are achieved to get rid of the unwanted solutions and chemicals. These washings may cause cell bursts and causes cytosolic and organelle-based protein contamination. To reduce this contamination, a lower centrifugation speed can be considered. Lowering the centrifugation speed may cause cell loss and centrifugation speed is also important to get rid of the dead cells' residues. During the centrifugation, healthy cells will settle to the bottom and burst cells and remaining particles will be in the liquid part and will be discarded. However, after centrifugation cell pellet should be suspended again. During this step, even the gentle pipetting may cause cellular disruption. Comparing the abundance of cytoplasmic and surface proteins, cellular disruption can adversely affect the enrichment of surface proteins. To reduce these troubles, cells may be centrifuged with a supportive layer such as filter-based washing systems. These types of approaches can decrease the cellular burst and reduce intracellular protein contamination.

With the N-linked glycoprotein enrichment method, 229 proteins were identified with Max Quant Gene Ontology Cellular Component analysis showed that 155 of the identified proteins are annotated as 'Membrane', 132 of those annotated as 'Extracellular Space', 78 of those annotated as 'Plasma membrane' 32 of those annotated as 'Membrane protein complex', and 33 of those annotated as 'Cell surface' (Table 3.5). Sun and colleagues have analyzed the surface glycoproteins of the CML cell lines. They have identified 180 surface glycoproteins such as CD156C, CD54, BSG which are also identified in our data [91]. Hofmann and colleagues have analyzed the surface proteins of acute myeloid leukemia cell lines (HL60 and NB4) with the Cell Surface Capturing (CSC) technology. In that study, surface glycoproteins are oxidized and labeled with biocytin hydrazide and isolated using streptavidin agarose beads. With this method, they were able to identify CD45, CD11c, and CD71 protein respectively [96].

When we compare the data obtained from this thesis study to previously published studies, we identified less number of surface glycoproteins. There could be several reasons for the low number of protein identification. Firstly, during our glycoprotein enrichment studies, we had several technical issues. The methods we tried and applied depended on the enrichment of glycopeptides, not the glycoproteins. Enrichment of the glycoproteins directly was not successful due to the precipitation of proteins during

downstream enrichment steps. The major disadvantage of glycopeptide enrichment is that the protein identifications depend on the detection of that specific glycopeptide with the MS instrument. Considering the sizes and charges of those peptides, it is possible that some of those glycopeptides cannot be detected and identified using LC-MS/MS analysis. To overcome this issue, it is required to optimize glycoprotein enrichment. If the glycoproteins could be enriched successfully, from a protein many peptides could be obtained and detected. For example, CD19 which is a marker protein was detected on the surface of CCRF-SB cells using surface biotinylation and flow cytometry analysis. Although N86 residue of CD19 is N-linked glycosylated, it was not listed in our glycopeptide analysis data. Once we analyzed possible tryptic peptides from the CD19 sequence, the N86 glycosylation site was too big to be identified with the standard MS and MS/MS data acquisition protocol. Therefore, the glycoprotein enrichment technique needs to be optimized. For the digestion using only trypsin may cause the elimination of the peptide at the MS1 analyzing stage. Surface proteins, which are few in number compared to intracellular proteins, will be even less in number in the event of cellular disruption and are even less likely to be captured during the enrichment. Chen and colleagues, clearly showed that using multiple digestive enzymes is significantly increasing the N-glycoprotein coverage for the MS analysis [97]. Secondly, PNGase F, a glycosidase is commonly employed to liberate N-glycans from the peptide backbone, resulting in a peptide in which the asparagine, the residue is deamidated and transformed to aspartic acid. However, the PNGase F deglycosylation may affect the determination of the N-glycosylated proteins due to long incubation times, incomplete deglycosylation, and spontaneous nonenzymatic deamidation of asparagine residues [98].

Chapter 5

Conclusion and Future Prospects

5.1 Conclusion

B-ALL is a fast-growing and aggressive type of leukemia. It is characterized by malignant B cell accumulation. It is compromised 80% of the childhood leukemias. Because of this high rate, it is important to generation of new diagnostic and treatment approaches. In the light of this information surface, proteins are important molecules for these studies because of their easily accessible and cancer representative features. Any potential cancer biomarker could be a candidate for diagnostic, treatment, or both approaches. The fact remains that, their low abundance level, hydrophobicities, and highly dynamic range cause underrepresentation in the proteomic data and it makes challenging the study on surface protein mapping. Therefore, to map the surfaceome of the suspended cells, it is required to use two different methods together. In addition to all the challenging sites of the surface proteomic, our methods need to be improved.

We have applied two different surface protein enrichment techniques. Both approaches have advantages and drawbacks. Enrichment of glycosylated proteins is a very efficient strategy for analyzing surface proteins considering 80% of them are glycosylated [92]. But considering their less abundant level compared to an intracellular protein, the remaining ~20% non-glycosylated membrane proteins are a huge loss for surface proteomics studies. Also, the glycoprotein enrichment method includes several steps like oxidation, coupling, enzyme elution, this approach requires more optimization. Taking consideration of this, another enrichment approach, biotinylation has been chosen. For the biotinylation strategy, un-accessible primary amines can be considered as drawbacks however, since their inaccessibility may be due to heavy N-linked modifications, the glycol-enrichment method will also be able to cover the disadvantages at this point. Identification of the enriched surface protein can result in a list of proteins that would include possible candidates for diagnostic or therapeutic approaches. Comparison of the two methods showed that each method has the ability to identify

different surface proteins subsets. The combined use of the two approaches for the identification of glycosylated and non-glycosylated surface proteins can boost yield due to their complementing features.

5.2 Social Impact and Contribution to Global Sustainability

Leukemia is a common type of cancer. B-ALL covers 80% of the leukemia types seen in children. Currently, there are different diagnostic and treatment methods that could be used in all the cases sufficiently. Because of the high incidence and challenges in diagnosis and treatment, leukemia is a serious health problem and global economic burden. This thesis, it was aimed to determine the potential surface biomarkers for diagnosis and treatment approaches.

Examination of surface biomarkers is of great importance for the early detection of the disease and the development of alternative treatment methods. These markers are expected to contribute to the fight against leukemia, which threatens public health and is very expensive to treat, by providing solutions to combat leukemia. If contributions are discussed under a different title, finding new biomarkers will contribute the society in terms of health and economy. For example, at the detection stage, diagnosing cancer in the early stages will help reduce the amount of medication a patient would be exposed to if the cancer was diagnosed in later stages. Also, early detection will help reduce the burden on health systems. In the therapy stage, it is known that cancer therapies are complete in long term and follow the same treatment path for different patients. In these treatment paths, patients are exposed to high medication. Finding specific surface markers will help to the improvement of personalized medicine and lower the high medication exposure [99]. Lowering the medicine usage is an important approach for improving the patients' life quality especially when it is considered B-ALL is a pediatric disease. Low medication usage may also contribute to the economy. The identification of potential biomarkers will lead to the development of high value-added diagnostic kits and the development of new generation treatment methods (such as smart drugs, CAR-T cell therapy).

5.3 Future Prospect

ALL is the most frequent kind of leukemia in children, and B-ALL is the most common type of ALL, accounting for 80% of cases. Overexpression of B malignant cells is a hallmark of B-ALL. Depending on the patient's age and a number of distinct genetic patterns, current therapy regimens take different pathways. On this path, proteins have become more appealing to research using various methodologies and for various objectives as a result of their significance in cancer. The use of surface biomarkers in the early detection of disease and the development of alternative treatment techniques is critical. In the light of this information, our protein list which includes surface biomarkers that are already in use is a promising list to show possible potential biomarkers. However, in order to reflect a protein as a biomarker comparison of CCRF-SB protein data with the B-ALL patient sample should be done. Also, patient and cell line-based protein data should be compared to healthy tissues to understand protein targeting probabilities.

To improve the number of identified proteins to map the surface proteins, the methods need to be optimized. First, instead of trypsin which is used for the digestive enzyme in this thesis study, other proteases need to be applied. Trypsin digest the peptides from their arginine and lysine residues. However, cell surface proteins are mostly embedded in the lipid bilayer and consist of non-polar (hydrophobic) amino acids. Hence, using only trypsin for protein digestion may end up with longer peptides than mass spectrometry can analyze [100]. To eliminate this possibility, another batch digestion enzyme such as chymotrypsin can be used.

Secondly, in this thesis study as an MS method, DDA was used where peptides are selected depending on their intense values. In the case of our study even if the peptides are enriched it is possible to lose the data from the MS1 stage because of their low intensities. More comprehensive methods such as SWATH analysis could be preferred for the low abundant protein studies like surface proteins or related parameters could be arranged for DDA analysis.

Lastly, the identification of peptide sequences by mass spectrometry (MS) relies on the quality of the protein sequence database. The peptides, analyzed by the instrument cannot be identified if it is missing in the databases. Considering the number of mutations and protein variants in cancer, the use of standard human protein databases will fail the identification of tumor-specific peptides and proteins which are products of mutations,

indels, splice isoforms, etc. Using proteogenomic tools, we want to utilize transcriptome data to create tumor-specific custom databases [101, 102].

BIBLIOGRAPHY

[1] B. Van Ness, "Applications and limitations in translating genomics to clinical practice," *Transl. Res.*, no. May, 2015.

[2] D. Bausch-Fluck *et al.*, "A mass spectrometric-derived cell surface protein atlas," *PLoS One*, vol. 10, no. 4, pp. 1–22, 2015.

[3] R. L. Siegel, K. D. Miller, and A. Jemal, "Cancer statistics, 2018," *CA. Cancer J. Clin.*, vol. 68, no. 1, pp. 7–30, 2018.

[4] T. Terwilliger and M. Abdul-Hay, "Acute lymphoblastic leukemia: a comprehensive review and 2017 update," *Blood Cancer J.*, vol. 7, no. 6, p. e577, 2017.

[5] C. Ariöz, "Exploring the interplay of lipids and membrane proteins," 2014.

[6] D. Hanahan and R. A. Weinberg, "Hallmarks of cancer: The next generation," *Cell*, vol. 144, no. 5, pp. 646–674, Mar. 2011.

[7] M. B. Sporn and A. B. Roberts, "Autocrine growth factors and cancer," *Nat. 1985 3136005*, vol. 313, no. 6005, pp. 745–747, 1985.

[8] S. Meshinchi and F. R. Appelbaum, "Structural and Functional Alterations of FLT3 in Acute Myeloid Leukemia," *Clin. Cancer Res.*, vol. 15, no. 13, pp. 4263–4269, Jul. 2009.

[9] R. Hong, Y. Hu, and H. Huang, "Biomarkers for Chimeric Antigen Receptor T Cell Therapy in Acute Lymphoblastic Leukemia: Prospects for Personalized Management and Prognostic Prediction," *Front. Immunol.*, vol. 12, p. 210, Feb. 2021.

[10] R. C. Sterner and R. M. Sterner, "CAR-T cell therapy: current limitations and potential strategies," *Blood Cancer J. 2021 114*, vol. 11, no. 4, pp. 1–11, Apr. 2021.

[11] N. Coccaro, L. Anelli, A. Zagaria, G. Specchia, and F. Albano, "Next-generation sequencing in acute lymphoblastic Leukemia," *Int. J. Mol. Sci.*, vol. 20, no. 12, 2019.

[12] S. Chiaretti, G. Zini, and R. Bassan, "Diagnosis and subclassification of acute lymphoblastic leukemia," *Mediterr. J. Hematol. Infect. Dis.*, vol. 6, no. 1, 2014.

[13] K. Mrózek, D. P. Harper, and P. D. Aplan, "Cytogenetics and Molecular Genetics of Acute Lymphoblastic Leukemia," *Hematol. Oncol. Clin. North Am.*, vol. 23, no. 5, pp. 991–1010, 2009.

[14] J. S. Woo, M. O. Alberti, and C. A. Tirado, "Childhood B-acute lymphoblastic leukemia: A genetic update," *Exp. Hematol. Oncol.*, vol. 3, no. 1, pp. 1–14, 2014.

[15] C. G. Mullighan *et al.*, "Genome-wide analysis of genetic alterations in acute lymphoblastic leukaemia," *Nature*, vol. 446, no. 7137, pp. 758–764, 2007.

[16] Amriah Buang, "WHO Klassifikation 2016," *Blood*, vol. 2, no. 20, pp. 58–71, 2006.

[17] Z. Jiang, D. Wu, S. Lin, and P. Li, "CD34 and CD38 are prognostic biomarkers for acute B lymphoblastic leukemia," *Biomark. Res.*, vol. 4, no. 1, pp. 80–83, 2016.

[18] M. L. Den Boer *et al.*, "A subtype of childhood acute lymphoblastic leukaemia with poor treatment outcome: a genome-wide classification study," *Lancet Oncol.*, vol. 10, no. 2, pp. 125–134, 2009.

[19] M. Maďarová, D. Dobransky, and T. Dobransky, "New Protein Markers of Chronic Lymphocytic and Acute Lymphocytic Leukemia," in *Advances in Hematologic Malignancies*, IntechOpen, 2019.

[20] C. H. Pui *et al.*, "Childhood acute lymphoblastic leukemia: Progress through

collaboration," *J. Clin. Oncol.*, vol. 33, no. 27, pp. 2938–2948, 2015.
[21] N. Gökbuget and D. Hoelzer, "Recent approaches in acute lymphoblastic leukemia in adults," *Reviews in Clinical and Experimental Hematology*, vol. 6, no. 2. pp. 114–141, Jun-2002.
[22] C. Gao *et al.*, "Detecting Drug-Resistant PDAC Detection of Chemotherapy-Resistant Pancreatic Cancer Using a Glycan Biomarker."
[23] Y. Zhou *et al.*, "Advances in the molecular pathobiology of B-lymphoblastic leukemia," *Hum. Pathol.*, vol. 43, no. 9, pp. 1347–1362, 2012.
[24] T. Zuckerman and J. M. Rowe, "Pathogenesis and prognostication in acute lymphoblastic leukemia," *F1000Prime Rep.*, vol. 6, no. July, pp. 3–7, 2014.
[25] C. Malouf and K. Ottersbach, "Molecular processes involved in B cell acute lymphoblastic leukaemia," *Cell. Mol. Life Sci.*, vol. 75, no. 3, pp. 417–446, 2018.
[26] M. Kato and A. Manabe, "Treatment and biology of pediatric acute lymphoblastic leukemia," *Pediatr. Int.*, vol. 60, no. 1, pp. 4–12, 2018.
[27] L. Kuhlmann, E. Cummins, I. Samudio, and T. Kislinger, "Cell-surface proteomics for the identification of novel therapeutic targets in cancer," *Expert Rev. Proteomics*, vol. 15, no. 3, pp. 259–275, 2018.
[28] Y. Z. Zheng and L. J. Foster, "Biochemical and proteomic approaches for the study of membrane microdomains," *J. Proteomics*, vol. 72, no. 1, pp. 12–22, 2009.
[29] D. Vuckovic, L. F. Dagley, A. W. Purcell, and A. Emili, "Membrane proteomics by high performance liquid chromatography-tandem mass spectrometry: Analytical approaches and challenges," *Proteomics*, vol. 13, no. 3–4, pp. 404–423, 2013.
[30] H. Lodish, A. Berk, S. L. Zipursky, P. Matsudaira, D. Baltimore, and J. Darnell, "Membrane Proteins," 2000.
[31] B. A. Macher and T. Y. Yen, "Proteins at membrane surfaces - A review of approaches," *Mol. Biosyst.*, vol. 3, no. 10, pp. 705–713, 2007.
[32] S. Tan, H. T. Tan, and M. C. M. Chung, "Membrane proteins and membrane proteomics," pp. 3924–3932, 2008.
[33] K. Chou and D. W. Elrod, "Prediction of Membrane Protein Types and Subcellular Locations," vol. 153, no. June 1998, pp. 137–153, 1999.
[34] K. B. Chandler and C. E. Costello, "Glycomics and glycoproteomics of membrane proteins and cell-surface receptors : Present trends and future," pp. 1407–1419, 2016.
[35] R. S. Hegde and R. J. Keenan, "The mechanisms of integral membrane protein biogenesis," *Nat. Rev. Mol. Cell Biol.*, 2021.
[36] Y. Peng, J. Li, and L. Zhu, *Cancer and non-coding RNAs*. Elsevier, 2019.
[37] S. I. Grivennikov, F. R. Greten, and M. Karin, "Immunity, Inflammation, and Cancer," *Cell*, vol. 140, no. 6, pp. 883–899, Mar. 2010.
[38] P. Janovska *et al.*, "Autocrine signaling by Wnt-5a deregulates chemotaxis of leukemic cells and predicts clinical outcome in chronic lymphocytic leukemia," *Clin. Cancer Res.*, vol. 22, no. 2, p. 459, Jan. 2016.
[39] J. M. Boag *et al.*, "Altered glucose metabolism in childhood pre-B acute lymphoblastic leukaemia," *Leuk. 2006 2010*, vol. 20, no. 10, pp. 1731–1737, Aug. 2006.
[40] R. A. Gatenby and R. J. Gillies, "Why do cancers have high aerobic glycolysis?," *Nat. Rev. Cancer 2004 411*, vol. 4, no. 11, pp. 891–899, Nov. 2004.
[41] A. Lorentzian, A. Uzozie, and P. F. Lange, "Origins and clinical relevance of proteoforms in pediatric malignancies," *https://doi.org/10.1080/14789450.2019.1575206*, vol. 16, no. 3, pp. 185–200,

Mar. 2019.
[42] Y. Shang and F. Zhou, “Current Advances in Immunotherapy for Acute Leukemia: An Overview of Antibody, Chimeric Antigen Receptor, Immune Checkpoint, and Natural Killer,” *Front. Oncol.*, vol. 9, p. 917, Sep. 2019.
[43] E. Dotan, C. Aggarwal, and M. R. Smith, “Impact of Rituximab (Rituxan) on the Treatment of B-Cell Non-Hodgkin’s Lymphoma,” *Pharm. Ther.*, vol. 35, no. 3, p. 148, Mar. 2010.
[44] H. M. Kantarjian *et al.*, “Inotuzumab Ozogamicin versus Standard Therapy for Acute Lymphoblastic Leukemia,” *N. Engl. J. Med.*, vol. 375, no. 8, pp. 740–753, Aug. 2016.
[45] N. N. Shah *et al.*, “Bispecific anti-CD20, anti-CD19 CAR T cells for relapsed B cell malignancies: a phase 1 dose escalation and expansion trial,” *Nat. Med. 2020 2610*, vol. 26, no. 10, pp. 1569–1575, Oct. 2020.
[46] K. Sharma *et al.*, “Phase II study of alemtuzumab (CAMPATH-1®) in patients with HTLV-1-associated adult T-cell leukemia/lymphoma,” *Clin. Cancer Res.*, vol. 23, no. 1, p. 35, Jan. 2017.
[47] K. Tsukasaki and K. Tobinai, “Clinical Trials and Treatment of ATL,” *Leuk. Res. Treatment*, vol. 2012, pp. 1–12, Jan. 2012.
[48] C. G. Gahmberg and M. Tolvanen, “Why mammalian cell surface proteins are glycoproteins,” *Trends Biochem. Sci.*, vol. 21, no. 8, pp. 308–311, 1996.
[49] A. Illiano, G. Pinto, C. Melchiorre, A. Carpentieri, V. Faraco, and A. Amoresano, “Protein Glycosylation Investigated by Mass Spectrometry: An Overview,” *Cells*, vol. 9, no. 9, Aug. 2020.
[50] C. Reily, T. J. Stewart, M. B. Renfrow, and J. Novak, “Glycosylation in health and disease,” *Nature Reviews Nephrology*, vol. 15, no. 6. Nature Publishing Group, pp. 346–366, 01-Jun-2019.
[51] P. Stanley, N. Taniguchi, and M. Aebi, “N-Glycans,” *Essentials Glycobiol.*, 2017.
[52] P. P. Hsu and D. M. Sabatini, “Cancer Cell Metabolism: Warburg and Beyond,” *Cell*, vol. 134, no. 5, pp. 703–707, Sep. 2008.
[53] O. Warburg, “On the origin of cancer cells,” *Science (80-.).*, vol. 123, no. 3191, pp. 309–314, Feb. 1956.
[54] A. Schietinger *et al.*, “Supporting Online Material A Mutant Chaperone Converts a Wild-Type Protein into a Tumor-Specific Antigen,” *13. I. Gusarov, E. Nudler, Mol. Cell*, vol. 5, p. 90, 2004.
[55] S. Kellokumpu, R. Sormunen, and I. Kellokumpu, “Abnormal glycosylation and altered Golgi structure in colorectal cancer: dependence on intra-Golgi pH.”
[56] S. Julien, M. Bobowski, A. Steenackers, X. Le Bourhis, and P. Delannoy, “How Do Gangliosides Regulate RTKs Signaling?,” *Cells*, vol. 2, no. 4, pp. 751–767, Dec. 2013.
[57] E. Gorelik, U. Galili, and A. Raz, “On the role of cell surface carbohydrates and their binding proteins (lectins) in tumor metastasis,” *Cancer Metastasis Rev.*, vol. 20, no. 3–4, pp. 245–277, 2001.
[58] A. Magalhães, H. O. Duarte, and C. A. Reis, “Aberrant Glycosylation in Cancer: A Novel Molecular Mechanism Controlling Metastasis,” *Cancer Cell*, vol. 31, no. 6, pp. 733–735, Jun. 2017.
[59] K. J. Mccann *et al.*, “Remarkable selective glycosylation of the immunoglobulin variable region in follicular lymphoma,” 2007.
[60] K. H. Yu and M. Snyder, “Omics Profiling in Precision Oncology,” *Mol. Cell. Proteomics*, vol. 15, no. 8, pp. 2525–2536, Aug. 2016.
[61] M. R. Wilkins *et al.*, “Progress with Proteome Projects: Why all Proteins

Expressed by a Genome Should be Identified and How To Do It," *Biotechnol. Genet. Eng. Rev.*, vol. 13, no. 1, pp. 19–50, 1996.
[62] R. Aebersold and M. Mann, "Mass spectrometry-based proteomics," *Nat. 2003 4226928*, vol. 422, no. 6928, pp. 198–207, Mar. 2003.
[63] C. Ho *et al.*, "Electrospray Ionisation Mass Spectrometry: Principles and Clinical Applications," *Clin. Biochem. Rev.*, vol. 24, no. 1, p. 3, 2003.
[64] P. Hö, C. A. Miller, T. Preckel, and C. Wenz, "Efficient Fractionation and Improved Protein Identification by Peptide OFFGEL Electrophoresis* □ S," 2006.
[65] B. Manadas, V. M. Mendes, J. English, and M. J. Dunn, "Peptide fractionation in proteomics approaches," *Expert Rev. Proteomics*, vol. 7, no. 5, pp. 655–663, 2010.
[66] S. Saha, S. Walia, K. Sharma, and K. Banerjee, "Suitability of stationary phase for LC analysis of biomolecules," *Crit. Rev. Food Sci. Nutr.*, vol. 60, no. 17, pp. 2856–2873, 2020.
[67] K. K. Murray, R. K. Boyd, M. N. Eberlin, G. John Langley, L. Li, and Y. Naito, "Definitions of terms relating to mass spectrometry (IUPAC Recommendations 2013)," *Pure Appl. Chem.*, vol. 85, no. 7, pp. 1515–1609, Jun. 2013.
[68] J. V. Olsen, B. Macek, O. Lange, A. Makarov, S. Horning, and M. Mann, "Higher-energy C-trap dissociation for peptide modification analysis," *Nat. Methods 2007 49*, vol. 4, no. 9, pp. 709–712, Aug. 2007.
[69] J. Mitchell Wells and S. A. McLuckey, "Collision-Induced Dissociation (CID) of Peptides and Proteins," *Methods Enzymol.*, vol. 402, pp. 148–185, Jan. 2005.
[70] A. Guthals and N. Bandeira, "Peptide Identification by Tandem Mass Spectrometry with Alternate Fragmentation Modes* □ S," 2012.
[71] G. Hart-Smith, "A review of electron-capture and electron-transfer dissociation tandem mass spectrometry in polymer chemistry," *Anal. Chim. Acta*, vol. 808, pp. 44–55, 2014.
[72] D. M. Good, M. Wirtala, G. C. Mcalister, and J. J. Coon, "Performance Characteristics of Electron Transfer Dissociation Mass Spectrometry* □ S," 2007.
[73] C. S. Su, "A new time-of-flight mass spectrometer," *Nucl. Instruments Methods Phys. Res.*, vol. 220, no. 2–3, pp. 431–432, 1984.
[74] R. E. March, "Quadrupole ion traps," *Mass Spectrom. Rev.*, vol. 28, no. 6, pp. 961–989, Nov. 2009.
[75] "TripleTOF ™ 5600 System Maintenance," no. January, pp. 2–5, 2011.
[76] X. Li *et al.*, "Ion Trap Array Mass Analyzer: Structure and Performance," *Rapid Commun. Mass Spectrom*, vol. 12, no. 2, pp. 4840–4846, 2001.
[77] Q. Hu, R. J. Noll, H. Li, A. Makarov, M. Hardman, and R. G. Cooks, "The Orbitrap: a new mass spectrometer," *J. Mass Spectrom.*, vol. 40, no. 4, pp. 430–443, Apr. 2005.
[78] Z. R. Gregorich and Y. Ge, "Top-down Proteomics in Health and Disease: Challenges and Opportunities," *Proteomics*, vol. 14, no. 10, p. 1195, 2014.
[79] N. L. Kelleher, "PROTEOMICS."
[80] J. R. Yates, C. I. Ruse, and A. Nakorchevsky, "Proteomics by Mass Spectrometry: Approaches, Advances, and Applications," *http://dx.doi.org/10.1146/annurev-bioeng-061008-124934*, vol. 11, pp. 49–79, Jul. 2009.
[81] J. Li and H. J. Zhu, "Liquid chromatography-tandem mass spectrometry (LC-MS/MS)-based proteomics of drug-metabolizing enzymes and transporters," *Molecules*, vol. 25, no. 11, 2020.
[82] L. C. Gillet, A. Leitner, and R. Aebersold, "Mass Spectrometry Applied to Bottom-Up Proteomics: Entering the High-Throughput Era for Hypothesis Testing," *Annu. Rev. Anal. Chem.*, vol. 9, pp. 449–472, 2016.

[83] X. Ye *et al.*, *Cell surface protein enrichment for biomarker and drug target discovery using mass spectrometry-based proteomics*, 2nd ed. Elsevier Inc., 2019.
[84] Y. Li, H. Qin, and M. Ye, "An overview on enrichment methods for cell surface proteome profiling," *J. Sep. Sci.*, vol. 43, no. 1, pp. 292–312, 2020.
[85] C. C. Wu, M. J. Maccoss, K. E. Howell, and J. R. Y. Iii, "R ESEARCH A RTICLE A method for the comprehensive proteomic analysis of membrane proteins," vol. 21, no. May, 2003.
[86] R. J. Christie, D. J. Anderson, and D. W. Grainger, "Comparison of Hydrazone Heterobifunctional Crosslinking Agents for Reversible Conjugation of Thiol-Containing Chemistry," *Bioconjug. Chem.*, vol. 21, no. 10, p. 1779, Oct. 2010.
[87] S. A. Hanke, A. Kerner, W. M. Nadler, A. Trumpp, Y. Zhang, and C. P. Rösli, "The pivotal role of reactivity in the design of novel biotinylation reagents for the chemical-proteomics-based identi fi cation of vascular accessible biomarkers," *J. Proteomics*, vol. 141, pp. 57–66, 2016.
[88] Y. Tian, Y. Zhou, S. Elliott, R. Aebersold, and H. Zhang, "Solid-phase extraction of N-linked glycopeptides," *Nat. Protoc.*, vol. 2, no. 2, pp. 334–339, 2007.
[89] H. Zhang, X. Li, D. B. Martin, and R. Aebersold, "Identification and quantification of N-linked glycoproteins using hydrazide chemistry , stable isotope labeling and mass spectrometry," vol. 21, no. 6, pp. 660–666, 2003.
[90] M. Han *et al.*, "Support vector machines coupled with proteomics approaches for detecting biomarkers predicting chemotherapy resistance in small cell lung cancer," pp. 2233–2238, 2012.
[91] Z. Sun *et al.*, "Identification of Chemoresistance-Related Cell-Surface Glycoproteins in Leukemia Cells and Functional Validation of Candidate Glycoproteins," 2014.
[92] K. W. Moremen, M. Tiemeyer, and A. V. Nairn, "Vertebrate protein glycosylation: diversity, synthesis and function," *Nat. Rev. Mol. Cell Biol. 2012 137*, vol. 13, no. 7, pp. 448–462, Jun. 2012.
[93] C. A. McDonald, J. Y. Yang, V. Marathe, T. Y. Yen, and B. A. Macher, "Combining Results from Lectin Affinity Chromatography and Glycocapture Approaches Substantially Improves the Coverage of the Glycoproteome," *Mol. Cell. Proteomics*, vol. 8, no. 2, p. 287, Feb. 2009.
[94] N. E. Ö. Küçük, E. Şanal, E. Tan, T. Mitchison, and N. Özlü, "Labeling Carboxyl Groups of Surface-Exposed Proteins Provides an Orthogonal Approach for Cell Surface Isolation," *J. Proteome Res.*, vol. 17, no. 5, pp. 1784–1793, 2018.
[95] P. R. Karhemo *et al.*, "An optimized isolation of biotinylated cell surface proteins reveals novel players in cancer metastasis," *J. Proteomics*, vol. 77, p. 87, 2012.
[96] A. Hofmann *et al.*, "Proteomic cell surface phenotyping of differentiating acute myeloid leukemia cells," *Blood*, vol. 116, no. 13, 2010.
[97] R. Chen *et al.*, "Glycoproteomics analysis of human liver tissue by combination of multiple enzyme digestion and hydrazide chemistry," *J. Proteome Res.*, vol. 8, no. 2, pp. 651–661, 2009.
[98] Y. Yang, V. Franc, and A. J. R. Heck, "Glycoproteomics: A Balance between High-Throughput and In-Depth Analysis," *Trends in Biotechnology*, vol. 35, no. 7. Elsevier Ltd, pp. 598–609, 01-Jul-2017.
[99] V. C. Hodgkinson, G. L. Eagle, P. J. Drew, M. J. Lind, and L. Cawkwell, "Biomarkers of chemotherapy resistance in breast cancer identified by proteomics: Current status," *Cancer Lett.*, vol. 294, no. 1, pp. 13–24, 2010.
[100] P. Giansanti, L. Tsiatsiani, T. Y. Low, and A. J. R. Heck, "Six alternative proteases for mass spectrometry–based proteomics beyond trypsin," *Nat. Protoc. 2016 115*,

vol. 11, no. 5, pp. 993–1006, Apr. 2016.
[101] R. Salz *et al.*, “Personalized Proteome: Comparing Proteogenomics and Open Variant Search Approaches for Single Amino Acid Variant Detection,” *J. Proteome Res.*, vol. 20, no. 6, pp. 3353–3364, Jun. 2021.
[102] K. V. Ruggles *et al.*, “An Analysis of the Sensitivity of Proteogenomic Mapping of Somatic Mutations and Novel Splicing Events in Cancer,” *Mol. Cell. Proteomics*, vol. 15, no. 3, pp. 1060–1071, Mar. 2016.

CURRICULUM VITAE

2015 – 2019 B.Sc., Molecular Biology and Genetics, Necmettin Erbakan University,

Konya, TURKEY

2019 – Present M.Sc., Bioengineering, Abdullah Gul University,

Kayseri, TURKEY