BRNO UNIVERSITY OF TECHNOLOGY
VYSOKÉ UČENÍ TECHNICKÉ V BRNĚ

FACULTY OF CHEMISTRY
FAKULTA CHEMICKÁ

INSTITUTE OF FOOD SCIENCE AND BIOTECHNOLOGY
ÚSTAV CHEMIE POTRAVIN A BIOTECHNOLOGIÍ

PROTEOMIC ANALYSIS OF POSTTRANSLATION MODIFICATIONS IN BREAST CANCER CELL LINE PROFILES

PROTEOMICKÁ ANALÝZA POSTTRASNLAČNÍCH MODIFIKACÍ BUNĚČNÉ LINIE RAKOVINY PRSU.

MASTER'S THESIS
DIPLOMOVÁ PRÁCE

AUTHOR **Bc. Nikola Predná**
AUTOR PRÁCE

SUPERVISOR **Ing. Dana Strouhalová, Ph.D.**
VEDOUCÍ PRÁCE

BRNO 2022

BRNO UNIVERSITY OF TECHNOLOGY
FACULTY OF CHEMISTRY

Assignment Master's Thesis

Project no.:	FCH-DIP1721/2021	Academic year: 2021/22
Department:	Institute of Food Science and Biotechnology	
Student:	**Bc. Nikola Predná**	
Study programme:	Chemistry and Technology of Foodstuff	
Study field:	Food Science and Biotechnology	
Head of thesis:	**Ing. Dana Strouhalová, Ph.D.**	

Title of Master's Thesis:

Proteomic analysis of posttranslation modifications in breast cancer cell line profiles

Master's Thesis:

Deadline for Master's Thesis delivery: 13.5.2022:

Master's Thesis should be submitted to the institute's secretariat in a number of copies as set by the dean This specification is part of Master's Thesis

Bc. Nikola Predná student	Ing. Dana Strouhalová, Ph.D. Head of thesis	prof. RNDr. Ivana Márová, CSc. Head of department

In Brno dated 1.2.2022

prof. Ing. Michal Veselý, CSc.
Dean

ABSTRACT

Estrogen and progesterone receptors, as well as HER2 protein, are currently the most clinically useful metabolic markers in breast cancer. These markers allow for the determination of the type of tumor and its best treatment options. However, one of the most aggressive types of this disease, triple-negative breast cancer (TNBC), lacks these clinically established biomarkers. This means that hormone therapy or targeted drugs are not an option, leaving fewer treatment options to choose from. In order for new tailored drugs to be developed, the understanding of the molecular basis of the disease is crucial. Recently, many studies aim their search for biomarkers at the protein level using proteomics. Proteins, notably their post-translational modifications (PTM), are at the core of many cellular events and their uncovering may help in the understanding of breast cancer mechanisms.

In order to discover the molecular features of TNBC, this study aims to compare proteomic data of untreated cancer cell lines with cells that underwent retinoid therapy. The focus will be on the PTMs, notably glycosylation and phosphorylation, of Vimentin and CD44, which were proposed as potential TNBC biomarkers in previous studies.[1] Protein separation will be carried out using 1D and 2D gel electrophoresis or by SEC-HPLC. The samples will also be subdued to enzymatic cleavage before being identified using MALDI-TOF Mass Spectrometry. In the case of phosphoprotein selective capture, enrichment will be performed by affinity chromatography using TiO_2 phosphopeptide enrichment tips (TopTip). Glycosylated proteins will be enriched using WGA lectin affinity based chromatography. Proteins with significant differences in PTMs between the treated and untreated cells will be evaluated using protein databases (MASCOT, STRING, and more). The data acquired from the study will eventually be used to propose potential biomarkers for TNBC.

ABSTRAKT

Estrogenové a progesteronové receptory, stejně jako HER2 protein, jsou v současnosti klinicky nejužitečnějšími metabolickými markery u karcinomu prsu. Tyto markery umožňují určit typ nádoru a nejlepší možnosti jeho léčby. Jeden z nejagresivnějších typů tohoto onemocnění, triple-negative breast cancer (TNBC), však tyto klinicky stanovené biomarkery postrádá. To znamená, že hormonální terapie nebo cílené léky nepřicházejí v úvahu, takže je na výběr méně možností léčby. Aby bylo možné vyvinout nové léky na míru, je zásadní pochopení molekulárního základu onemocnění. V poslední době se mnoho studií zaměřuje na hledání biomarkerů na úrovni proteinů pomocí proteomiky. Proteiny, zejména jejich post-translační modifikace (PTM), jsou jádrem mnoha buněčných událostí a jejich odhalení může pomoci při pochopení mechanismů rakoviny prsu.

Pro objevení molekulárních rysů TNBC, je cílem této studie porovnat proteomická data neléčených rakovinných buněčných linií s buňkami, které podstoupily retinoidní terapii. Důraz bude kladen na PTM, zejména glykosylaci a fosforylaci, Vimentinu a CD44, které byly navrženy jako potenciální biomarkery TNBC v předchozích studiích.[1] Proteinová separace bude provedena pomocí 1D a 2D gelové elektroforézy nebo pomocí SEC-HPLC. Vzorky budou také podrobeny enzymatickému štěpení před identifikací pomocí MALDI-TOF hmotnostní spektrometrie. V případě fosfoproteinového selektivního záchytu bude obohacení provedeno afinitní chromatografií s použitím hrotů pro obohacení fosfopeptidu TiO2 (TopTip). Glykosylované proteiny budou obohaceny pomocí WGA lektinové afinitní chromatografie. Proteiny s významnými rozdíly v PTM mezi ošetřenými a neošetřenými buňkami budou blíže hodnoceny pomocí proteinových databází (MASCOT, STRING a další). Data získaná ze studie budou případně použita k navržení potenciálních biomarkerů pro TNBC.

KEY WORDS

Triple-negative breast cancer, proteomics, MDA-MB-231, MCF-7, biomarkers, ATRA, retinoids, Vimentin, CD44, EMT, cell stemness, PTMs

KLÍČOVÁ SLOVA

Triple-negativní karcinom prsu, proteomika, MDA-MB-231, MCF-7, biomarkery, ATRA, retinoidy, Vimentin, CD44, EMT, kmenové buňky, PTM

PREDNÁ, Nikola. *Proteomic analysis of posttranslation modifications in breast cancer cell line profiles*. Brno, 2022. Available at: https://www.vutbr.cz/studenti/zav-prace/detail/140423. Master's thesis. Brno University of Technology, Faculty of Chemistry, Institute of Food Science and Biotechnology. Head of thesis Ing. Dana Strouhalová, Ph.D.

DECLARATION

I declare that the Master's thesis has been worked out by myself and that all the quotations from the used literary sources are accurate and complete. The content of the Master's thesis is the property of the Faculty of Chemistry of Brno University of Technology and all commercial uses are allowed only if approved by both the supervisor and the dean of the Faculty of Chemistry, BUT.

.....
student's signature

ACKNOWLEDGEMENTS

I thank once again my supervisor, Ing. Dana Strouhalová, Ph.D., for a very pleasant and informative experience on this meaningful project. Despite the odds not always being in our favor, you did your best to grant me access to many new techniques and methods of a proteomics laboratory and greatly prioritized our success. Thank you to the Institute of Analytical Chemistry of the Czech Academy of Sciences for letting me work in their facility. Thanks also to my supportive partner, family and friends.

TABLE OF CONTENTS

1 INTRODUCTION

When encountering a disease like cancer, time is of upmost importance. The stage at which the defect is detected sets the probability at which the patient may be cured. This is why great emphasis is put on the study of various methods of early detection. One of these methods is the identification of characteristic compounds, biomarkers, which are highly expressed in patients with cancer. In other words, these are compounds which reflect that something is wrong with an individual's health. The idea is that in practice, a simple blood test would suffice in the detection of cancer. A number of gene and protein based biomarkers are clinically recognized, including PSA (prostate specific antigen) for prostate cancer, CEA for colorectal cancer, S100 proteins for melanoma, HER-2 for breast cancer, and many others. In combination with DNA information of the patient, predictions of an individual's response to drug doses may be made, making cancer treatment options much more reliable than in the past. The study of proteomes of cancerous cells through proteomics may help in making this a reality. [2]

2 THEORETICAL PART

2.1 Breast cancer

2.1.1 Introduction

Breast cancer is a type of malignant carcinoma that affects mainly women, with men making up about 1 % of the affected. As long as the estrogen hormone is active, the risk of breast cancer increases. Females who begin their first menstrual cycle at a younger age are more likely to develop the cancer, as do females with a later onset of menopause. Studies have shown that women bearing their first child at a higher age or women with no children at also carry a higher risk.[3] Another factor determining the likelihood of cancer is age; it is believed that cancer may develop as a result of an error occurring during the copying of DNA before cell division. [4] More factors include obesity, harmful use of alcohol, family history of breast cancer, history of radiation exposure, reproductive history (such as age that menstrual periods began and age at first pregnancy), tobacco use and postmenopausal hormone therapy. [5] Unfortunately, even if these potential risk factors are controlled, the risk of cancer development is said to be reduced by only about 30 %.

The incidence of breast cancer has been slowly increasing due to the lack of effective diagnostic tools and early detection. The mortality rates have reached up to 522 000 deaths in 2017 alone.[6] To compare, in 2020, there were 685 000 deaths by breast cancer globally. Breast cancer is the most frequently diagnosed cancer as well as the leading cause of death by cancer among females. In 2015, it accounted for 23 % of the total cancer cases and 14 % of deaths by cancer. [7]

2.1.2 Classification

Breast cancer types are identified by the nature of the cells which become cancerous. The subtypes are characteristic by their cell origin, somatic changes and their overall molecular mechanism. As such, breast cancer cannot be treated by the same treatment procedures in all cases and it is important to be able to identify each type individually. [8]

The molecular subtype of an invasive breast cancer is characterized by the genes, expressed by the cancer cells, which regulate how the cells behave. Researchers have identified five main molecular subtypes of invasive breast cancer; "luminal A", "luminal B", "luminal B-like", "HER2-enriched" and "triple negative" or "basal-like" breast cancer. Each subtype is also based on the levels of certain biomarkers, some of which include estrogen receptors (ER), progesterone receptors (PR), the human epidermal growth factor receptor 2 (HER2) or the Ki-67 protein.[9]

The "luminal-A" breast cancer is ER and PR-positive, HER2-negative and it's levels of the protein Ki-67, which plays a role in how fast the cancer cells grow, are low. "Luminal B" is ER-positive, PR-

negative and HER2-negative, as well as high in levels of Ki-67. “Luminal B-like” is ER-positive and HER2-positive, has any level of Ki-67 and may be PR-positive or negative. [10] “HER2-enriched” cancer is ER-negative, PR-negative and, as the name suggests, HER2-positive. [11,12] Finally, some subtypes cannot be characterized by either of the recognized biomarkers. These types are often referred to as “triple negative” breast cancer (TNBC).

2.1.3 Triple negative breast cancer

Lacking the expression of neither ER, PR or HER2 biomarkers, TNBC poses a challenge for detection in general. Moreover, TNBC tends to grow and spread faster than other types of invasive breast cancer and thus treatment options are very limited, making it one of the most difficult types of cancer to deal with. [13]

Beacuse of the lack of common breast cancer biomarkers, targeted therapy may not be an opiton for treatment. Thus, chemotherapy or surgery (provided that the cancer has not spread) is often used. Radiation may also be an option depending on the type of tumor and surgery that the patient had underwent. Generally, patients with TNBC are of high risk of the cancer relapsing and progression. So far, there is no preferred standard form of chemotherapy for patients with TNBC. [14 15]

2.1.4 Treatments options

Breast cancer may be effectively treated, with survival probabilities reaching 90 % or higher, especially when identified in the early stages. Diagnosis is performed either through physical examinations, where the breast tissue undergoes mammography imaging either by a physician or by the patient herself, or by performing biomarker detection tests directly on the biopsy material taken from the tissue or on the serum sample. Currently, the most commonly used biomarkers used in the clinic are the CA 15-3, CA 27-29, carcinoembryonic antigen (CAE), estrogen receptor (ER), progesterone receptor (PR) and the human epidermal growth factor 2 (HER2).[16]

Cancers that express the estrogen receptor (ER) and/or progesterone receptor (PR) usually respond to endocrine therapies. These are taken orally for longer periods of time and may reduce the chance of recurrence of these types of cancer nearly by half.

As of now, hormone receptor negative cancers need to be treated with chemotherapy, as there is no targeted treatment option.

Some breast cancers that overexpress the molecule called HER2 may be cured with targeted biological agents, antibodies, such as trastuzumab, which can also be combined with chemotherapy for better effectiveness.

Lastly, radiotherapy may also greatly aid in curing breast cancer is the cancer is diagnosed at an early stage. Radiation therapy may prevent a woman from having to undergo a mastectomy. If, however, the cancer is already at a later stage, radiotherapy plays a big role in the reduction of recurrence risk even after mastectomy is performed.[17 18]

2.1.5 Research tools

When analyzing the behavior of cancerous cells, it is important that researchers work on equal ground, ideally using the same set of cells to allow for better comparison between the analysis results. For this reason, cancer cell lines were obtained and are cultivated continuously to provide an indefinite source of biological material for research purposes. Some of the most used breast cancer cell lines include MDA-MB-231 and MCF-7. These two cell lines, together with cell line T-47D, account for more that two-thirds of all studies mentioning breast cancer cell lines. [19] Tumor cell lines are used as in vitro models serving as an essential tool in present-day cancer research.

2.1.5.1 MDA-MB-231

The breast cancer cell line MDA-MB-231 obtained from pleural effusion is classified as adenocarcinoma by tumor type. Characterized by low expression of ER, PR and HER2, it is referred to as triple-negative (TNBC).

Established in 1974[20], MDA-MB-231 is specifically used as a model for TNBC research and is one of the most commonly used breast cancer cell lines in the world. The cell line contains a high level of Vimentin, a protein corresponding to the to the mesenchymal phenotype, which makes Vimentin a big candidate for being a TNBC biomarker.

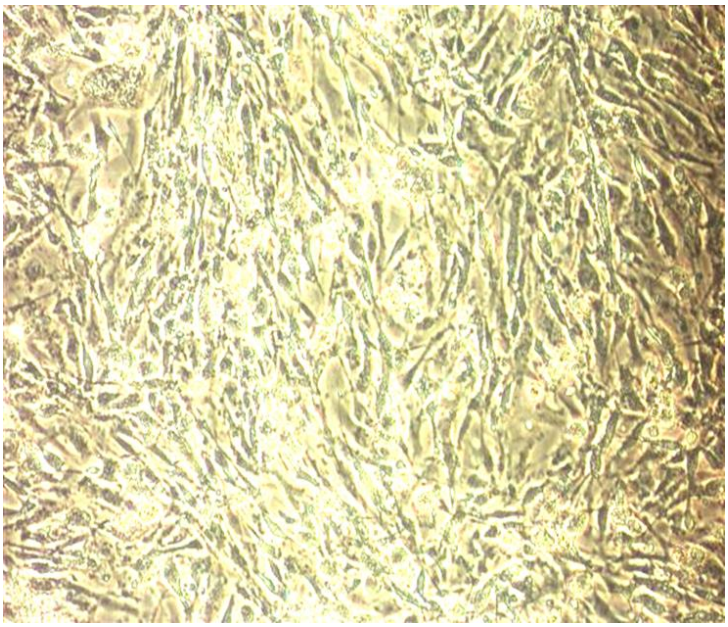

Fig. 1: A view on the MDA-MB-231 cells under a microscope.

2.1.5.2 MCF7

MCF-7 is a breast cancer cell line isolated in 1970 from a 69-year-old white woman with metastatic disease.[21] This cell type expresses markers of the luminal epithelial breast cancer phenotype. They are a great tool for the study of hormone response thanks to their high hormone sensitivity through the expression of ER. It was not possible for cancer researchers to obtain a mammary cell line that was capable of living longer than a few months, prior to MCF-7. As such, it is the first hormone-responsive breast cancer cell line.[22]

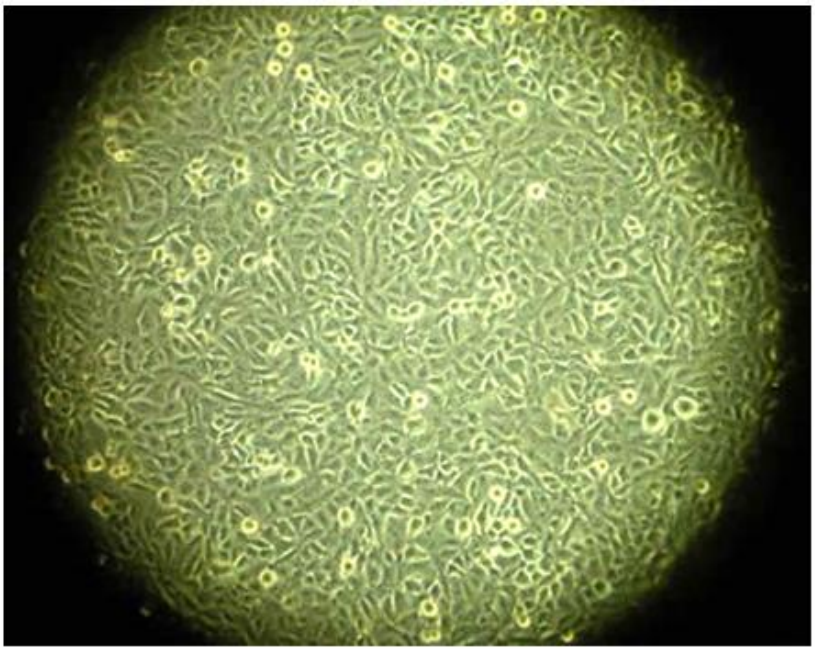

Fig. 2:A view on the MCF-7 cells under a microscope

2.2 Proteomics

2.2.1 Introduction

The study of proteins or proteomes is yet another stage of understanding the molecular basis of cellular processes after the succesful deciphering of the human genome. The term „proteome“ was first used in 1994 by an Australian post-doctoral named Mark Wilkins. According to his first definition, the proteome refers to the total set of proteins present in a cell at a given time. [23]

Proteomics serves to describe the structure and function of functional products from gene expression, together with the identification of each protein, its quantification, location in the cell and to discover protein-protein interactions and characteristics of post-translational modifications (namely phosphorylation, glycosylation, acetylation, methylation, and others), which play an important role in cellular processes. [24]

2.2.2 A Proteomics Laboratory

A proteomics laboratory can be used for a variety of studies, however all of these experiments may be characterised by a number of analytical tasks. What follows is a brief description of these tasks which are essential capabilities of a proteomics laboratory. It should be noted that not all experiments consist of all of the tasks listed. Thus, emphasis will be placed on methods used specifically in the thesis.

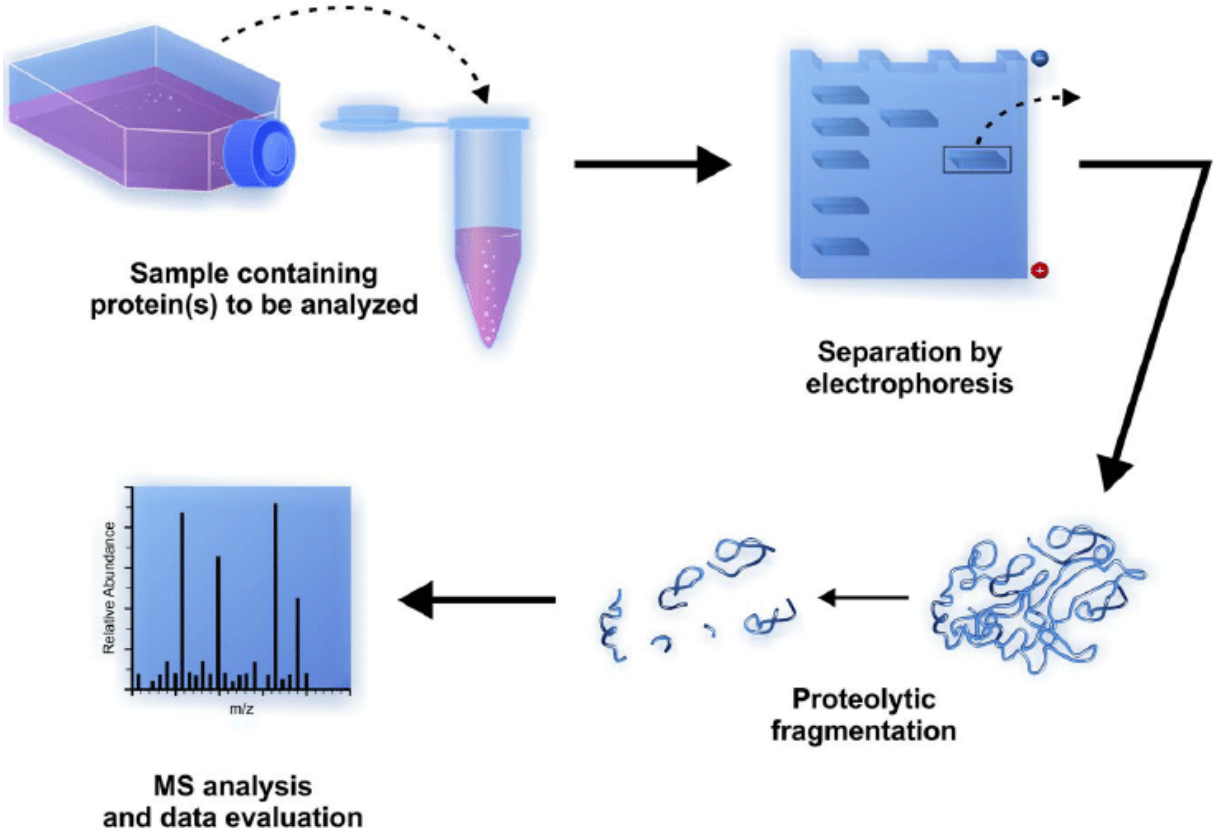

Fig. 3: Steps of proteomic analysis using MS after separation and in-gel digestion[25]

2.2.2.1 Cell cultivation

As of today, cell cultures are widely used both in research and in the production of, for example, antibodies, protein molecules and peptides. Cells for cultivation are taken mainly from humans or laboratory animals. Cultures can be organ, tissue or cell (cell culture occurs when the cells leave the original tissue and divide separately in the medium). The use of cell cultures in experimental work has several advantages over other biological models. The experiment is performed on a single cell of a precisely defined type, the results are not affected by reactions with other tissues. In a relatively short

time, the amount of precisely characterized homogeneous material can be obtained. However, creating physiological conditions is not possible, which may lead to phenotypic changes in cultured cells and, as a result, skew the conclusions.[26]

2.2.2.2 Protein and peptide separation

Most proteome analyses begin with a mixture of various proteins. Nowadays, there are many extraction and franctional methods. These methods usually exploit differences in protein size, physico-chemical properties, binding properties or biological activity. It is essential that all interfering elements are removed from the analyzed sample. The number of separation steps is determined by the complexity of the sample and the goals of the research. Some of the methods most widely used are one- and two-dimensional gel electrophoresis or high performance liquid chromatography (HPLC).

2.2.2.2.1 Gel electrophoresis

The method of gel electrophoresis allows the fractionation of proteins by their isoelectric point (pI) and molecular weight (M_r). The gel is usually composed of polyacrylamide, but gels from agarose or starch may also be used.[27]

The first dimension of two-dimensional (2D) gel elecrophoresis is isoelectric focusing (IEF) and separates the proteins by their pI, which is the unique pH value where a protein has no net charge. For this technique, proteins are introduced into an immobilized pH gradient gel. An electric current is passed through the gel and the molecules are pulled through the medium by the cathode and anode ends to their respective isoelectric points, where they cease to migrate.[28]

The second dimension is SDS-polyacrylamide gel electrophoresis (SDS-PAGE). In this techique, sodium dodecyl-sulphate (SDS) is used as a surfactant, binding strongly to the protein chains in order to contribute a negative charge. This negative charge overlaps any potential charge of the protein.[29] In addition, the number of SDS molecules that bind to a protein is proportional to the length of the protein molecule[30], thus allowing the sample to be separated only by molecular weight (size) and not by charge, since the final charge of the protein after association with SDS is always negative and proportional to the protein size. The gel is placed in an electrophoresis buffer with suitable electrolytes which maintain the pH at a constant value and provide ions to carry the current during the process of electrophoresis. Thereafter, a voltage is applied to the gel loaded with samples and the negatively charged molecules migrate towards the positively charged anode. Small molecules pass through the gel relatively easily, while larger molecules are retained, thereby separating the molecules by their molecular size. During electrophoresis, a set of standards is loaded onto the gel together with the samples. These standards are often referred to as protein ladders and allow for identification of the approximate size of a sample molecule. SDS-PAGE is usually paired with IEF for 2D gel separation of proteins, but may be also used alone for 1D SDS-PAGE separationg.

Following electrophoresis, dyes are used for better visualisation of the separated molecules. This process is referred to as protein staining. Coomassie® Blue are a family of dyes most commonly used to stain proteins on polyacrylamide gels. This treatment allows the visualization of proteins as blue bands or spots on a clear background.[31] [32]

The advantages of gel electrophoresis are its relative speed (the process takes approximately 30 minutes), low cost and simplicity. In addition, there are many post electrophoresis options for detection including colorimetric and fluorescent total stains, and specific detection with mass spectrometry. Its disadvantages are its low reproducibility, although precast gels and ready-made buffers improve the situation. A big hurdle is removing the macromolecules from the gel itself and getting them into the right condition for mass spectrometry.

2.2.2.2.2 Liquid chromatography

Chromatography includes all experimental techniques that use intentionally chosen reagents, conditions or systems in the pursuit of separating the constituent parts of a mixture. Different manifestations of

chromatography include High-Performance Liquid Chromatography (HPLC), gel-permeation chromatography (size exclusion chromatography), color chromatography, column chromatography, thin-layer chromatography (TLC), and more.[33] Gel electrophoresis is a chromatographic method as well.

Liquid chromatography is a variant of column chromatography. The eluent containing the analyte undergoing separation is referred to as the mobile phase. A column filled with an adsorbent (commonly silica or alumina), is analogically the stationary phase. The separation occurs on the notion that the different parts of the analyte will exhibit different rates of elution based on their unique interaction with the stationary phase. This characteristic reaction is almost entirely based on the intermolecular attraction of polarity between the two phases. In practice, the liquid functioning as the mobile phase drives the analyte through the column where separation occurs.[34] An interesting type of liquid chromatography is the use of tips with chromatography media fixed at its` ends, such as C18 chromatography tips. They present a simple and reproducible method for concentration and purification of protein and peptide samples.

In proteomics research, HPLC is generally used for the separation of peptides after digestion. The instrumentation in HPLC consists of pumping systems, separation columns and detectors which allow for separation under high pressure.[35] Unlike in conventional HPLC, all system components are downscaled for a proteomics laboratory.[36] Compounds are determined based on their retention time in the column, which is usually mesured by a UV detector through the intensity of absorbance. Other types of detectors may be used, such as mass spectrometry (LC-MS).[37]

2.2.2.3 Protein Digestion

While it is possible to study intact proteins and their modifications by mass spectrometry (MS), the most common proteomic approaches currently utilize digestion. MS instrument sensitivity toward whole proteins is lower than that for cleaved proteins. This is due to the reduction of the m/z ratio which occurs during the cleavage of proteins to peptides. Peptides are easier to characterize, they can be separated by reverse phase supports using high performance liquid chromatography (HPLC) with a C18 column. With that in mind, protein digestion is considered an essential step in the analysis of proteomes. Although both enzymatic and nonenzymatic digestion are possible, enzymes have the ability to reduce the activation energy of the reaction, allowing working under mild conditions with high protein yields. Enzymatic digestion is thus preferred. Their disadvantage is their high cost.

The most widely used protease is Trypsin, partially due to its high purity and affordable price. It is proven to be very effective in the specific cleavage of Lysine and Arginine residues.[38]

2.2.2.3.1 In-gel digestion

Gel cleavage is mainly used to digest proteins after gel separation (1D or 2D). Its advantage is in the anchoring of the proteins in the gel, which ensures that the proteins stay unmodified. However, the process is quite time consuming. Not only does the digestion itself take more time, as the pores of the gel impede the availability of the enzyme to the protein, but also the digested fragments need to be extracted from the gel by repeated use of acetonitrile. This means the approach in significantly more complex than in-solution digestion. In addition, because of the gel making it difficult for the enzyme to access the proteins, it is generally advised to use larger quantities of the enzyme.[39]

2.2.2.3.2 In-solution digestion

In-solution digestion is usually used in LC-MS/MS analysis. Since the digestion occurs in a solution, the proteins are freely available for an enzyme. As a result, the enzyme does not have to be in excess and the digestion products dont require extraction. This approach tends to be simpler that in-gel digestion in terms of sample handling and speed, however it requires sophisticated LC-MS instrumentation which requires constant maintenance. Another disadvantage is a higher risk of protein modifications, as the reagents supplied remain in the solution together with the sample.[40]

2.2.2.4 Protein Identification

The most widely used technique for protein identification is Mass Spectrometry (MS).

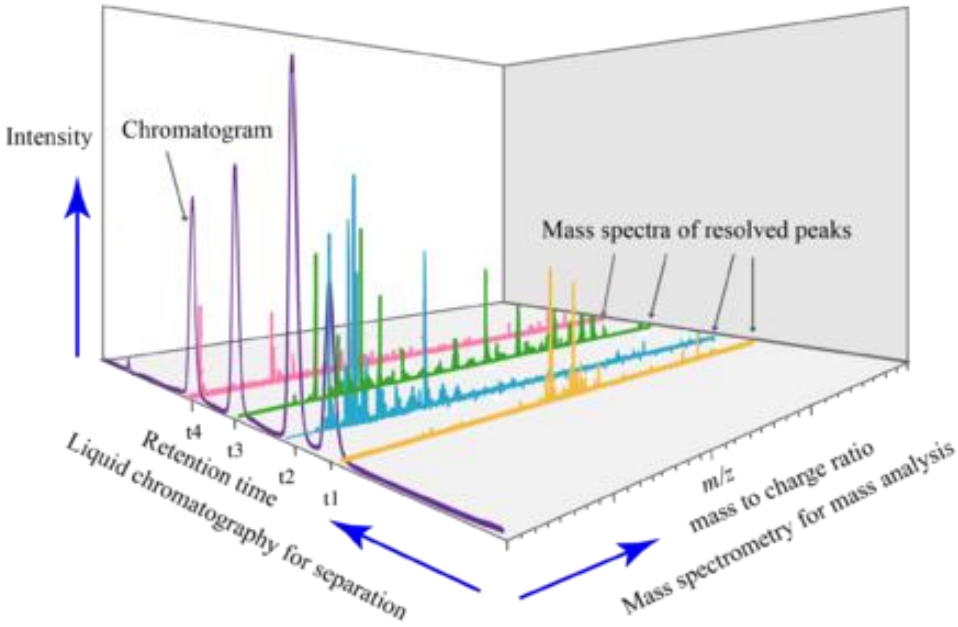

Fig. 4: LC-MS spectrum 3D analysis[41]

2.2.2.4.1 Mass Spectrometry

In mass spectrometry, proteins or peptides (gained using prior digestion) are ionized and separated according to their specific mass to charge ratio (m/z) and their relative abundance is measured to obtain a mass spectrum of the analyzed molecules. The mass spectrometer can thus be thought of as two distinct components: the ionization source and the mass analyzer.

2.2.2.4.1.1 Ionization

The first step in the mass spectrometric analysis of proteins is the production of ions. The sample of interest is ionized, with a positive or negative charge, and then desorbed into the gas phase. This is possible thanks to numerous techniques, one of them being electronspray ionization (ESI) and the other matrix-assisted laser desorption/ionization (MALDI). The latter is preferred for quick protein identification. In MALDI, it is typical for ionized analytes to have a single plus charge. Both techniques are considered as soft (low fragmentation) ways of obtaining ions of large molecules in the gas phase, however MALDI tends to produce far fewer multi-charged ions. [42] In MALDI, the samples are placed on a metal plate together with a laser energy absorbing matrix which allows for the desorption of the sample when it is hit by a laser, forming ions in the process.[43] [44]

2.2.2.4.1.2 Mass Analysis

Following ionization, the ions are injected and accelerated in an electric or magnetic field to the detector. This segment is called the mass analyzer. The most widely used instruments for peptide mass analysis are MALDI-TOF instruments, TOF standing for time-of-flight.[45] Ions are accelerated and separated by mass in a field-free drift region before detection. These instruments are preferred as they permit obtaining peptide mass fingerprints (PMFs) in a short amount of time.

2.2.2.4.1.3 Protein Identification

The end result is shown on a mass spectrum, which is a plot representing intensity vs m/z. This may be seen in *Fig.4*. Intensity loosely represents the abundance of a particular fragment while the m/z is useful for qualitative analysis. Each peak corresponds to an individual ion fragment. The identification of

proteins is possible through deduction from the fragments identified by database search. Raw data (e.g., mass spectra) is analyzed against theoretical spectra from protein databases to generate derived data (e.g., lists of peptide and protein identifications corresponding to the spectra). These databases and servers are created for scientists to be able to access a wide range of resources in many different domains. Analysis and review of the data set tends to be the most time-consuming part in a proteomics laboratory.

2.2.2.4.1.4 Different approaches

There are more than one approaches to MS. It is possible to identify a highly purified protein based on the measured molecular weight (the „top-down" method), however identification based on the mass spectrum of a protein's peptides (the „bottom-up" method) is more accurate and desirable in complex samples of abundant proteins that require sensitive separation.[46] The mass spectra of a protein's peptide fragments is referred to as „peptide fingerprint". Peptide mapping alone works well for simple mixtures of proteins, however to confirm the identification provided by a peptide fingerprint of a more complex mixture, individual peptides can be further fragmented directly in the mass spectrometer. This process is commonly referred to as tandem MS or MS/MS and provides separation of ions or peptides with similar m/z ratios.[47]

In the method of MS/MS, the sample is first ionized and analyzed as in a regular MS approach. Fragments of a specific m/z are selected and once again fragmented by another mass spectrometer. This way, fragments of, for example, peptides, may be further fragmented to pieces of amino acids and an amino acid sequence may be achieved. It is very useful in sample analysis where peptide masses may overlap.

LC is often paired with MS or MS/MS in order to maximize identification of the sample components. LC separates the compound mixture after digestion in order to resolve the main peaks from the minor contaminants in the chromatogram. LC-MS/MS is thus more specific and sensitive than standard MS/MS in detecting components.[48]

2.2.2.5 Quantitative Analysis

In order to understand the proteome dynamics in biology, comparative analyses involving quantification are essential. Simply comparing samples visually in gels after 2D electrophoresis usually does not suffice and requires the use of more sophisticated methods. Common methods for relative quantification of peptides in different samples use labeling of peptides using isobaric tags and analysis by tandem MS, although label-free methods also exist.[49] One of the commercially available tags are Isobaric tags for relative and absolute quantitation (iTRAQ).

iTRAQ, or isobaric tag for relative and absolute quantification, are special reagents with stable isotope molecules which bind covalently to the peptide via free amines at the peptide N-terminus after digestion.[50] The basic overview of a labeled peptide is shown in *Fig.5*. Each sample is labeled with a different iTRAQ tag and the labeled samples are then combined into one mixture. LC-MS/MS is most commonly used for both identification and quantification using this method.[51] The reagents consist of a mass balancing group (carbonyl group) and a reporter group based on N-methylpiperazine. When both groups are present, the iTRAQ tags are chromatographically indistinguishable, thus the ion peak of each labelled peptide is detected simultaneously in the mass spectrometer. After MS/MS is applied, the mass balancing group is released as a neutral fragment and the peptides gain a specific m/z based on the isotope-encoded reporter ions. Labelled peptides do not show a mass shift in MS as the tags are isobaric. Labeled peptides are then identified using databases. An analysed labelled peptide will have a specific m/z based on the tag used, so peaks of the peptide from differnet samples may be compared for relative quantification.

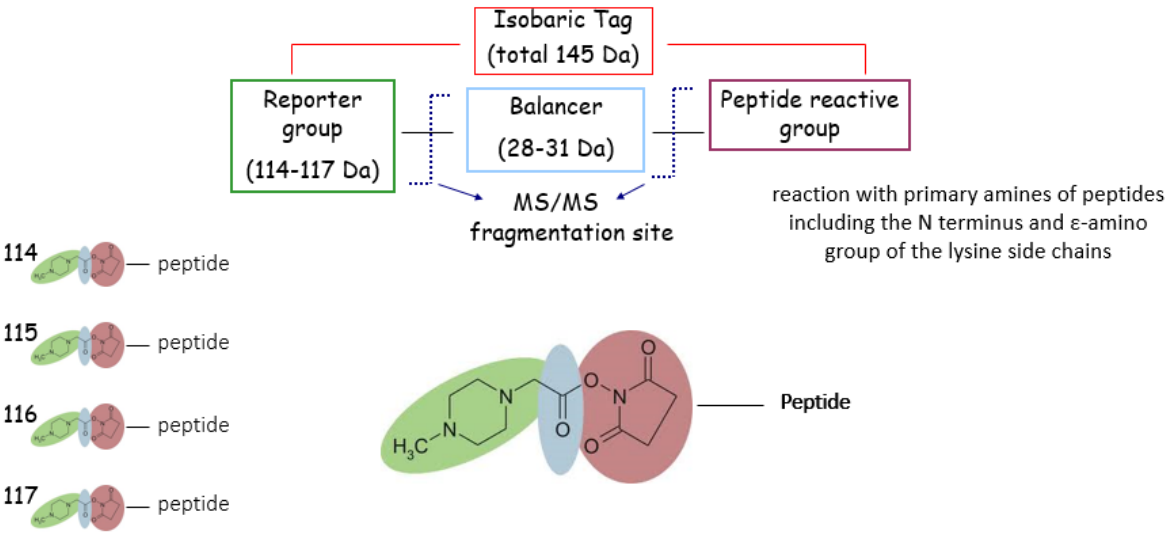

Fig. 5: Basic overview of an iTRAQ tag[52]

2.2.2.6 Proteome Profiling

Protein profiling is a rapidly evolving technique. The aim is to obtain mass spectra which would represent the intact proteins present in the sample. The method is most commonly used for serum and tissue samples, but is applicable to any complex sample.[53]

2.2.3 Proteomics in cancer research

Cancer has long been understood as a disease caused by genomic instability.[54] [55] Defects in DNA repair, alterations in abundance and sequence all contribute to disease onset and progression. Whilst gene therapy is nowadays more and more accessible, it is becoming apparent that the field of proteomics must play a role in clinical practice, as proteins cause actions depending on the type of cell they dock with.

Quantitative proteomics could greatly enhance the current understanding of the molecular mechanisms in cancer and in doing so help identify relevant biomarkers for these diseases.[56] Ideally, clinicians might recommend combinations of molecularly targeted agents and other therapies on the basis of an individual patient's proteomic profile.

Personalized approaches will hopefully be the basis of cancer treatment in the future. Appropriate molecularly directed therapy will be matched to individual patients thanks to the identification of critical molecules necessary for tumor growth and survival.[57]

2.3 Retinoids

2.3.1 Introduction

Retinoids are by definition a class of compounds structurally related to vitamin A. They are involved in complex physiological processes and developmental reactions in many tissues of higher vertebrate animals. They take part in embryonic development, in the process of vision, reproduction, bone formation and hematopoiesis. They also greatly affect the metabolism, growth and differentiation of different cell types, apoptosis and the carcinogenesis process.[58] [59]

It has come to the conclusion that the bioavailability of most retinoids can be increased if taken on a high fat diet. Retinoids cross the placenta and are also excreted in breast milk, as shown in animal experiments. Clinical use is still limited by the fact that these substances can act as teratogens if consumed in higher amounts.[60]

2.3.2 Classification of retinoids

Retinoids comprise of three groups. First-generation retinoids include non-aromatic compounds such as tretinoin (retinoic acid; ATRA} and isotretinoin (13-cis-retinoic acid). The second generation of

retinoids are made up of monoaromatic compounds such as etretinate or acitretin. These are now used to treat severe forms of psoriasis. Third-generation retinoids include adapalene or tazarotene, which are newly synthesized polyaromatic retinoids often referred to as arotenoids.

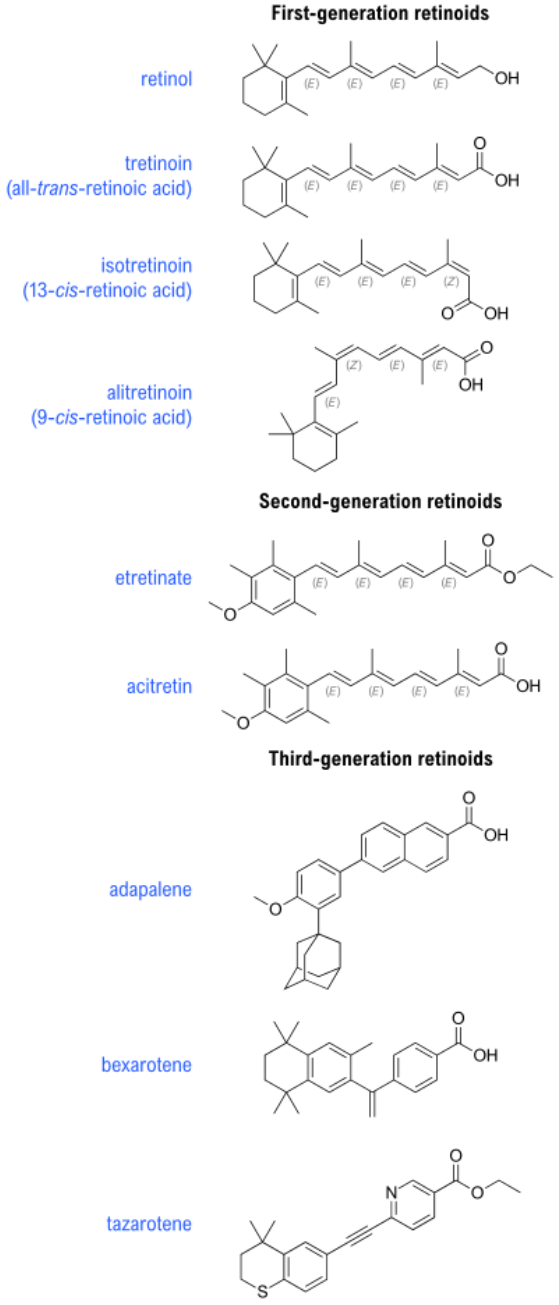

Fig. 6: Retinoid classes and their representants.[61]

2.3.3 Retinoid structure and function

Retinoids, as well as rexinoids, are polyisoprenoid compounds which structurally form a cyclohexenyl ring including retinol (vitamin A) as its building block together with its natural metabolites or synthetic derivatives. The polyprenoid side chain, terminated by a carbonyl functional group, affects retinol's susceptibility to metabolic transformations as well as protein interactions. The particular structure of retinoids causes an increase of sensitivity to oxidation and UV radiation.[62]

Natural retinol metabolites contain all-trans-, 9-cis and 13-cis retinoic acids. Of these, the all-trans variant is biologically the most active and stable. All-trans retinoic acid is synthetized by oxidation of

all-trans retinol to all-trans retinaldehyde and subsequential irreversible oxidation to all-trans retinoic acid.

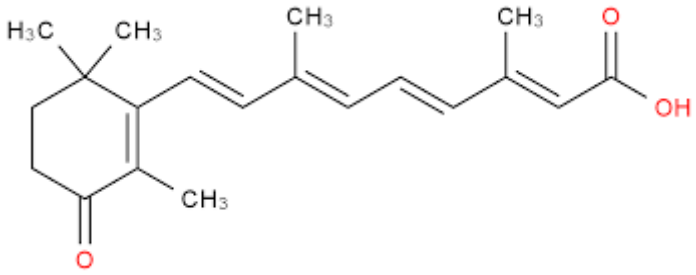

Fig. 7: Tretinoin (ATRA) molecular structure

Retinoids are involved in a number of processes, including embryonic development, immune responses, bone formation, hematopoesis, metabolism, reproduction, vision as well as in various processes of carcinogenesis or apoptosis.[63]

Retinoids acting through their cognate nuclear receptors may be beneficial to health, however they may also be detrimental. They are known to have tumour-suppressive activity but on the other hand, they are teratogenic[64], thus they are often not recommended, especially in pregnancy.[58]

2.3.4 Retinoid metabolism

The metabolism of vitamin A is crucial for the synthesis of retinoic acid, which is needed to maintain gene expression and tissue differentiation.

Retinoids are transported by the blood to cells, where they are taken up by retinol binding proteins (RBPs). After binding to proteins, they are transported to the cell, where they then bind to specific nuclear receptors. Here, oxidation occurs by the enzymes retinol dehydrogenase and retinaldehyde, which results in retinoic acid (RA), which is the most important compound in this signaling cascade.

Retinoic acid serves as a major ligand for retinoid receptors in the nucleus. RA is transported by a carrier protein to the cell nucleus, where it binds to the RAR (retinoic acid receptor) or RXR (retinoic acid receptor X), each of which has three alpha, beta and gamma subunits.[65 66] Each of these types of receptors has different functions in different tissues. The triggering of the receptors itself depends on the form in which the RA has. For example, only all-trans retinoic acid (ATRA) binds to RAR; in contrast, 9-cis-retinoic acid and 13-cis-retinoic acid bind to RXR. These receptors then bind to sections of DNA called responsive elements (RARE or RXRE), which are located on the promoter of the target genes and transcribed.[67]

2.3.5 Retinoids in cancer research

Today, tretinoin (ATRA) is greatly useful in the treatment of acute promyelocytic leukemia while alitretinoin (9-cis-retinoic acid) serves as a topical solution to treat skin lesions from Kaposi's sarcoma.[68]

Furthermore, retinoid treatment, alone or in combination, has been suggested as a treatment of a number of conditions, some of which include lung cancer[69], ovarian cancer[70 71], kidney cancer[72], bladder cancer[73], and other forms of cancer. Clearly, many data have shown that retinoids inhibit carcinogenesis, suppress tumour growth and its invasion in a variety of tissues.[74] Finally, data also suggests that retinoids are able to inhibit human breast cancer.[75]

The idea is that these retinoid drugs are to bind to specific receptors (ER and PR, as explained in previous chapters) and this way allow for inhibition of cell growth and differentiation. An emphasis has been put on the use of retinoids against cancer stem cells, as they are relatively resistant to conventional therapies such as radiotherapy and chemotherapy. [76 77] The effectiveness of ATRA treatment in inducing differentiation of CSCs has been supported in the past. [78] so consideration has been given to how ATRA and its derivatives could be included in clinical practice. For example, various combinations of ATRA and 9-cis RA have been tested. All 18 combinations have been found to cause strong inhibition of proliferation and induction of proliferation. In conclusion of these studies, retinoids cause apoptosis. [79]

2.4 Protein markers in breast cancer

2.4.1 Introduction

Proteomic cancer research has been so far very successful in the proposal of numerous biomarkers. Some of these have already been validated and are currectly used in clinical practice. A number of cancer markers have been recommended by the American Society of Clinical Oncology (ASCO) in 2007, such as urokinase plasminogen activator (uPA) and plasminogen activator inhibitor (PAI-1), which have been associated with angiongenesis, invasive and metastatic breast cancer. Other proposed biomarkers included the estrogen receptor (ER), progesterone receptor (PR) and human epidermal growth factor receptor 2 (HER2) [80]. However, many of the proposed markers, excluding the three latter, proved generally insufficient sensitivity and specificity to be used in clinical practice.[81] Some of these markers include the much discussed Ki67, a marker of proliferation.

This proves that proteomic profiling and successful protein biomarker discovery is not an easy task. One must not forget that the proteome is very complex. Biomarker research is dependent on the current proteome coverage and the sensibility of the bioanalytical instruments, making the process quite lenghtly. Nevertheless, the search for new disease markers is highly desirable from the clinical point of view. Breast cancer biomarkers are of various nature depending on the conditions of the tumor formation.

2.4.2 Vimentin (EMT markers)

An important process in tumor transport is the EMT process, in which a cell's, originally an epitheial cell's, proteome profile changes in a way that causes the cell to adopt a more mesenchymal nature.

Currently, the protein Vimentin may serve as a marker for the EMT process. This is discussed in a clinical study from 2015 [82], where cancer patients' samples have been tested for Vimentin and E-cadherin presence. In conclusion of the study, the patients with negative detection of E-cadherin and positive detection of Vimentin had significantly poorer prognoses. On the other hand, patients whose samples tested E-cadherin-positive and Vimentin-negative had a higher chance of curing the cancer. This suggests that the loss of E-cadherin is usually observed during the process of EMT as an epithelial character, while higher levels of Vimentin signify mesenchymal phenotype presence. Overall, cancer cells with an EMT phenotype have been linked to greater invasiveness and resistance to chemotherapy.[83]

The levels of Vimentin have been shown to be significantly higher in MDA-MB-231 cell lines than in their ATRA treated counterparts. This confirmed that the EMT process is ongoing in TNBC as well as that ATRA may be capable of selectively targeting the Vimentin protein, inhibiting its function.[84]

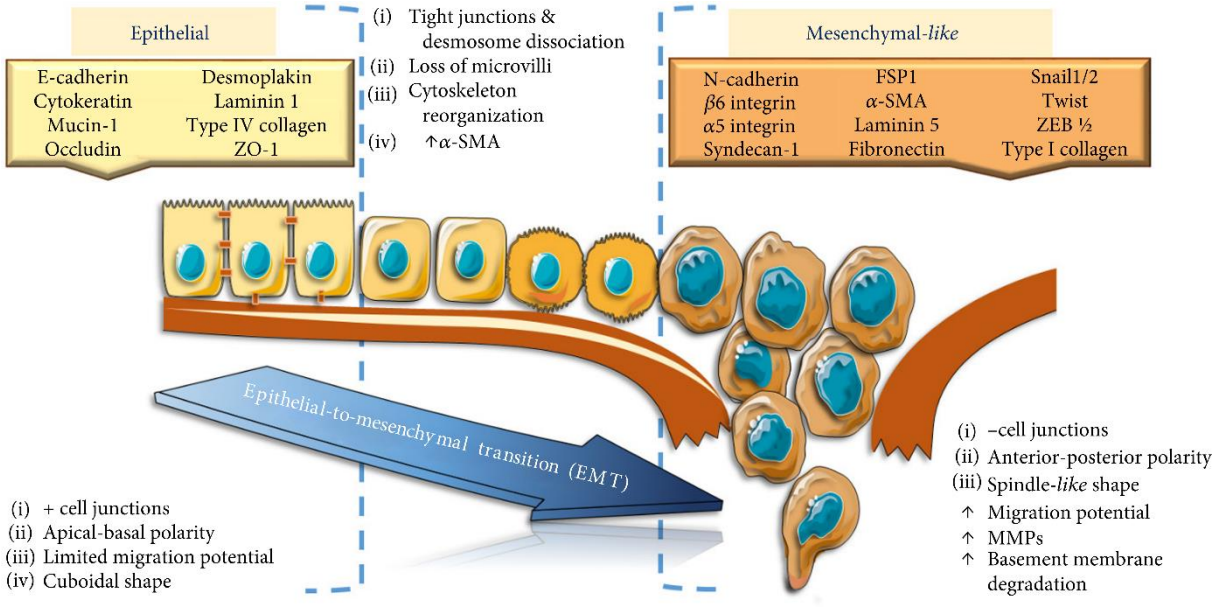

Fig. 8: The process of EMT with respective biomarkers for epithelial and mesenchymal cells[85]

2.4.2.1 Epithelial cells

By definition, an epithelial cell phenotype has strong cell-to-cell contact, allowing clusters of these cells to be well organized, enough to form tissue. The are generally less capable of invasion and migration. Tissues made up of epithelial cells often line the outer surface of organs or the inner cavities of a number of internal organs. E-cadherin (epithelial cadherin) serves as the most common epithelial factor. The loss of its function of cellular adhesion has been linked to greater tumor metastasis. [86]

2.4.2.2 Mesenchymal cells

A mesenchymal cell phenotype is characterized by its migration capability and invasive nature. These cells are often much more resistant to therapeutic approaches than epithelial cells. Tumor repopulation is possible by the migration of mesenchymal cells to different parts of the body and subsequential retransformation (MET; mesenchymal to epithelial transition) into more stable, epithelial cells. Biomarkers for mesenchymal cells include Vimentin[87] [88] [89] [90], N-cadherin[91] or Fibrionectin. Vimentin is a part of the mesenchymal cell cytoskelet, allowing for the loose binding of the cells. N-cadherin, previously responsible for cell-to-cell adhesion as E-cadherin, provides a mechanism for migration. [92]

2.4.2.3 EMT

The epithelial to mesenchymal transition (EMT) process is complex type of transformation between an originally epithelial cell type and a mesenchymal cell type. This event plays a crucial role in the progression of a variety of diseases, for example inflammation or fibrosis.[93] Frequency of the process is especially elevated not only during the migration of malignant cells undergoing invasion and metastasis, but also in drug resistance. [94] Normally, it occurs during embryogenesis as well as in cell migration during wound healing. [95] For embryogenesis, vimentin is often used as a biomarker for indication of the EMT process.

Two processes are characteristic for the transition; the loss of epithelial factors and the gain of mesenchymal factors[96]. Cancer cells with an EMT phenotype have been shown to be increasingly invasive and are more resistant to chemotherapy.[97]

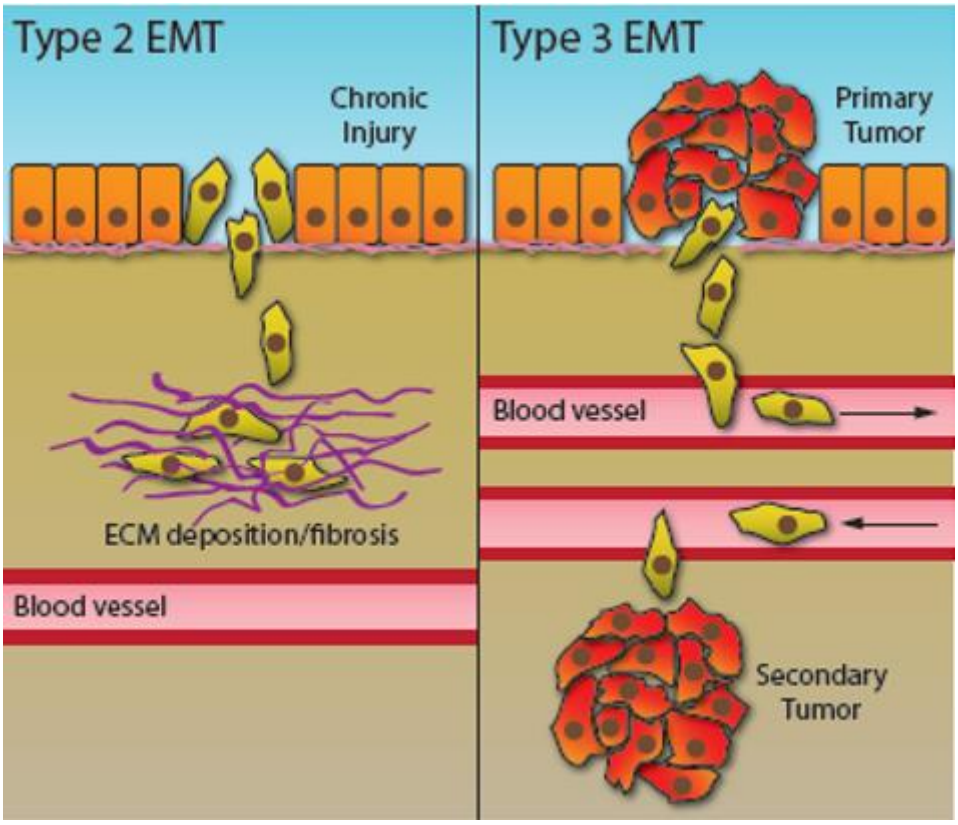

Fig. 9: The process of EMT and subsequential transport of the mesenchymal cells by blood vessels[98] *Type 2 EMT occurs in response to chronic injury (inflammation,...) and leads to tissue fibriosis. Type 3 EMT leads to the formation of tumor cells that become metastatic and form secondary tumors at distant locations*

2.4.3 CD44 (cancer stem cells)

CD44 is a is a transmembrane glycoprotein linked to cell differentiation, proliferation, migration and angiogenesis [99] It has been found to be overexpressed in solid tumors and marked as a specific biomarker together with CD133 and CD24.

In the case of CD44, studies show often opposing conclusions as to if the protein signals inhibition of promotion of cancer progression. Some have detected low levels of CD44 in cancer cells as opposed to healthy cells [100], while others have detected elevated levels of CD44 in cancerous cells [101] [102]. However, a new approach has been introduced in the previous study[103], suggesting that CD44 expressions vary in the cytoplasm fraction as opposed to the membrane fraction of the cell. [104]

Cancer stem cells (CSCs) are a small population of tumor cells that have to ability to initiate tumor growth and cause relapses. They were first identified in the late 1990s and have ever since been an inteste cancer research focus.[105] They have the same abilities as normal stem cells, meaning they have the ability to give rise to all cell types in a cancer sample, and are thus tumorigeneric. They are not often targeted through chemotherapy, thus they remain untouched and may cause relapse.

Cancer stem cells often share a similarity with cells undergoing EMT. Tumor cells undergoing an EMT may be precursors for metastatic cancer cells, or even metastatic CSCs. [106] It has been shown than the expression of known tumor-associated glycans has been enhanced in cancer wells which underwent EMT. [107]

2.4.4 PTM markers

2.4.4.1 Post translational modifications

The proteome is highly dynamic. Depending on the physiological state of the cells, the amount and type of expressed proteins is not always the same. Proteins are often altered after their biosynthesis in response to physiological conditions of the tissue where the proteins are expressed. It is believed that

post-translational modifications (PTMs) change the protein's physical and chemical properties, which may affect their cellular processes, as this determines their activity, stability, structure and distribution. Currently, more than 100 different PTMs have been described. The common PTMs include but are not limited to glycosylation, phosphorylation, sulfation, and acetylation. Few of the several hundred known PTMs have been shown to be reversible and therefore of regulatory importance in biological systems and processes.

2.4.4.1.1 Glycosylation in cancer research

In terms of occurence, glycosylation is a substantial PTM with more than 50 % of proteins being glycosylated.[108] The major types of glycans include N-linked glycans, O-linked glycans and proteoglycans.

Solid tumors are organ-like complexes embedded in an extracellular matrix which consists of a network of proteoglycans and glycoproteins such as collaged, laminin or fibronectin.[109] These proteins form a network together with surrounding cells of their microenvironment and determine the nature of the tumor. The cell itself is surrounded by a coat of glycans. These glycans mediate many biological activites, such as cell to cell adhesion and cell to cell contact. Cancer cells, just like healthy cells during embryogenesis, undergo rapid growth and adhere to other surrounding cells and cell matrices, suggesting correlation between the level of glycosylated proteins and the development and progression of cancer. Assays and studies have supported the conclusion that glycan changes are linked to numerous aspects of tumor cell behaviour and may serve as important biomarkers as well as provide a set of specific targets for therapeutical intervention. Glycoproteins are of increasing interest, as the demand for protein drugs, many of which are glycosylated, is high.[110] [111]

2.4.4.1.2 Phosphorylation in cancer research

Phosphorylation is characterized by the addition of a covalently bound phosphate group to an amino acid residue. This changes the structure of the protein, modifying its‘ interaction with other molecules and its‘ overall functions.

Phosphoproteins are of major importance in biological processes. They are responsible for functions such as signaling gene expression, cell cycle, cell adhesion, proliferation and differentiation. Protein phosphorylation could be a key determinant of cellular physiology such as early-stage cancer, however the mapping of the development of phosphoproteins in biofluids for cancer screening and diagnosis has not yet been achieved and only a few phosphoproteins have been developed as disease markers.[112]

2.4.4.1.3 Approachesfor indentifying PTMs

The characterisation of PTMs is quite demanding in terms of requirements on both sample preparation and instrumentation due to their vast structural heterogeneity. The presence of PTMs often complicates or eventually prevents the use of classical methods for protein sequence analysis. Methods utilizing differences in molecular weight of proteins may have decreased accuracy of measurement due to the presence of lipid or carbohydrate covalent attachments on proteins.

Nonetheless, post-translational processing of proteins is part of the many biological processes, and investigating their diversity is critical for understanding the mechanism of cell regulation. MS is an essential tool for detecting and mapping covalent modifications (mostly phosphorylation and dephosphorylation) and quantifying their changes. MS alone may provide evidence of modification, however further fragmentation must be performed for further analysis. MS-MS spectra provide information about the sequence of the peptides as well as the weight and position of the modified group.[113] It has been shown that ESI MS analysis of intact glycoproteins has better success over MALDI MS.[114] Upon obtaining spectra of the peptides, modified portions must be determined by the user. Available protein identification databases such as Sequest and Mascot allow for the identification of fixed and variable modifications.

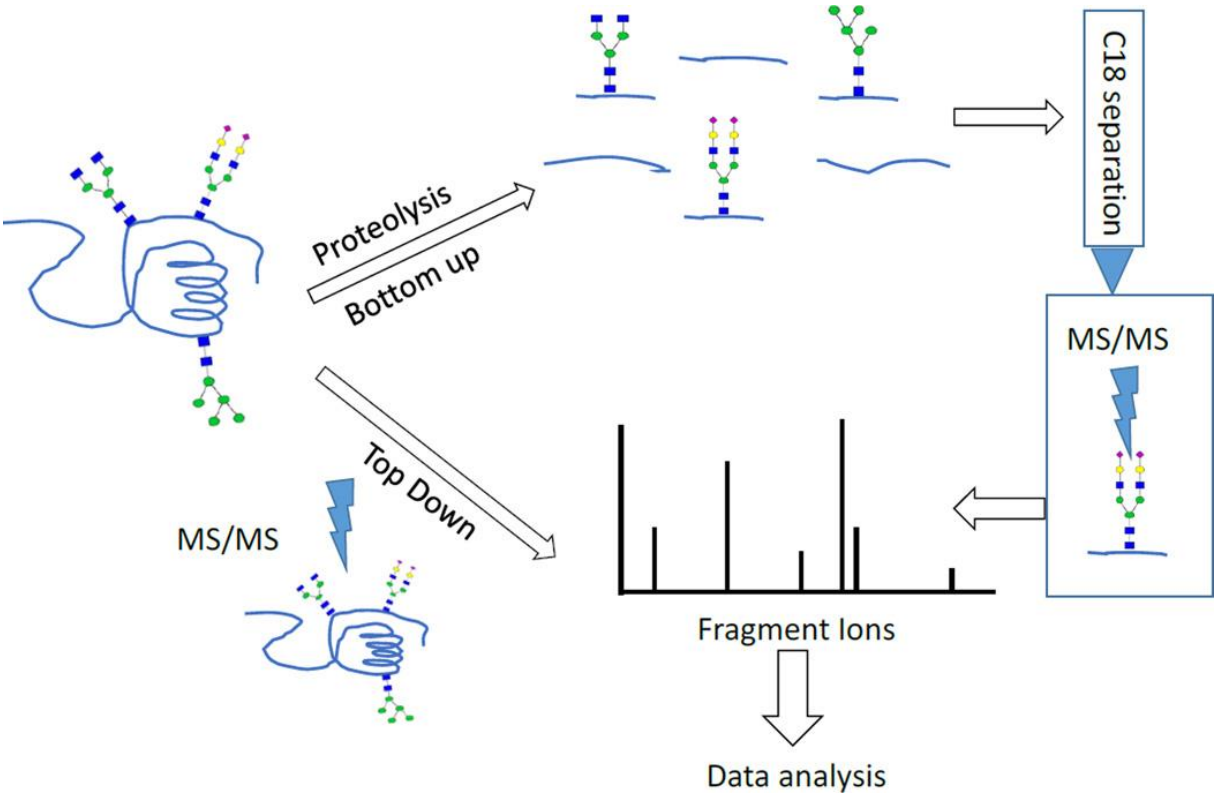

Fig. 10: The glycoprotein and glycopeptide analysis procedure[115]

2.4.4.1.3.1 Identification of glycosylated proteins

The biggest challenge in the analysis of glycoproteins is their low abundance compared to that of the unmodified proteins.[116] Mass spectral signals will thus be of low intensity, as ionization efficiency of glycoproteins will be low. In addition, distribution of their signal among various glycoforms, which share a common peptide sequence, can render their detection an overwhelming task.[116] Enrichment steps are performed in order to increase probability for detecting glycoproteins. The goal is to eliminate the most abundant unmodified proteins from competing for charge during the ionization process of MS. One of the techniques used in a proteomics laboratory are lectin affinity columns, which capture glycoproteins and can effectively enrich glycopeptides.[117] The samples can then be analyzed using methods such as MALDI-TOF or coupled HPLC with MS detection.

2.4.4.1.3.2 Identification of phosphorylated proteins

Phosphorylation posttranslational modifications may occur at low stoichiometry, so just as with glycoproteins, MS detection of phosphorylated proteins is challenged by proteins of high abundance generating an overwhelming amount of peptides following proteolytic digestion. Often, low-abundant proteins or peptides co-elute with high-abundant peptides, so the first step is always the reduction of sample complexity. Strategies for the enrichment of phosphoproteomes are necessary prior to analysis to improve its‘ detectability. Some of the enrichment methods used are strong cation exchange chromatography or metal oxide affinity chromatography.[118] The affinity of titanium oxide (TiO2) for organic phosphates was recently introduced for the purpose of selective enrichment of phosphopeptides and has been proved effective when coupled with HPLC.[119] [120]

3 AIMS

This master’s thesis analyses the proteomic profile of cell line models, namely MDA-MB-231 and MCF-7, both of which represent cancerous behaviour of cells with different characteristics. The aim is to

compare the protein map of control samples to that of samples treated with retinoids, specifically with ATRA. This master's thesis is a direct follow up on the bachelor's thesis, which concluded that the proteins Vimentin and CD44 may be viewed as potential biomarkers for breast cancer based on the significant difference of expression of these proteins in control and ATRA treated samples. Now, this study aims to explore the role of Vimentin and CD44 by further analyzing their protein profiles in terms of post-translational modifications including phosphorylation and glycosylation. Ultimately, these observations should provide a better understanding of the function of both proteins in breast cancer mechanisms.

The analysis will be carried out by using a combination of a number of protein enrichment strategies for PTM detection. Some of these strategies include TiO2 ZipTips, PhosTag gels, WGA lectin ZipTips, HPLC, and finally MS.

4 EXPERIMENTAL PART

4.1 The analyzed cell culture

For this study of breast cancer, the cell lines MDA-MB-231 and MCF-7 were used. The cell material was provided by the Slovak Academy of Sciences in Bratislava, Institute of Experimental Endocrinology, who cooperated on the study.

4.1.1 Preparation of the cell culture

The cell cultures were prepared at the Slovak Academy of Sciences, Institute of Experimental Endocrinology. The cells were grown as monolayer cultures. They were seeded in Petri dishes in Dulbecco's modified Eagle's mediem (DMEM) containing 10% fetal bovine serum (FBS) and antibiotics (penicillin, streptomycin, gentamicin), providing optimal environment for cell proliferation. Cultivation was carried out at 37 °C in a humidified atmosphere of 5% CO2 and 95% air. After incubation, the cells were washed in cold PBS and the cell lysates were then stored at -70 ºC before being sent to the Department of Analytical Chemistry of the Czech Academy of Sciences in Brno.

4.2 Protein purification and isolation

The cell lysis was conducted following an instruction manual of the RIPA buffer. The Lowry assay was used to assess protein concentrations, an equal amount of total protein was established for each cell sample. The samples were then stored at −70°C to be transported to the Department of Analytical Chemistry of the Czech Republic in Brno.

In order to desalt the cell samples, they were dialyzed against deionized water using Slide-A-Lyzer dialysis cassettes (Thermo Fisher Scientific, Waltham, MA, USA) with a 2 kDa cut off. The purified samples were then lyophilized to be stored at −70°C before the next steps..

4.3 Protein separation

Serving as a follow up to the previous study done on the cell line MDA-MB-231 (CITACE), this study aims to focus solely on the proteins Vimentin and CD44 and the analysis of their PTMs. Thus, this segment of the experimental analysis aims to isolate these two proteins using two different protein separation techniques; gel electrophoresis and HPLC.

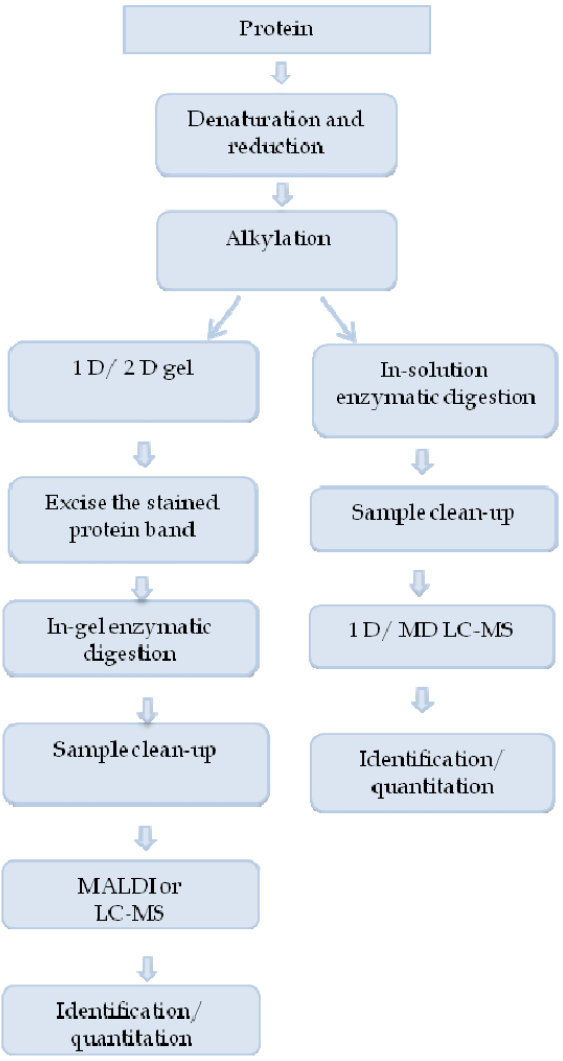

Fig. 11: Workflows of in-gel(left) and in-solution(right) digestion and subsequent LC-MS analysis on a protein sample[121]

4.3.1 1D Gel electrophoresis

The gel used for 1D gel electrophoresis consisted of two parts which had to be prepared – the running gel (12%) and the stacking gel. Pre-made gels from the company BIO-RAD were also available and were often used.

For the preparation of the gels, a few solutions had to be prepared beforehand. Solution A consisted of 30 g of acrylamide and 0,8 g of bisakrylamide topped to 100 ml with distilled water. Solution B was prepared from 10 g of SDS topped to 100 ml with distilled water. Solution C consisted of 9,1 g of TRIS dessolved in 50 ml of distilled water, which was titrated with HCl until pH 8,8 and topped to 100 ml with distilled water. Solution D consisted of 3 g of TRIS dessolved in 50 ml of distilled water, which was titrated with HCl until pH 6,8 and topped to 100 ml with distilled water. Solution E was prepared by dissolving 10 mg of bromophenol blue in 1 ml of solution D. Lastly, solution F was prepared from 1 g of ammonium persulphate topped to 10 ml with distilled water and had to be prepared fresh.

The running gel was prepared by mixing 12 ml of solution A, 15 ml of solution C, 0,3 ml of solution B and 2,4 ml of distilled water. 10 µl of of tetramethylethylendiamine (TEMED) and 75 µl of solution F were added just before pouring the gel between two glass plates and the gel was left ot polymerize. A layer of butanol was poured over the polymerizing gel to prevent oxygen access and to keep an even surface. This layer was removed before continuing to the next step.

The stacking gel was prepared by mixing 1 ml of solution A, 5 ml of solution D, 0,1 ml of solution B and 3,8 ml of distilled water. Just before pouring the gel on the polymerized running gel, 5 µl of TEMED and 0,15 ml of solution F were added. Solutions A, B, and F are identical to the separation gel. Once the focusing gel was poured, a comb was immediately placed between the two glasses to create dispensing wells in the gel.

After the gel solidified, the comb was removed and electrode buffer was poured into the electrophoresis chamber to coat the gel between the electrodes. The electrode buffer consisted of 3 g of TRIS, 14,4 g og glycine and 1 g of SDS dissolved in 1000 ml of distilled water.

The lyophilized samples were rehydrated in 180 µl of the Laemmli sampling buffer and set to boil for approximately 6 minutes together with a standard sample. Volumes of 13 µl of each sample, that is the control, ATRA treated and the standard sample, were then carefully pipetted onto the 12% SDS gel as pictured in *Fig. 12*. A constant voltage of 160 V was applied for the separation, which lasted for approximately 45 minutes. After the electrophoresis, the gel was removed from between the glass plates and into a fixing solution of trichloroacetic acid prepared from 36 g of the concentrated acid and 300 ml of distilled water. After about 10 minutes, the gel was washed with distilled water and put into a new fixing solution. The procedure was repeated 3 times. The gel was then kept overnight in the Coomassie Brilliant Blue G-250 Dye for visualization of the bands. The next day the dye was replaced by a 5% solution of acetic acid for decoloration. The solution was replaced every 10 minutes until the bands on the gel were fully visible and the gel was rid of any excess dye. A quick overview of the protein profiles in individial samples was obtained.

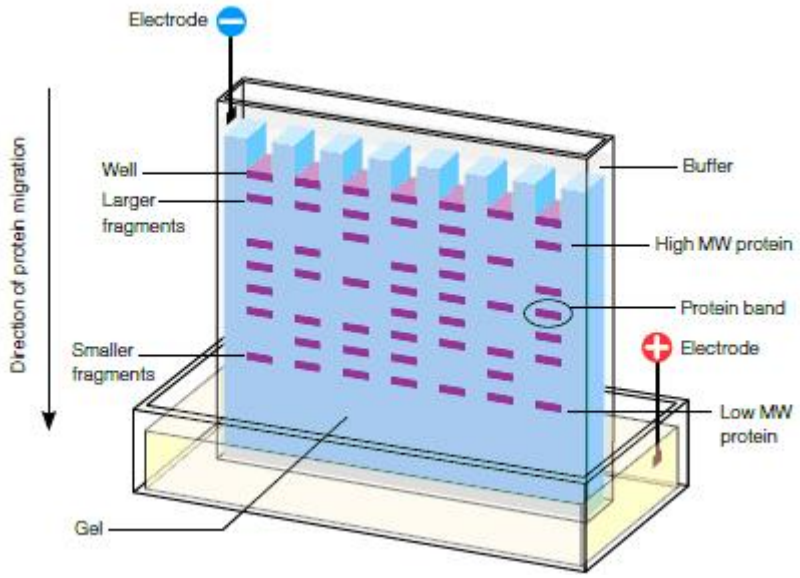

Fig. 12: Gel electrophoresis apparatus, the application of samples onto the gel.[122]

4.3.2 HPLC

Solely MDA-MB-231 cell lysates underwent separation by size exclusion chromatography (SEC) HPLC. This step followed after sample purification using 3 kDa cut-off dialysis cassettes (as seen in chapter 4.2.1). SEC-HPLC uses poreous stationary phases to trap larger molecules while letting smaller molecules pass through to pores and reach the detector quicker than the larger ones.

In this case, an Agilent Bio SEC-3 HPLC bio column containing 3 µm silica particles coated with a proprietary hydrophilic layer was used. The pore sizes used were 150 Å, and the column size was 7,8 x 150 mm. The flow rate was set to 0,5 and eventually 1 ml/min at a temperature of 25 ºC. Detection wavelength was set to 214 nm.

4.4 Sample preparation for protein identification

Before protein identification by MS, the proteins needed to be digested. By cleaving the proteins, the m/z ration is reduced, allowing for better sensibility. Enzymes have the ability to reduce the activation energy of the reaction, allowing working under mild conditions with high protein yields

4.4.1 Sample preparation after 1D gel electrophoresis

After separation, spots of the sample proteins were cut from the pure 1D and 2D gels with a scalpel. The band of the standard sample served as guidance in the 1D gel for molecular weight of the proteins in Da units. The pieces were then cut into cubes approximately 1 mm^2 in size and placed in microtubes.

4.4.1.1 Washing of the gel pieces

The gel pieces were washed twice every 15 minutes in water and a mixture of 1:1 acetonitrile and water. The amount used for washing was approximately twice the gel volume. Subsequently, the liquid was removed and the gel pieces were covered with acetonitrile, which was removed as soon as the gel precipitated and turned white. The samples then underwent centrifugation in the Eppendorf Concentrator 5301.

4.4.1.2 Reduction and alkylation

A solution of 10 mM dithiotreitol (DTT) in 0,1M ammonium bicarbonate (NH_4HCO_3) was prepared by dissolving 5 mg of DTT in 3,33 ml of NH_4HCO_3. This solution was used to coat the samples in each microtube. DTT acted as a reducing agent, preventing the formation of disulfide bonds to ensure that the proteins are unfolded and soluble, easy to purify.[123] The microtubes were then left to shake for 45 minutes at 56 °C. After the reduction, the supernatant was removed from the microtubes.

A solution of 55 mM iodoacetamide in 0,1M NH_4HCO_3 was prepared by dissolving 30 mg of iodoacetamide in 3 ml of NH_4HCO_3. The solution was used to coat the samples in each microtube and the samples were left for 30 minutes in the dark at laboratory temperature for alkylation. The supernatant was then removed.

Once reduced and alkylated, the sampled needed to be desalinated. This was achieved by adding approximately 60 µl of 100% ACN for 15 minutes. The samples were then left to dry in the Eppendorf Concentrator 5301.

4.4.1.3 Protein digestion

A solution containing 50 mM of NH_4HCO_3 and 12,5 µg/µl of trypsine was prepared by adding 1,5 ml of 50mM NH_4HCO_3 into a vial containing the enzyme. The solution was kept in the freezer for a short while before the following step.

On ice, approximately 50 µl of the trypsine solution were added to each sample for full coating. The samples were then left in the fridge for 45 minutes. Most of the trypsine solution had been absorbed by the gel pieces by then. The excess solution was removed and the samples were coated by approximately 70 µl of a 50mM NH_4HCO_3 solution. They were then left to incubate at 37 °C till the next morning.

4.4.1.4 Peptide extraction

The bulk of the peptides was released during the overnight trypsine cleavage. In order to cease the peptization process and to extract the peptides, a sufficient amount (approxiamately 60 µl) of a 5% trifluoroacetic acid (TFA) in 50% ACN solution was used to coat each sample, which were then left to incubate for 15 minutes. The supernatants containing the extracted peptides were transferred into new clean microtubes. This extraction process was repeated for another two cycles. The extracted peptides were then left to centrifuge to completely evaporate the water as well as the volatile ACN, which could interfere in the following steps.

4.4.1.5 Peptide purifica tion

An elution solution consisting of 5 ml of 0,1% TFA and 50 ml of 50% ACN was prepared. The samples were rehydrated in 10 µl of 0,1% TFA. The mixtures were then thoroughly homogenized using vortex and ultrasound.

The peptides were purified using the ZipTip C18 tips (Millipore). The tips were moistened using a wetting solution of 50% ACN and equilibrated with 0,1% TFA. Each sample was then loaded into a tip for the peptides to thoroughly bind to the media. A solution of 0,1% TFA was used for the washing of the tip in order to get rid of any potential contaminants. The peptides were eluted from the media using the elution solution prepared earlier and transferred into new clean microtubes. The purified peptides were then left to centrifuge. The remaining samples before extraction were stored for further use.

4.4.2 Sample preparation after HPLC

Just like before in-gel digestion, the disulfide bonds in proteins must be reduced, and the sulfhydryl groups must be alkylated to prevent the disulfides from re-forming. The protein samples are then incubated with trypsin for several hours to be digested into peptides.

4.4.2.1 Reduction and alkylation

A fresh 0,1M digestion buffer solution of 16 mg/ml of ammonium bicarbonate in water was prepared. Then, a solution containing 30 mg/ml of the reducing agent TCEP (Sigma D0632) in the digestion buffer was prepared. The alkylation agent used was iodoacetamide (Sigma 11149), prepared as a 0,1M solution with a concentration of 38 mg/ml in the digestion buffer.

The samples were soluted in a small amount of the digestion buffer to create sample solutions. Next, to each sample solutiom, 15 µl of the digestion buffer along with 3 µl of the reducing agent were added. The samples were left at 60 °C for an hour to reduce.

For the alkylation step, 3 µl of the iodoacetamide solution were added and the solutions were left to incubate in the dark for 20 minutes.

4.4.2.2 In-solution digestion

20 µg of lyophilized proteomics grade trypsin was dissolved in 100 µl of the digestion buffer. About 15 µl of this stock solution was added to each reduced and alkylated sample vial. The samples were left to digest at 37 °C overnight.

The next day, a TFA solution was prepared and used to stop the digestion process. This TFA solution had a concentraion of 1:9 with water and was used to coat each sample with approximately the same volume as the sample solution (about 48 µl) and then left to incubate for 20 minutes.

Following the digestion, a purification step needed to be added prior to MS. The purification of the samples was done using C18 ZipTips and a 0,1% solution of TFA together with a solution of ACN as washing solutions. The samples were then left to centrifuge.

4.5 Protein indentification

After succesful protein separation, the proteins samples underwent MS analysis in order to obatin an overview of the proteins present in both MDA-MB-231 and MCF-7. This step allowed for confirmation of the nature of the breast cancer and of the presence of the two main proteins which were the focus of this analysis; Vimentin and CD44. This step also allowed for the affirmation of the conclusion made from the previous work; ATRA treatment has a significant effect on the function of certain proteins in the breast cancer cell lines, mainly CD44, whose presence was almost completely erased after ATRA treatment.

4.5.1 Mass spectrometry

An α-Cyano-4-hydroxycinnamic acid (CHCA) matrix solution of 10 mg/ml was prepared in a 1:1 TFA:ACN extraction solution.

The samples were rehydrafed after centrifugation using a solution of TFA, applied in small quantities on the MALDI MS plate and left to crystallize. The samples were then processed in the mass spectrometer AB SCIEX TOF/TOF™ 5800 System (AB SCIEX, Framingham, MA, USA) equipped with a 1 kHz Nd:YAG laser.
The generated mass spectra were processed with the 4000 Series Explorer software and the data was submitted for search in the MASCOT database. Using the taxonomy of the NSBInr database, limited to Homo sapiens, protein identifications were assigned. Finally,the acquired proteomes of the control and ATRA treated samples from the MDA-MB-231 and MCF-7 cancer cell lines were compared.

4.5.2 Identification of potential glycosylation and phosphorylation in proteins

Individual functional properties, pathways and protein-protein interaction networks of the identified proteins were searched for on website databases such as ExPASy, STRING and UniProt. One of the characteristics analysed were protein PTMs.

For further study of the proteins Vimentin and CD44, potential glycosylation and phosphorylation peptide sites were searched for. These potential sites were to be compared with the results from MS analysis after PTM purification steps (as explained below).

4.6 Isolation steps for PTMs

As protein PTMs are low in abundance compared to their non-modified counterparts, purification steps are essential for their tracking. Moreover, many modifications are reversible and the bonds between the proteins and PTMs can be unstable.

4.6.1 Isolation of phosphorylated proteins

Phosphopeptides are difficult to analyze by MS because of lower ionization efficiency and lower relative abundance compared to non-modified peptides. Because of this, it is advised to conduct PTM purification prior to MS analysis. The techniques most commonly used for posphopeptide isolation include Immobilized metal affinity chromatography (IMAC) and TiO2 chromatography. The enrichment techniques allowed for the suppression of non-modified peptides in the mass spectra.

4.6.1.1 IMAC

The IMAC tips (Millipore) used utilised the affinity of iron ions. The workflow was based on the User Manual for ZipTip$_{MC}$ (www.millipore.com), which included steps for binding, washing and elution. A guide for the preparation of the needed solutions may be viewed in
Table 1. A solution of 0,1% acetic acid and 10% acetonitrile served as the loading buffer. The washing solutions used were the loading buffer and then Millipore grade water. Finally, elution was possible thanks to a solution of ammonium hydroxide of pH 10.

The peptide fragments were dissolved in a loading buffer and binded on tips. After the washing step (as seen in

Table 1) the phosphopeptides were eluted with ammonium hydroxide.

4.6.1.2 TiO_2 ZipTips

Commercial tips from Glygen Corporation were used. Optimization steps were necessary for the enrichment protocol by the manufacturer. *Table 1* shows the solutions prepared.

The loading buffer comprised of 20g/l DHB in 80% acetonitrile and 0,1% TFA. The washing solutions were as follows: the loading buffer, 0,1% TFA in 80% ACN and finally Millipore grade water. The elution solution was ammonium hydroxide of pH 10. The peptide fragments were dissolved in a loading buffer and binded on tips. After the washing step (as seen in

Table 1) the phosphopeptides were eluted with ammonium hydroxide

Method	Loading buffer	Washing solution	Elution solution
TiO_2 chromatography	DHB ($300\,g\,l^{-1}$) in 80% acetonitrile, 0.1% TFA	1. DHB ($20\,g\,l^{-1}$) in 80% acetonitrile, 0.1% TFA 2. 80% acetonitrile, 0.1% TFA 3. Water	Ammonium hydroxide (pH 10.5)
IMAC	0.1% acetic acid, 10% acetonitrile	1. 0.1% acetic acid, 10% acetonitrile 2. Water	Ammonium hydroxide (pH 10.5)

Table 1: The conditions for the phosphopeptides enrichment[124]

4.6.2 Isolation of glycosylated proteins

4.6.2.1 WGA LAC

The loading buffer comprised of 50mM Tris (ph 6,5), 150mM NaCl, 0,1mM $MnCl_2$, and 0,5% Triton X-100. The eluting buffer was provided by the manufacturer.

4.6.2.2 Carbon ZipTip

Glycopeptides obtained from WGA LAC needed to be purified using carbon Supel-Tips. The carbon tips were activated using 50% acetonitrile and washed with water. The sample was then carefully aspirated into the tip, washed with water and finally the bound peptides were eluted using 30% acetonitrile and directly spotted on the MALDI target.

4.6.3 Identification of protein PTMs

After the enrichment procedures, the PTM samples were spotted directly on the plate and analyzed using MALDI-TOF MS.

For the analysis of phosphopeptides, 15 mg/ml of DHB in 50% acetonitrile together with 6% phosphoric acid were used as matrix. Phosphoric acid was recommended for acidification after elution.

For the analysis of glycosylated peptides, a solution of 30 mg/ml DHB in 30% ACN and 0,1% TFA was used as the matrix.

5 DISCUSSION AND RESULTS

This study aimed to further study the proteomic profiles of MDA-MB-231 as well as MCF-7 cell lines, paying further attention to two particular proteins and their PTMs; Vimentin and CD44.

The results below include identified proteins of individual whole cell lysates, their potential (theoretical) modifications and characteristics based on online protein database research and finally the identification of these protein modifications in the samples themselves. Individual protein separation steps will also be discussed. Emphasis will also be put on the comparison of the proteomic profiles of the control and ATRA treated samples in both cell lines as well as the comparison of proteomes between the two cell lines.

5.1 Evaluation of HPLC (MDA-MB-231 cell line)

5.1.1 HPLC SEC

The first step of protein separation was SEC-HPLC. This step was conducted only for the MDA-MB-231 cell line.

The method allowed for the first overview of the proteins and their distribution in the sample based on their size. The below chromatograms are result of the measurements. The chromatograms also allow for a quick quantity analysis of the measured proteins, as the peak size directly correlates with the quantity or concentration of the detected protein.

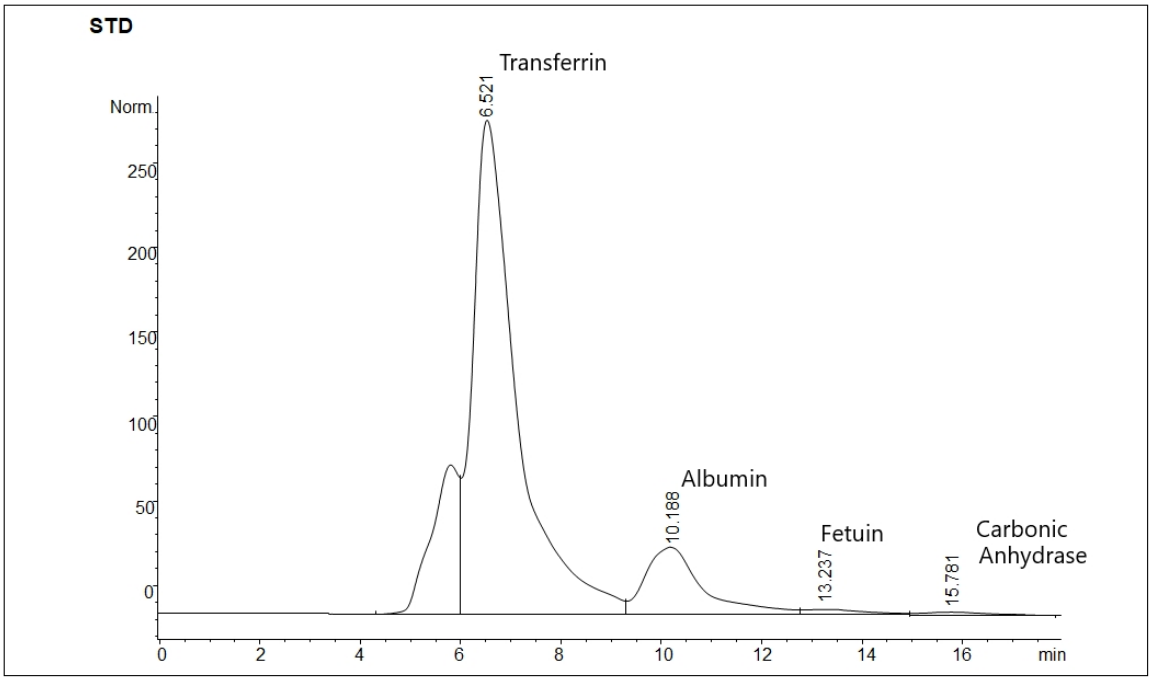

Fig. 13: Separation of the standard used, including compounds with known masses for reference

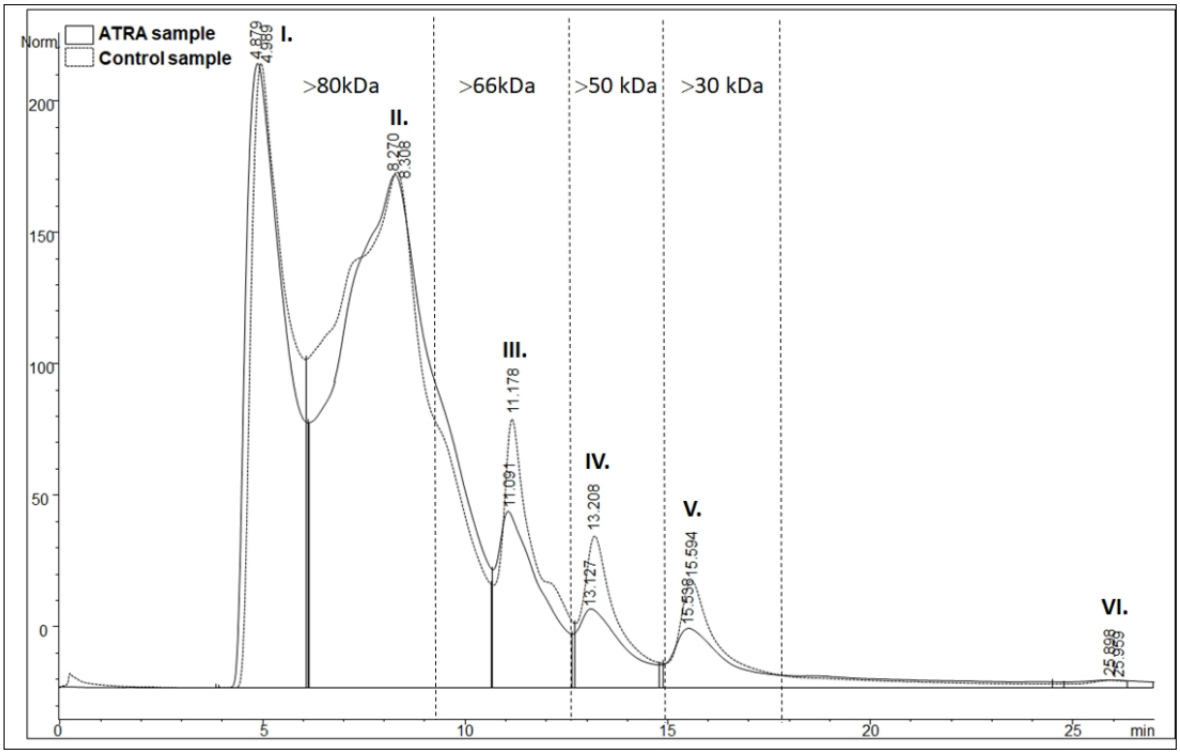

Fig. 14: Separation of ATRA and control samples with respective protein masses

The chromatogram for the standard, including the well know compounds Transferrin (80 kDa), Albumin (69 kDa), Fetuin (48.4 kDa) and carbonic anhydrase (29 kDa), served as markers for the protein molecular weights. As seen in *Fig. 14*, where the chomatograms of both ATRA and control samples are shown, significant reduction in peak sizes of the ATRA treated sample may be observed.

As peak height corresponds to the concentration of measured proteins, this indicates that the levels of proteins present in the cancerous cells have dropped after ATRA treatment. A number of these proteins may be responsible for the well-being of the cancerous nature of the cell and their reduction proves that ATRA treatment may be a valuable option for clinical use.

5.1.2 Protein identification

After separation by HPLC, the proteins underwent in-solution digestion by Trypsin and were subdued to MS analysis for protein identification by MS/MS. Below is an overview of the proteins identified together with their distribution along the chromatogram based on mass.

Each protein has been researched via protein databases, namely UniprotKB, for any potential PTMs, especially phosphorylation and glycosylation, as these will be further evalutaed in the PTM isolation steps. Both Vimentin and CD44 were found in their respective mass group.

Peak	Accession	Mr	Protein	PTMs (phosphorylation)	PTMs (glycosylation)
I	UBE4B_HUMAN	146185	Ubiquitin conjugation factor E4 B	Phosphoserine	
	URGCP_HUMAN	105675	Up-regulator of cell proliferation	Phosphoserine	
	ACTN1_HUMAN	103058	Alpha-actinin-1	Phosphoserine, phosphotyrosine	
	ENPL_HUMAN	92701	Endoplasmin	Phosphoserine, phosphothreonine	
	HS90B_HUMAN	86671	Heat shock protein HSP 90-beta	Phosphoserine,phosphothreonine, phosphotyrosine	
	CD44_HUMAN	82009	CD44 antigen	Phosphoserine	
II	NUCL_HUMAN	76626	Nucleolin	Phosphoserine, phosphothreonine	
	GRP75_HUMAN	73925	Stress-70 protein, mitochondrial	Phosphothreonine, phosphoserine	
	HSP7C_HUMAN	71086	Heat shock cognate 71 kDa protein	Phosphoserine	
	BIP_HUMAN	74162	Endoplasmic reticulum chaperone BiP	Phosphoserine, phosphothreonine	
	MOES_HUMAN	69651	Moesin	Phosphoserine,phosphothreonine, phosphotyrosine	
	EZRI_HUMAN	69413	Ezrin	Phosphotyrosine,phosphoserine, phosphothreonine	
	CALX_HUMAN	67990	Calnexin	Phosphoserine, phosphothreonine	
III	SHC1_HUMAN	62822	SHC-transforming protein 1	Phosphoserine, phosphotyrosine	
	K1C10_HUMAN	62394	Keratin, type I cytoskeletal 10	Phosphoserine	
	TCPG_HUMAN	61792	T-complex protein 1 subunit gamma	Phosphoserine,phosphotyrosine, phosphotyrosine	
	IL2RB_HUMAN	61117	Interleukin-2 receptor subunit beta		N-linked asparagine
	KPYM_HUMAN	59197	Pyruvate kinase PKM	Phosphoserine, tyrosine, threonine	
	VIME_HUMAN	55492	Vimentin	Phosphoserine, threonine, tyrosine	O-linked serine, threonine
	PAK4_HUMAN	54940	Serine/threonine-protein kinase PAK 4	Phosphoserine, threonine	
	TBB1_HUMAN	52168	Tubulin beta-1 chain	Phosphoserine	
	EF1A1_HUMAN	51838	Elongation factor 1-alpha 1	Phosphoserine, threonine	
	TBA1B_HUMAN	50804	Tubulin alpha-1B chain	Phosphoserine	
	TBA1A_HUMAN	50136	Tubulin alpha-1A chain	Phosphoserine	

	TBB4A_HUMAN	50010	Tubulin beta-4A chain	Phosphoserine	
IV	PDIA6_HUMAN	49676	Protein disulfide-isomerase A6	Phosphoserine	
	ACTBL_HUMAN	44282	Beta-actin-like protein 2		
	KLRG2_HUMAN	43508	Killer cell lectin-like receptor subfamily G member 2	Phosphoserine	
	ACTA_HUMAN	42388	Actin, aortic smooth muscle		
	ACTB_HUMAN	42058	Actin, cytoplasmic 1		
	ROA1_HUMAN	38839	Heterogeneous nuclear ribonucleoprotein A1	Phosphoserine	
	ROA2_HUMAN	37465	Heterogeneous nuclear ribonucleoproteins A2/B1	Phosphothreonine, serine, tyrosine	
	G3P_HUMAN	36201	Glyceraldehyde-3-phosphate dehydrogenase	Phosphotyrosine, serine, threonine	N-beta-linked arginine
	RPC6_HUMAN	35684	DNA-directed RNA polymerase III subunit RPC6		
	HNRPC_HUMAN	33708	Heterogeneous nuclear ribonucleoproteins C1/C2	Phosphoserine, threonine	
	NPM_HUMAN	32729	Nucleophosmin	Phosphoserine, tyrosine, threonine	
	VDAC2_HUMAN	32069	Voltage-dependent anion-selective channel protein 2	Phosphoserine, tyrosine, threonine	
V + VI	SRS10_HUMAN	31340	Serine/arginine-rich splicing factor 10	Phosphoserine	
	PRDX3_HUMAN	28023	Thioredoxin-dependent peroxide reductase, mitochondrial	Phosphothreonine	
	SRSF7_HUMAN	27582	Serine/arginine-rich splicing factor 7	Phosphoserine	
	SODM_HUMAN	24909	Superoxide dismutase [Mn], mitochondrial		
	RL14_HUMAN	23533	60S ribosomal protein L14	Phosphoserine	
	PPR1B_HUMAN	22963	Protein phosphatase 1 regulatory subunit 1B	Phosphoserine, threonine	
	PRDX1_HUMAN	22328	Peroxiredoxin-1	Phosphothreonine, serine	
	SRSF3_HUMAN	19550	Serine/arginine-rich splicing factor 3	Phosphoserine	
	PPIA_HUMAN	18229	Peptidyl-prolyl cis-trans isomerase A	Phosphoserine, threonine	N-linked asparagine
	NDKA_HUMAN	17149	Nucleoside diphosphate kinase A	Phosphoserine	
	PROF1_HUMAN	15216	Profilin-1	Phosphoserine, tyrosine	
	H2A1A_HUMAN	14225	Histone H2A type 1-A	Phosphoserine, threonine	
	H4_HUMAN	11360	Histone H4	Phosphoserine, tyrosine, threonine	

Table 2: Indentified proteins separated by HPLC SEC. As seen in Fig. 13, section I. and II. correspond to masses greater than 80 kDa, section III. corresponds to masses greater than 66 kDa, section IV. corresponds to masses greater than 50 kDa, and section V. And VI. correspond to proteins of the lowest masses

5.2 Evaluation after 1D gel electrophoresis

In this workflow, both MDA-MB-231 and MCF samples of control and ATRA were used for comparison. Special emphasis was put on two proteins, Vimentin and CD44, and their PTMs

The results acquired include gel scans from 1D SDS and tables of identified proteins from MS analysis together with an exploration of the potential PTMs of the identified proteins.

5.2.1 1D SDS gel electrophoresis sample separation

1D SDS gel electrophoresis was conducted for whole cell lysates of control and ATRA treated samples for both MDA-MB-231 and MCF-7 cell lines. Protein separation by this method provided a quick overview of the protein distribution in the samples based on their molecular weight. *Fig. 15* shows the gel scan of the MDA-MB-231 cell line together with that of MCF-7 as well as a list of some of the anticipated proteins (to be identified in the MS protein identification step), distributed by mass based on the standard.

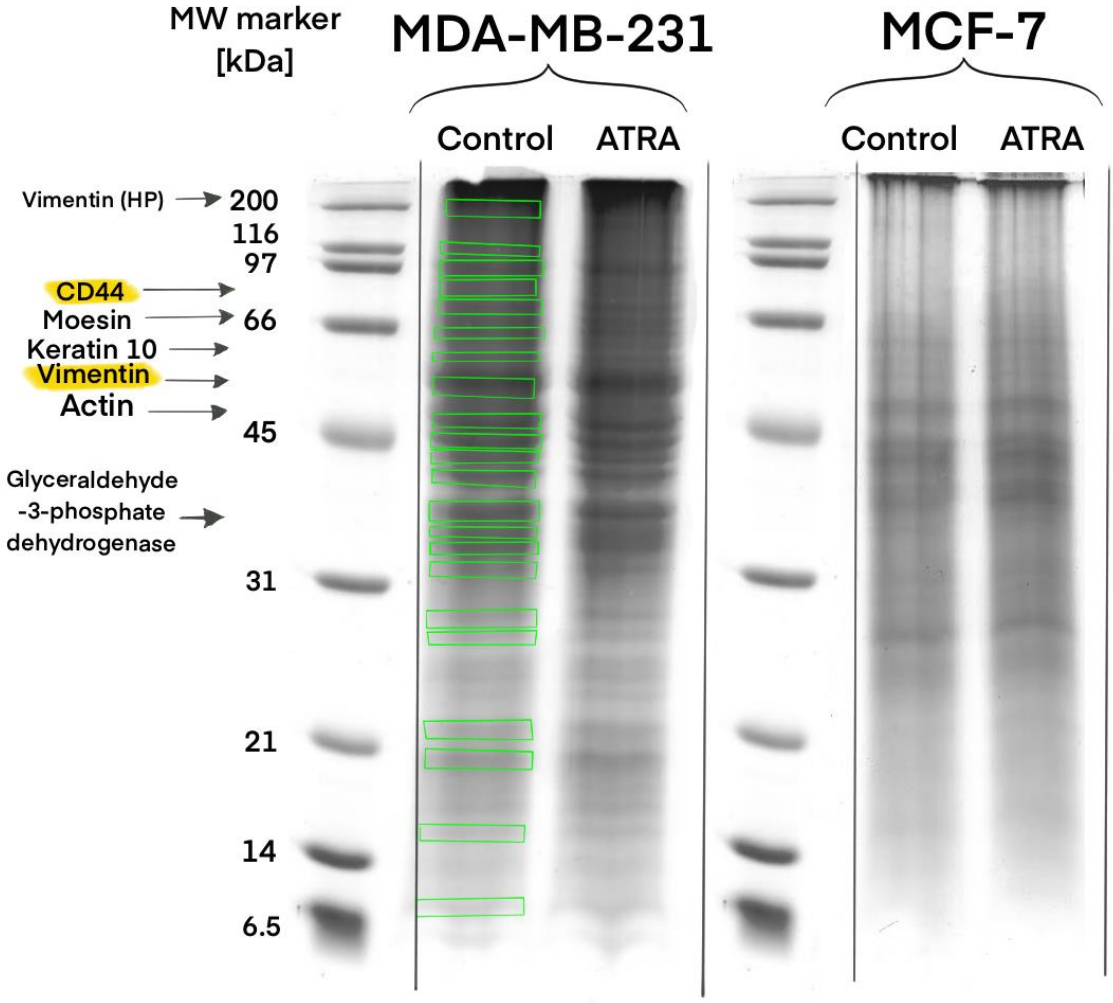

Fig. 15: Basic overview of the separated proteins using 1D SDS gel electrophoresis on whole cell lysates. The anticipated proteins CD44 and Vimentin may be spotted. Vimentin has a mass of 53681 Da, while CD44 has a mass of 82009 Da

5.2.2 Protein identification

After protein separation, the proteins were in-gel digested using Trypsin and underwent a number of PTM enrichment steps for better MS sensitivity and PTM identification. The PTM isolation steps included IMAC and TiO_2 ZipTips for phosphopeptides and WGA lectin affinity chromatography for glycopeptides.

Before MS analysis, Vimentin and CD44 peptide sequences were evaluated using protein and peptide databases to look for potential modifications, which would then be anticipated in MS analysis after PTM enrichment. The anticipated modifications in peptides together with the experimentally discovered modifications may be seen in *Fig. 21* for Vimentin and in *Fig. 24* for CD44.

MS analysis generated a number of peptide maps and identified proteins from whole cell lysates. The MS data was evaluated using website databses such as ExPASy, STRING and UniProt. From the information gathered, it was possible to group a few proteins with interesting roles with regards to cancer. These proteins showed a connection to apoptosis, angiogenesis or any other contribution to tumor growth (see *Table 4: Selected identified proteins from the control (cancerous) sample and their respective roles in the organism.*). This however represents only a small fraction of the identified proteins, as a thorough study of the whole proteome would be too time-consuming. Both proteins of interest, vimentin and CD44, were successfully identified in the samples.

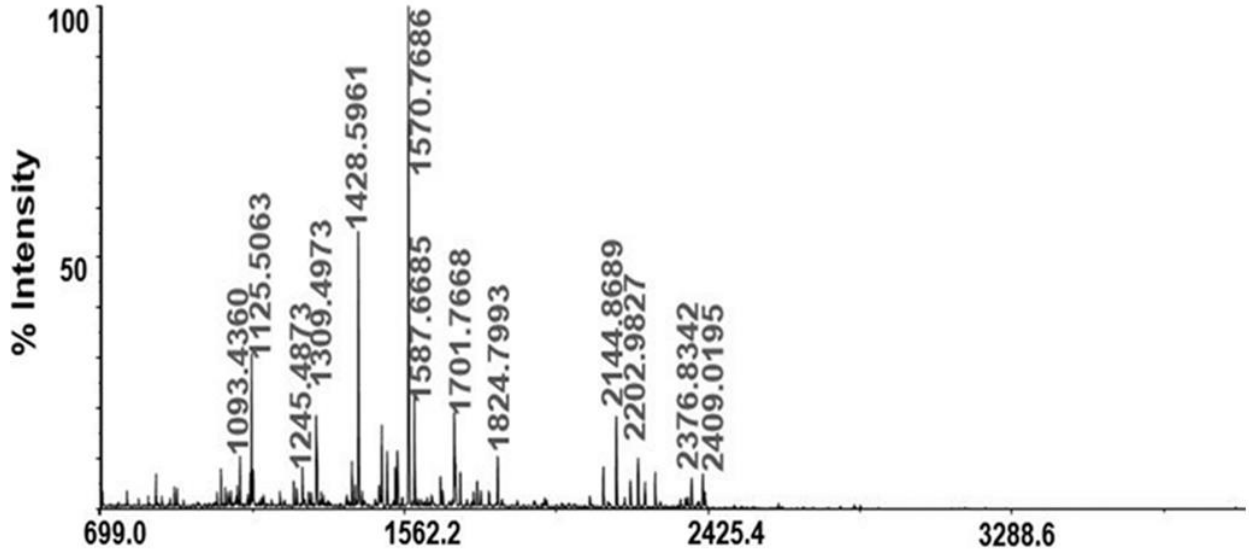

Fig. 16: The chosen MS spectrum of identified vimentin found in MDA-MB-231

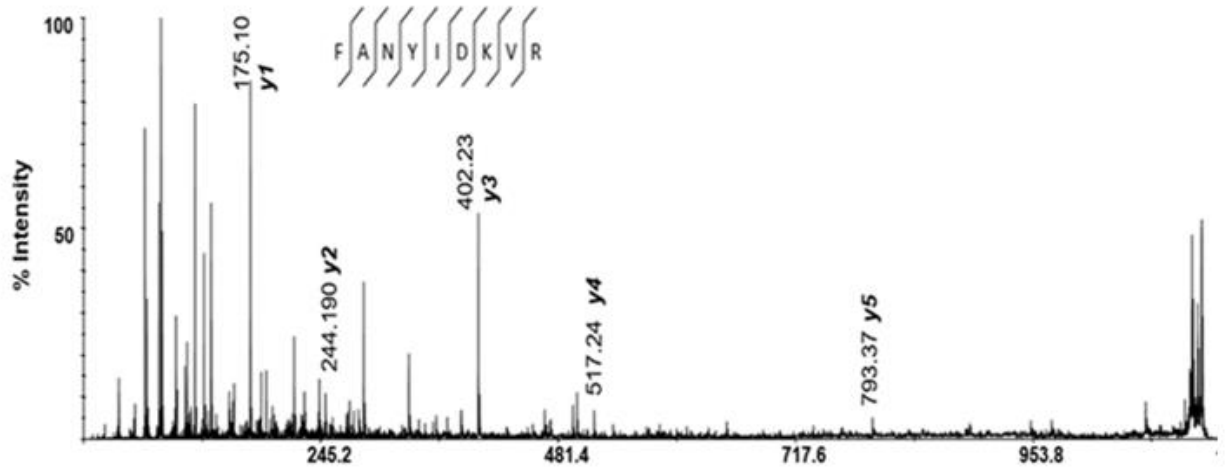

Fig. 17: An example of fragmentation MS/MS spectrum at 1125 m/z corresponding to Vimentin

The figures above serve as a presentation of proteomic MS spectra, with *Fig. 16* being the MS spectrum of Vimentin found in the MDA-MB-231 cell line. Based on the molecular weights measured with each significant peak, these peaks are identified as peptides, allowing for the identification of the protein through database searches. *Fig. 17* acts as an example of a peptide MS/MS spectrum of Vimentin, allowing for their amino acid sequences to be generated.

MCF7	MDA-MB-231	Protein	Mass (Da)	UniProt KB number
Y	Y	calreticulin=calcium binding protein	3737	
Y		similar to heat shock 70kDa protein 8 isoform 2; heat shock cognate protein	13379	A4D110
Y		cytokeratin 8 (279 AA)	30840	P05787
	Y	40S ribosomal protein SA	32833	P08865
Y	Y	ribosomal protein L6	32841	Q02878
Y	Y	Chain R, Twinning In Crystals Of Human Skeletal Muscle D-Glyceraldehyde-3- Phosphate Dehydrogenase	35853	P04406
	Y	heterogeneous nuclear ribonucleoproteins A2/B1 isoform A2	35984	P22626
Y	Y	glyceraldehyde-3-phosphate dehydrogenase	36031	P04406
Y		actin prepeptide, partial	36783	Q13707
Y		fructose-1,6-bisphosphatase	36805	P09467
	Y	annexin A2	38552	P07355
	Y	annexin A1	38690	Q5TZZ9,P04083
Y	Y	aldolase A	39307	P04075
	Y	keratin 10, partial	39718	P13645
Y	Y	actin, cytoplasmic 1	41710	Q1KLZ0,P60709
Y	Y	beta-actin-like protein 2	41976	Q562R1
Y	Y	actin, alpha skeletal muscle	42024	P68133
Y		Chain A, Crystal Structure Of A Heat Shock 70kda Protein 2	42130	P54652
Y		40-kDa keratin protein, partial	44065	P08727
Y		40-kDa keratin protein, partial	44079	P08727
Y	Y	phosphoglycerate kinase 1	44586	P00558
	Y	actin-related protein 2 isoform b	44732	P61160
	Y	eukaryotic initiation factor 4AII	46365	Q14240
Y	Y	alpha-enolase isoform 1	47139	P06733
Y		cytokeratin 18 (424 AA)	47305	P05783
	Y	calreticulin precursor	48112	P27797
	Y	elongation factor Tu	49509	P49411
Y	Y	tubulin beta-5 chain	49639	P99024
Y	Y	tubulin beta-4B chain	49799	P68371
Y		glial fibrillary acidic protein isoform 1	49850	P14136
Y	Y	elongation factor 1-gamma	50087	Q53YD7,P26641
Y	Y	tubulin alpha-1B chain	50120	P05213
Y	Y	transformation upregulated nuclear protein	51040	P61978

Y		unnamed protein product	51727	B4DRW1
Y		peptidyl-prolyl cis-trans isomerase FKBP4	51772	Q02790
Y		cytokeratin 8	53529	P05787
	Y	vimentin	53681	P08670
Y		glucose-6-phosphate dehydrogenase	54789	Q2VF42
	Y	unnamed protein product	54987	B4E1Q1
Y		UDP-glucose 6-dehydrogenase isoform 1	54989	O60701
Y		ATP synthase subunit beta, mitochondrial precursor	56525	P06576
Y		thyroid hormone binding protein precursor	57069	P07237
	Y	lamin A protein, partial	57686	P02545
Y	Y	pyruvate kinase	57841	P14618
Y		Keratin 10	58792	P13645
Y	Y	60 kDa heat shock protein, mitochondrial	61016	P10809
Y		unnamed protein product	63885	P08107
Y		keratin 1	65978	P04264
Y		growth regulated nuclear 68 protein	66881	
Y	Y	calnexin precursor	67526	P27824
Y		transketolase	67751	P29401
	Y	moesin	67778	P26038
Y	Y	HSP90AA1 protein	68329	Q2VPJ6
Y		dolichyl-diphosphooligosaccharide--protein glycosyltransferase subunit 1 precursor	68527	P04843
	Y	X-ray repair cross-complementing protein 6 isoform 1	69799	P12956,B1AHC8
Y	Y	HSP70-2	69982	P08107
Y	Y	heat shock cognate 71 kDa protein isoform 1	70854	P11142,Q53HF2
Y	Y	heat shock 70 kDa protein 6	70984	P17066
Y	Y	serum albumin	71316	P02768
Y	Y	GRP78 precursor, partial	72071	P11021
Y		myotubularin-related protein 14 isoform 2	72158	Q8NCE2
Y		protein disulfide-isomerase A4 precursor	72887	P13667
Y	Y	Heat shock 70kDa protein 9 (mortalin)	73682	P38646
Y	Y	tumor necrosis factor type 1 receptor associated protein TRAP-1 - human	75295	
	Y	calelectrin	75857	P08133
Y	Y	splicing factor, proline- and glutamine-rich	76102	Q86VG2,P23246
Y	Y	nucleolin	76298	P19338
	Y	heat shock protein 70	78945	P34932
Y	Y	CD44	82009	P08670
Y	Y	heat shock protein HSP 90-beta isoform a	83212	P08238
Y		nuclear corepressor KAP-1	88479	Q13263

Y	Y	heterogeneous nuclear ribonucleoprotein U isoform b	88924	Q00839
Y	Y	transitional endoplasmic reticulum ATPase	89266	P55072,Q96IF9
Y	Y	endoplasmin precursor	92411	P14625
Y	Y	elongation factor 2	95277	P13639
	Y	importin subunit beta-1 isoform 1	97108	Q14974
	Y	100 kDa coactivator	99628	Q7KZF4
Y	Y	alpha actinin 4	102204	O43707
	Y	alpha-actinin-1 isoform b	102993	P12814
	Y	alpha-actinin	103229	Q08043
Y	Y	Glucosidase II	106833	Q14697
Y	Y	ubiquitin activating enzyme E1	117715	P22314
	Y	exportin-1	123306	B3KWD0,O14980
	Y	karyopherin beta 3	123512	O00410
	Y	myosin-9	226392	P35579
	Y	talin	269550	Q9Y490
Y		fatty acid synthase	272919	P49327
Y	Y	filamin A	277332	Q60FE6
Y	Y	actin-binding protein homolog ABP-278	278018	O75369
	Y	plectin 1, intermediate filament binding protein 500kDa, isoform CRA_c	289821	
	Y	plectin	518173	Q15149
Y		protein kinase PKNbeta		Q6P5Z2

*Table 3: *Y = the protein was identified in the cell line. An overview of the identified proteins after gel electrophoresis separation and a comparison of their expression between MDA-MB-231 and MCF-7 cell lines.*

The table above displays the various identified proteins in both cell lines compared. Among the proteins to be noted, moesin, found solely in the MDA-MB-231 cell line, plays a role in cancer cell invasiveness as well as in the process of metastasis characteristic for the ER-negative breast cancers (Carmeci et al. 1998). Significant reduction of the protein and other migration related proteins was found after the use of ATRA, causing an inhibition in cell migration (Flamini et al. 2014).

Concerning the proteins found in the MCF-7 cell line, fatty acid synthase espression seems to be regulated by the G protein coupled estrogen receptor in fibroblasts associated with cancer, which contribute to the progression of cancer (Santolla et al. 2012).

Some well know breast cancer markers may also be found, such as annexin A1 and A2. Annexin A1 seems to act as a suppressor of tumour in cells (Ang et al. 2009). Increase of expression of annexin A2 was equally frequently observed in a number of cancer cells (Wang and Lin 2014).

The presence of vimentin in the MDA-MB-231 and absence in the counterpart MCF-7 has been observed in previous studies (Ivaska et al. 2007, Flodrova et al 2012).

Below is a selection of proteins interesting in the context of cancer progression together with their molecular roles.

Accession	Mass	Protein	Biological process	Molecular function
FINC_HUMAN	272320	Fibronectin	*Involved in cell adhesion, cell motility, opsonization, wound healing, and maintenance of cell shape.* ***Superfibronectin (Fibronectin polymer) inhibits tumor growth, angiogenesis and metastasis.***	*Binds cell surfaces and various compounds including collagen, fibrin, heparin, DNA, and actin.*
IMB1_HUMAN	98442	Importin subunit beta-1	***Apoptotic*** *DNA fragmentation, protein localization(transport), host-virus interaction,* ***mitotic*** *congression and assembly.*	*Enzyme binding, HSP90 protein binding, Ran GTPase binding.*
CD44_HUMAN	82009	CD44 antigen	*N-Glycosylation, Phosphorylation, Proteoglycan.*	*Cell-cell interactions, cell adhesion and migration, helping them to sense and respond to changes in the tissue microenvironment, activation, recirculation and homing of T-lymphocytes, hematopoiesis, inflammation and response to bacterial infection, platform for signal transduction by assembling.*
NUCL_HUMAN	76626	Nucleolin	***Major nucleolar protein in growing eukaryotic cells, angiogenesis***	*RNA binding, DNA topoisomerase binding(enzyme changing the tertiary structure of DNA).*
ANXA6_HUMAN	76174	Annexin A6	***Apoptotic signaling pathway****, ion transmembrane transport.*	*Calcium ion binding, actin filament binding.*
MOES_HUMAN	67894	Moesin	*Phosphorylation, Acetylation.*	*Connects the actin cytoskeleton to the plasma membrane and thereby regulates the structure and function of specific domains of the cell cortex; regulation of cell size and shape, Host-virus process.*
K2C1_HUMAN	66173	Keratin, type II cytoskeletal 1	*Keratinization,* ***regulation of angiogenesis (blood vessels).***	*Carbohydrate binding, signaling receptor activity.*
IL2RB_HUMAN	61117	Interleukin-2 receptor subunit beta	*Host-virus interaction(regulation of phagocytosis,* ***apoptotic process****).*	*Receptor for interleukin-2(regulation for imunity).*
PAK4_HUMAN	54940	Serine/threonine-protein kinase PAK 4	***Apoptosis****, cell cycle (by controlling levels of the cell-cycle regulatory protein CDKN1A and by phosphorylating RAN).*	*Protein serine/threonine kinase actibvity(binds the phosphate group of a high-energy molecule ATP to another molecule), ATP binding.*
VIME_HUMAN	53677	Vimentin	***Maintenance of the cytoarchitecture and tissue integrity****, Involved in the formation of signaling complexes with cell signaling molecules.*	*Class-III intermediate filaments found in various non-epithelial cells.*

ANXA2_HUMAN	38812	Annexin A2	*Angiogenesis.*	*Calcium ion binding, calcium dependent protein binding, protease and virion binding.*
NPM_HUMAN	32729	Nucleophosmin	*Ribosome biogenesis, centrosome duplication, protein chaperoning, histone assembly, cell proliferation, and regulation of tumor suppressors p53/TP53 and ARF.*	*Chaperone, RNA-binding.*
VDAC1_HUMAN	30870	Voltage-dependent anion-selective channel protein 1	*In the plasma membrane it is involved in cell volume regulation and apoptosis; anion transport (membrane).*	*Ion channel binding, protein kinase binding.*
PRDX1_HUMAN	22328	Peroxiredoxin-1	*Cell protection against oxidative stress by detoxifying peroxides, Might participate in the signaling cascades of growth factors and tumor necrosis factor-alpha by regulating the intracellular concentrations of H2O2. Natural killer cell activation.*	*Thiol-specific peroxidase that catalyzes the reduction of hydrogen peroxide and organic hydroperoxides to water and alcohols; peroxidase activity.*
CASPC_HUMAN	19136	Inactive caspase-12	*May reduce cytokine release in response to bacterial lipopolysaccharide during infections, apoptotic process.*	*Cysteine-type endopeptidase inhibitor activity.*
NDKA_HUMAN	17149	Nucleoside diphosphate kinase A	*Cell proliferation, differentiation and development. This protein is found in reduced amount in some tumor cells of high metastatic potential, neurogenesis.*	*Kinase, transferase.*

Table 4: Selected identified proteins from the control (cancerous) sample and their respective roles in the organism.

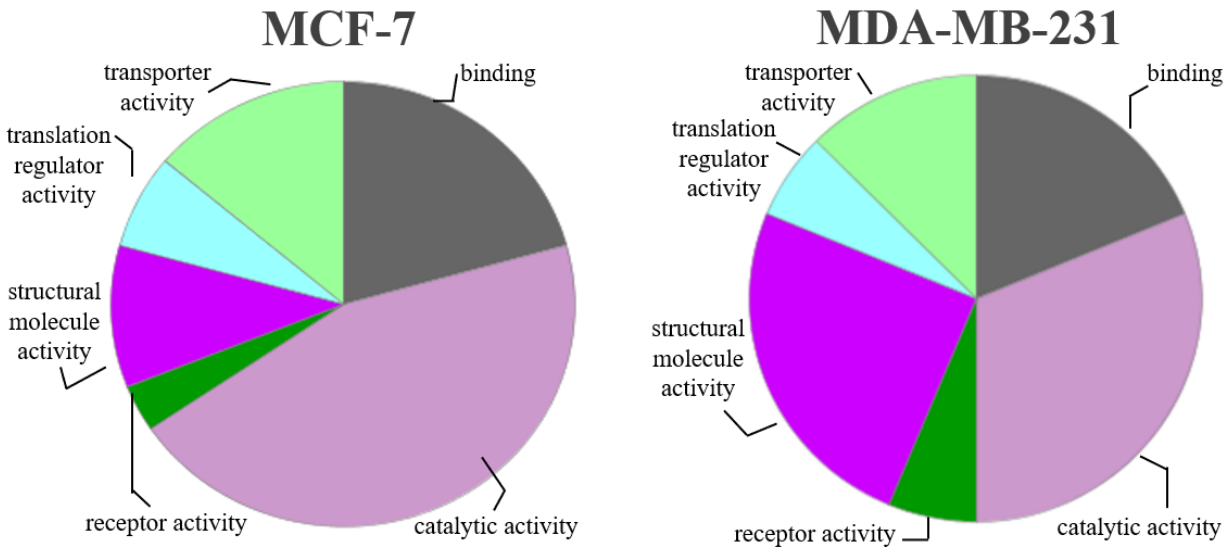

Fig. 18: Proteins categorization based on major biological, molecular functions and their potential association networks with known and predicted proteinï protein interactions

The data above show a proteomic overview of the cancer cell lines as well as the major biological and molecular functions of their proteins. As not many scientists have taken into account the protein differences between these two cancer cell lines, emphasis was put on the description of moleuclar functions of each protein.

A study by Strouhalová. D. et al. compared these two cell lines and revealed almost 100 proteins from both cell lines, with 30 of them seeming to be either up-regulated or specific to MDA-MB-231 in comparison to MCF-7 cells. Basic proteomic analysis can thus serve as a comparison method of protein patterns of cell lines, allowing easier decisionabout usage of the cell lines based on the similarities or dissimilarities.[cxxv]

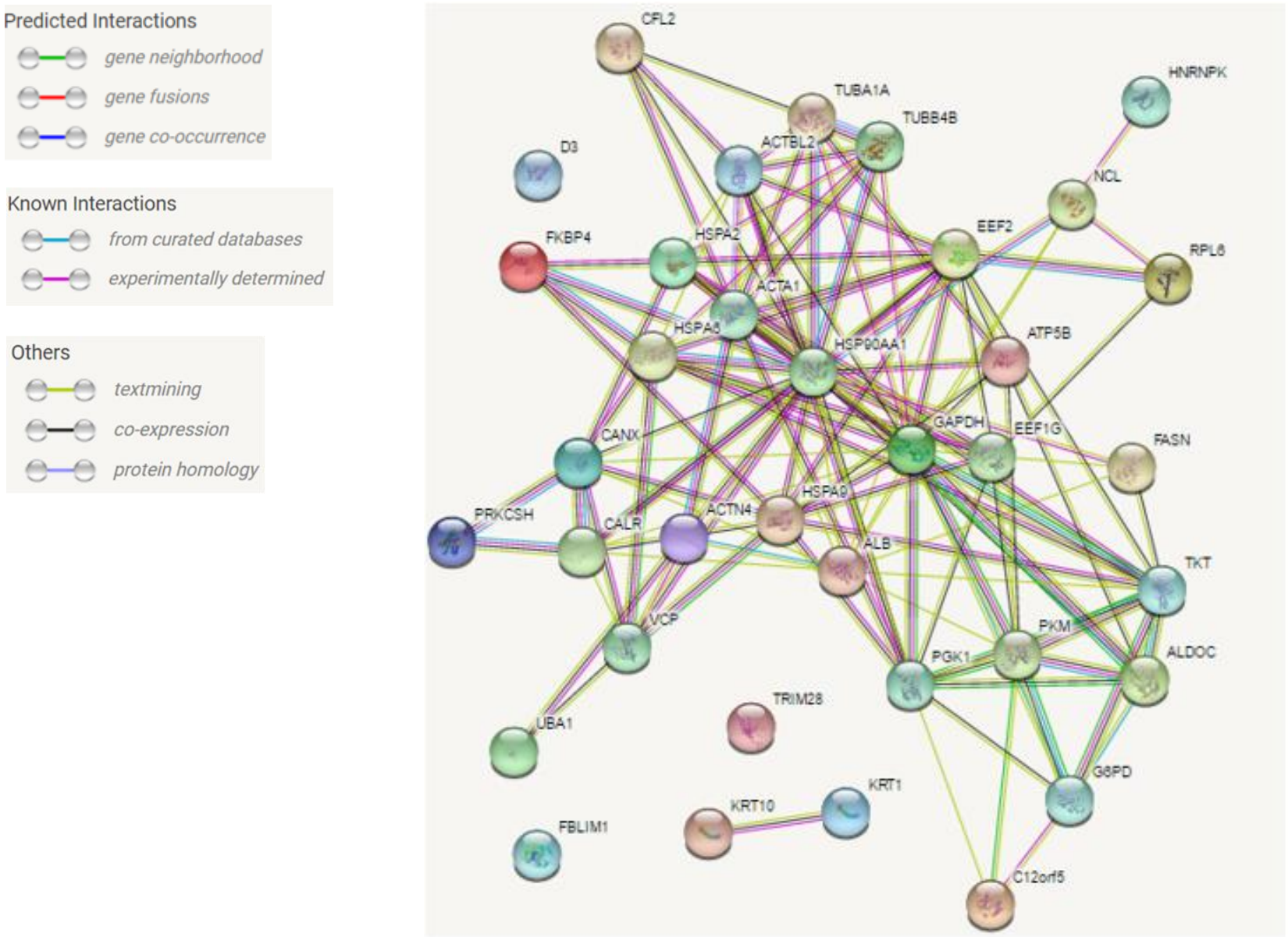

Fig. 19: An overview of protein-protein interactions in MCF-7, generated from STRING database

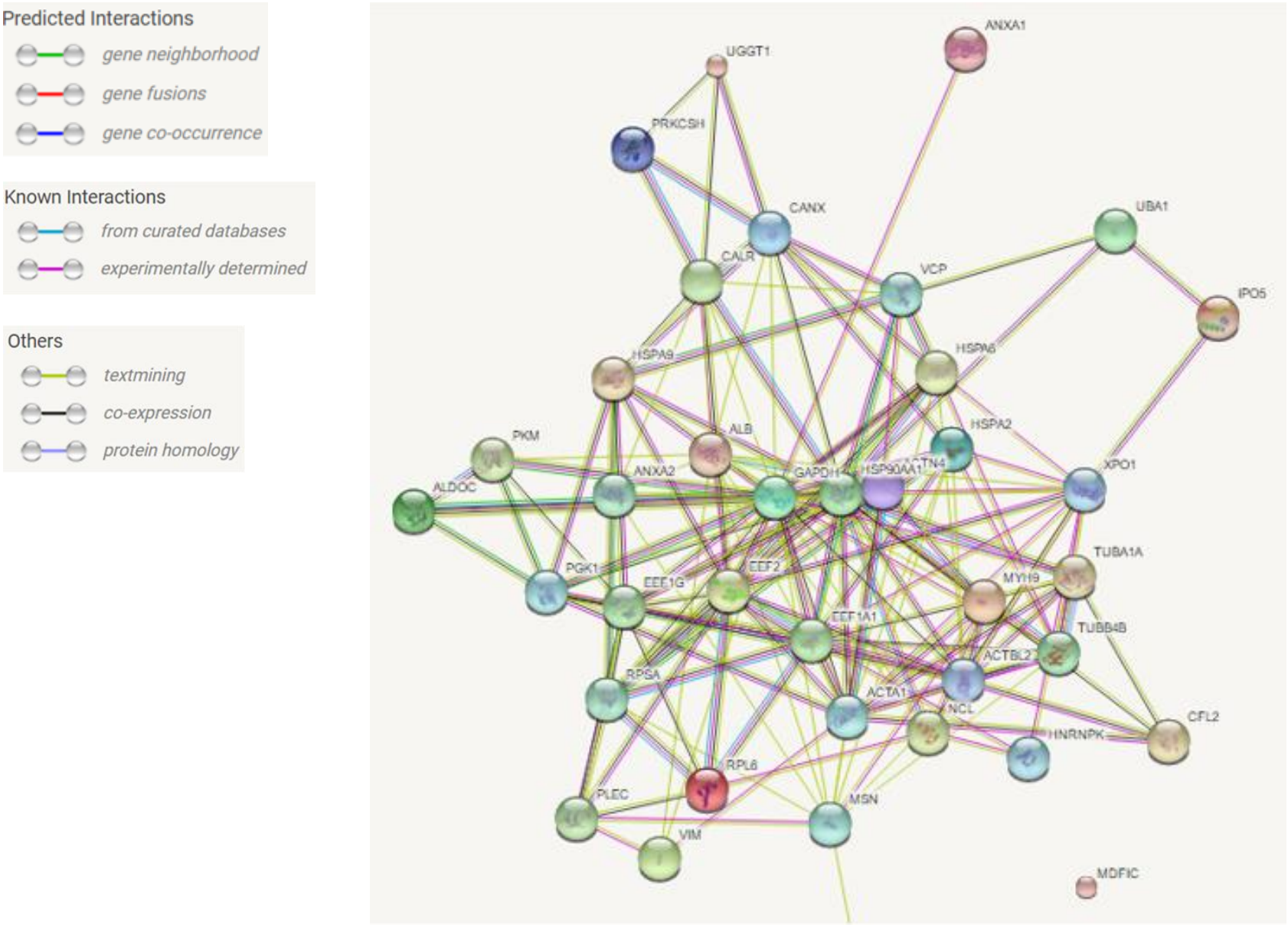

Fig. 20: An overview of protein-protein interactions in MDA-MB-321, generated from STRING database

1	MSTRSVSSSS	YRRMFGGPGT	ASRPSSSRSY	VTTSTRTYSL	GSALRPSTSR
51	SLYASSPGGV	YATRSSAVRL	RSSVPGVRLL	QDSVDFSLAD	AINTEFKNTR
101	TNEKVELQEL	NDRFANYIDK	VRFLEQQNKI	LLAELEQLKG	QGKSRLGDLY
151	EEEMRELRRQ	VDQLTNDKAR	VEVERDNLAE	DIMRLREKLQ	EEMLQREEAE
201	NTLQSFRQDV	DNASLARLDL	ERKVESLQEE	IAFLKKLHEE	EIQELQAQIQ
251	EQHVQIDVDV	SKPDLTAALR	DVRQQYESVA	AKNLQEAEEW	YKSKFADLSE
301	AANRNNDALR	QAKQESTEYR	RQVQSLTCEV	DALKGTNESL	ERQMREMEEN
351	FAVEAANYQD	TIGRLQDEIQ	NMKEEMARHL	REYQDLLNVK	MALDIEIATY
401	RKLLEGEESR	ISLPLPNFSS	LNLRETNLDS	LPLVDTHSKR	TLLIKTVETR
451	DGQVINETSQ	HHDDLE			

Phosphorylated
Glycosylated
PTMs detected by MS

Fig. 21: Vimentin peptide sequence with theoretical PTM sites and detected PTM sites.

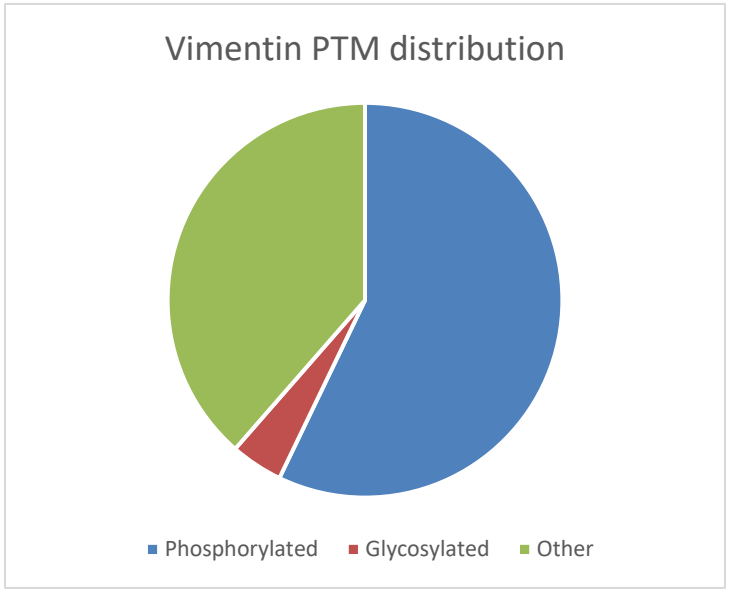

Fig. 22: Vimentin PTM distribution

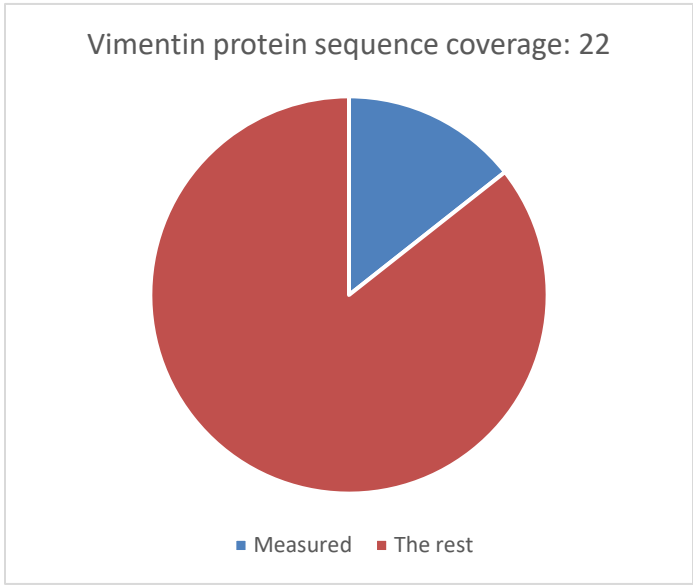

Fig. 23: Vimentin protein sequence coverage

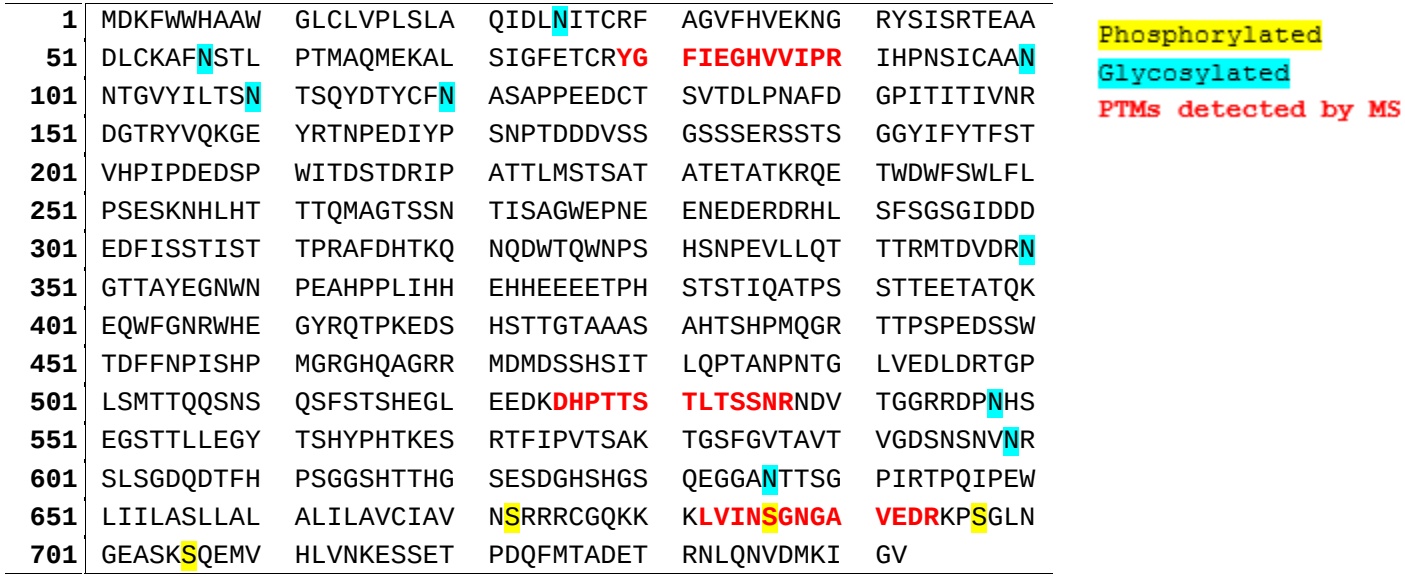

Fig. 24: CD44 peptide sequence with theoretical PTM sites and detected PTM sites

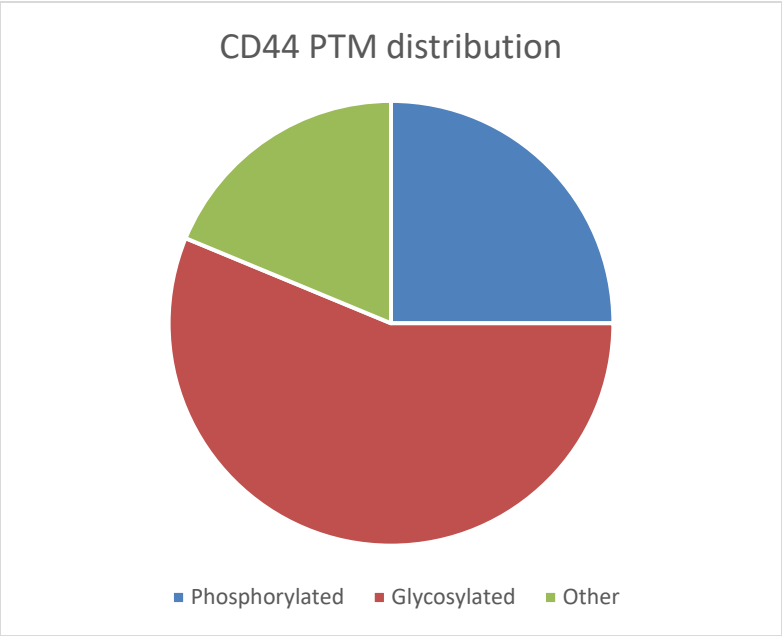

Fig. 25: CD44 PTM distribution

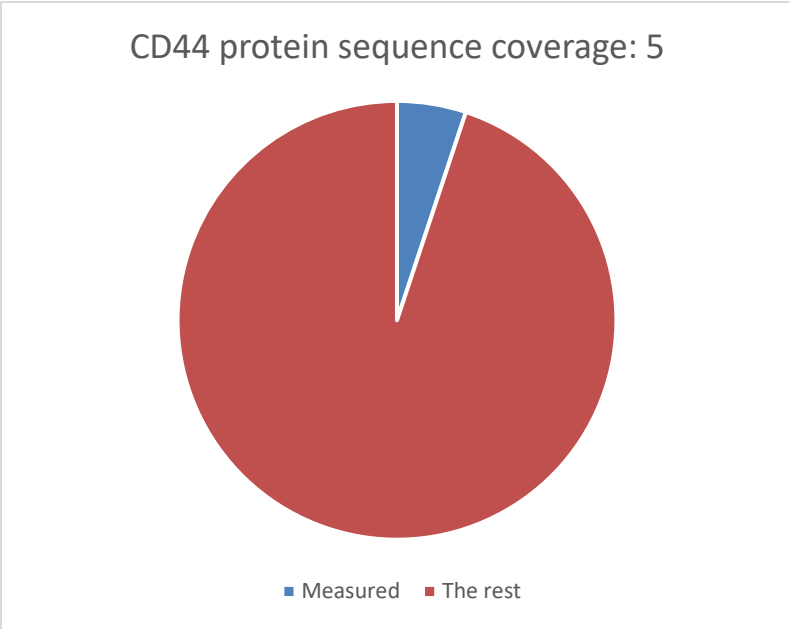

Fig. 26: CD44 protein sequence coverage

The graphs above serve as indicators of the PTM distribution in Vimentin (*Fig. 22*) and CD44 (*Fig. 25*). It si visible that the predominant PTM found in Vimentin is phosphorylation, accounting for 57 % of all PTMs known from peptide databses. As for CD44, glycosylation is more frequent and accounts for 56 % of all PTMs in CD44.

In comparison with the overall theoretical peptide profile of Vimentin, which has a total of 466 amino acids according to the UniprotKB database, 67 spots were succesfully detected, so the Vimentin protein sequence coverage was 22 % (*Fig. 27)*. For CD44, 38 spots out of 747 were detected, thus the protein sequence coverage for CD44 was 5 % (*Fig. 28).*

VIME_HUMAN									
Peptide origin	**Peptide (iTRAQ)**	**Sequence**	**Modification**	**Area ratio (114/117 m/z - Control/ATRA)**			**Downregulation**		
				whole lysates	**membrane**	**cytoplasm**	**whole lysates**	**membrane**	**cytoplasm**
1969,15	2113,15	ETNLDSLPLVDTHSKR	PHOS: 426, 430,436,438	3,88	ND	ND	74%	ND	ND
1093,51	1237,51	FADLSEAANR	PHOS: 299	3,89	1,49	1,23	75%	33%	20%
1428,83	1572,83	SLYASSPGGVYATR	PHOS: 51, 53,55,56,61	4,01	1,45	1,24	75%	32%	20%
1570,88	1714,98	ISLPLPNFSSLNLR	PHOS: 412, 419,420	3,91	1,45	1,24	74%	32%	20%
1050,64	1194,64	FLEQKNK	ACET: 129,SUCC: 129	3,91	1,46	1,24	74%	32%	20%
1828,16	1972,16	ILLAELEQLKGQGK	PHOS: 51, 53,55,56,61	3,89	ND	ND	75%	ND	ND
1269,81	1413,81	FANYIDKVR	ACET: 120, PHOS: 117,SUCC: 120	3,91	1,45	1,24	74%	32%	20%
1776,25	1920,25	FADLSEAANRNNDALR	PHOS: 299	3,91	ND	ND	74%	ND	ND

CD44_HUMAN									
Peptide origin	**Peptide (iTRAQ)**	**Sequence**	**Modification**	**Area ratio (114/117 m/z - Control/ATRA)**			**Downregulation**		
				whole lysates	**membrane**	**cytoplasm**	**whole lysates**	**membrane**	**cytoplasm**
1343,68	1487,23	LVINSGNGAVEDR	PHOS: 686	1,82	1,55	ND	45%	35%	ND
1386,76	1530,76	YGFIEGHVVIPR		1,81	1,57	ND	44%	36%	ND
1416,58	1560,58	DHPTTSTLTSSNR		1,82	1,57	ND	45%	36%	ND

Table 5: Summarized overview of found peptides and sequences corresponding to Vimentin and CD44 proteins after tryptic digest and iTRAQ labelling (MS).

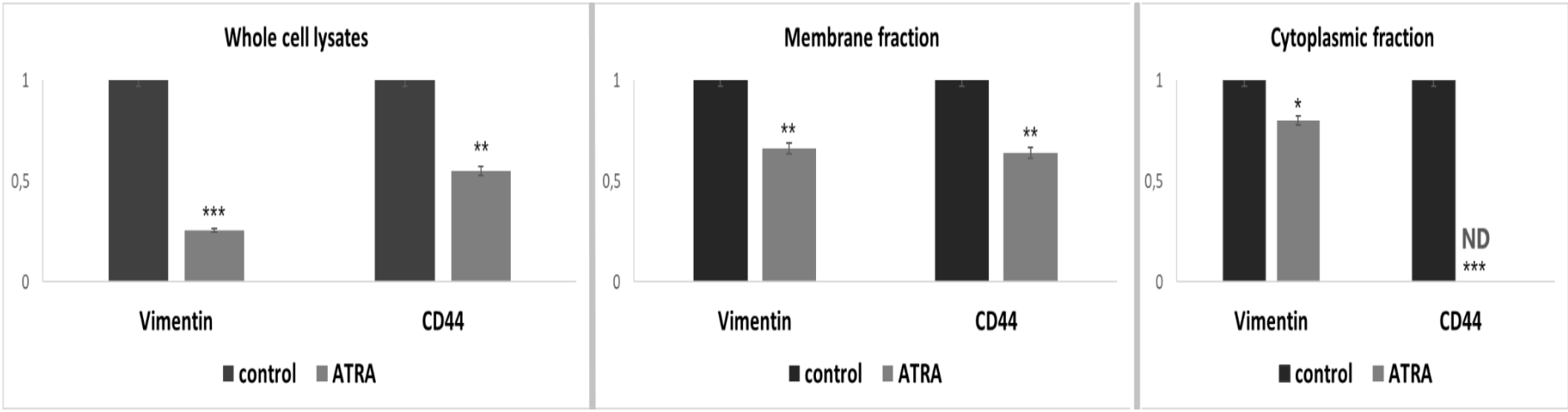

Fig. 29: Graphical expression of vimentin and CD44 proteins based on iTRAQ relative quantification assay. The ratio of each individual protein is expressed as an average value obtained from all identified labelled peptides corresponding to the given protein. The value of the control sample is set to 1, indicating a starting value. [126]

The graphs above show the quantitative analysis of vimentin and CD44 from the previous study (the obtained results were published in *General physiology and biophysics*[127]), taking into consideration their distribution in the cytoplasm and membrane fraction of the cells. The possible differences in distribution, especially in CD44, have been mentioned in a previous chapter(2.4.3). With the aim of trying to understand why this is happening, a new approach for monitoring ATRA treatment in TNBC cells had been used the previous work. This new approach included cell fractionation allowing for the analysis of not only whole cell lysates, but also cytoplasmic and membrane regions. The study has proven that the expression of both proteins after ATRA treatment vary greatly between the fractions, especially in the case of CD44, whose expression is almost nonexistant in the cytoplasmic region, whereas in the membrane fraction, ATRA treatment was only partially effective.

iTRAQ analysis in the case of Vimentin proved that ATRA treatment seems to have had more impact on the proteins detected in the membrane fraction, rather than the cytoplasm fraction. Nontheless, according to iTRAQ analysis, ATRA treatment of the whole cell lysates for Vimentin was effective.

As for CD44, ATRA treatment seems to have rendered the protein undetectable in the cytoplasm region, which was a success in terms of proving the capability of ATRA treatment in cancer. CD44 was detected only in the membrane fraction after treatment. Thus, ATRA treatment seems to be most effective in the cytoplasm region for this protein.

6 CONCLUSION

Some remarkable successes have been achieved thanks to the development of effective treatment steps for cancer. Acute promyelocytic leukemia, once considered the most malignant form, is currently one of the most curable forms of acute myeloid leukemia[128], all thanks to the endeavour of several generations of biomedical scientists and oncologists. It becomes evident that by opting for targeted treatments, which requires the indentification of compounds critical to the pathogenesis of the diseases, cure of cancer becomes an attainable option.

The aim of this work and its predecessor was to help in contributing to the mapping of the proteome of cancer cell lines and to analyse the the changes in representation of proteins in these cells after ATRA treatment using proteomic methods. As seen in the previous work, significant changes in the expression of certain proteins were demonstrated after ATRA treatment, proving its importance in the area of possible treatment of options for breast cancer, particularly in TNBC. Two specific proteins, Vimentin and CD44, were nominated for further study, as they proved to play important roles in processes characteristic to cancer, namely the EMT process of the problematic of stem cells. Significant reduction of both Vimentin and CD44 was observed in the previous study (see *Fig. 29*), which was confirmed in the work as well.

Both Vimentin and CD44 express a large number of variants based of their modifications (glycosylation, phosphorylation). The hypothesis was that some of these modifications may be specifically expressed in tumors. Proteoglycans such as Hyaluronic acid, when interacting with CD44[129], have been linked to poor prognosis and survival in cancer patients.[130] This work attempts to further understand how exactly these proteins and modifications are linked to cancer progression, which could be essential for the development of systems-biology-based synergistic targeting therapy and may in turn greatly improve the clinical outcome. The results indicate that modifications of both Vimentin and CD44 may be observed in the cancerous cells. Vimentin was predominantly phosphorylated, while CD44 was found to have more glycosylated sites. The modifications of the proteins show that the proteins are taking part in some some kinds of processes, most likely linked to the progression of the cancer, thus their observation may be used to prove the cancerous nature of the cells, ultimately marking the proteins and their PTMs as biomarkers.

7 LIST OF USED ABBREVIATIONS

TNBC – Triple negative breast cancer
MS – Mass spectrometry
EMT – Epithelial-mesenchymal transition
MALDI – Matrix-assisted laser desorption/ionization
TOF – Time of flight
ATRA – All-trans retinoic acid
ER – Estrogen receptor
PR – Progesterone receptor
HPLC – High pressure liquid chromatography
SDS – Sodium dodecyl sulfate
IEF – Isoelectric focusing
iTRAQ – Isobaric tags for relative and absolute quatitation
MS/MS – Tandem mass spectrometry
CSC – Cancer stem cells
PTM – Post-translational modification
IMAC - Immobilized metal affinity chromatography

8 LIST OF USED SOURCES

[1] PREDNÁ, Nikola. P*roteomic approach for the study of cancer cell line profiles.* Brno, 2020. Available at: https://www.vutbr.cz/studenti/zav-prace/detail/124246. Bachelor's thesis. Brno University of Technology, Faculty of Chemistry, Institute of Food Science and Biotechnology. Head of thesis Ing. Dana Strouhalová, Ph.D.

[2] Danny Hillis: Understanding cancer through proteomics. In: Youtube [online]. 16. 3. 2011 [8. 7. 2020]. https://www.youtube.com/watch?v=scITIoJvMJ0. Channel owner: TED.

[3] KLENER, Pavel, c2002. *Klinická onkologie*. Praha: Galén. ISBN 80-726-2151-3.

[4] CHOVANOVÁ, Silvia, 2011. *V Y U Ç I T ĉ R E T I N O I D š P ř I D I F E R E.N C I A Ļ N ĉ* Brno. Master's thesis. Masaryk University.

[5] HORÁK, Jaromír, Ctibor POVÝŠIL a Jitka ABRAHÁMOVÁ, 2000. *A t l a s n § P d a r G a d p r s u* ISBN 80-716-9771-0.

[6] YONETEN, K. KARAOSMANOGLU, M. KASAP, G. AKPINAR, A. GUNES, B. GUREL and N. Z. UTKAN, 2019. Comparative Proteome Analysis of Breast Cancer Tissues Highlights the Importance of Glycerol-3-phosphate Dehydrogenase 1 and Monoacylglycerol Lipase in Breast Cancer Metabolism. Cancer Genomics - *Proteomics* [online]. 16(5), 377-397 [cit. 2022-05-10]. ISSN 1109-6535. Available at: doi:10.21873/cgp.20143

[7] VELAZQUEZ BERUMEN, Adriana, Gabriela JIMENEZ MOYAO, Natalia M RODRIGUEZ, André M ILBAWI, Antonio MIGLIORE a Lawrence N SHULMAN, 2018. Defining priority medical devices for cancer management: a WHO initiative. *The Lancet Oncology* [online]. 19(12), e709-e719 [cit. 2022-05-10]. ISSN 14702045. Available at: doi:10.1016/S1470-2045(18)30658-2

[8] PEROU, Charles M., Therese SØRLIE, Michael B. EISEN, et al., 2000. Molecular portraits of human breast tumours. *Nature*. **406**(6797), 747-752. DOI: 10.1038/35021093. ISSN 0028-0836. Available also on: http://www.nature.com/articles/35021093

[9] SCHOLZEN, Thomas & Gerdes, JOHANNES. (2000). The Ki-67 protein: From the known and the unknown. *Journal of Cellular Physiology - J CELL PHYSIOL*. 182. 311-322. DOI: 10.1002/(SICI)1097-4652(200003)182:33.0.CO;2-9.

[10] PEROU, C. M. a A.-L. BORRESEN-DALE, 2011. Systems Biology and Genomics of Breast Cancer. *Cold Spring Harbor Perspectives in Biology*. **3**(2), a003293-a003293. DOI: 10.1101/cshperspect.a003293. ISSN 1943-0264. Available on: http://cshperspectives.cshlp.org/lookup/doi/10.1101/cshperspect.a003293

[11] CAREY, Lisa A., Charles M. PEROU, Chad A. LIVASY, et al., 2006. Race, Breast Cancer Subtypes, and Survival in the Carolina Breast Cancer Study. *JAMA*. **295**(21), a003293-a003293. DOI: 10.1001/jama.295.21.2492. ISSN 0098-7484. Available on: http://jama.jamanetwork.com/article.aspx?doi=10.1001/jama.295.21.2492

[12] STREHL, Johanna D., David L. WACHTER, Peter A. FASCHING, et al., 2011. Invasive Breast Cancer: Recognition of Molecular Subtypes. *Breast Care*. **6**(4), 258-264. DOI: 10.1159/000331339. ISSN 1661-3791. Available on: https://www.karger.com/Article/FullText/331339

[13] JHAN, Jing-Ru, Eran R ANDRECHEK, Peter A. FASCHING, et al., 2017. Triple-negative breast cancer and the potential for targeted therapy: Recognition of Molecular Subtypes. *Pharmacogenomics*. **18**(17), 1595-1609. DOI: 10.2217/pgs-2017-0117. ISSN 1462-2416. Available on: https://www.futuremedicine.com/doi/10.2217/pgs-2017-0117

[14] FOULKES, William D., Ian E. SMITH, Jorge S. REIS-FILHO, et al., 2010. Triple-Negative Breast Cancer: Recognition of Molecular Subtypes. *New England Journal of Medicine*. **363**(20), 1938-1948. DOI: 10.1056/NEJMra1001389. ISSN 0028-4793. Available on: http://www.nejm.org/doi/abs/10.1056/NEJMra1001389

[15] National Cancer Institute. Physician Data Query (PDQ). Breast Cancer Treatment – Health Professional Version. 2021. Accessed at https://www.cancer.gov/types/breast/hp/breast-treatment-pdq on August 24, 2021.

[16] DUFFY, M.J., N. HARBECK, M. NAP, R. MOLINA, A. NICOLINI, E. SENKUS a F. CARDOSO, 2017. Clinical use of biomarkers in breast cancer: Updated guidelines from the European Group on Tumor Markers (EGTM). *European Journal of Cancer* [online]. 75(12), 284-298 [cit. 2022-05-10]. ISSN 09598049. Available at: doi:10.1016/j.ejca.2017.01.017

[17] MUTEBI, Miriam, Benjamin O. ANDERSON, Catherine DUGGAN, et al., 2020. Breast cancer treatment: A phased approach to implementation. *Cancer* [online]. 126(S10), 2365-2378 [cit. 2022-05-10]. ISSN 0008-543X. Available at: doi:10.1002/cncr.32910

[18] GINSBURG, Ophira, Cheng-Har YIP, Ari BROOKS, et al., 2020. Breast cancer treatment: A phased approach to implementation. *Cancer* [online]. 126(S10), 2365-2378 [cit. 2022-05-10]. ISSN 0008-543X. Available at: doi:10.1002/cncr.32887

[19] LACROIX, Marc, Guy LECLERCQ, Ari BROOKS, et al., 2004. Relevance of Breast Cancer Cell Lines as Models for Breast Tumours: An Update. *Breast Cancer Research and Treatment* [online]. 83(3), 249-289 [cit. 2022-05-10]. ISSN 0167-6806. Available at: doi:10.1023/B:BREA.0000014042.54925.cc

[20] CAILLEAU, R., R. YOUNG, M. OLIVÉ, et al., 1974. Breast Tumor Cell Lines From Pleural Effusions 2: Recognition of Molecular Subtypes. *JNCI: Journal of the National Cancer Institute*. 53(3), 661-674. DOI: 10.1093/jnci/53.3.661. ISSN 1460-2105. Available on: https://academic.oup.com/jnci/article-lookup/doi/10.1093/jnci/53.3.661

[21] LEE, A. V., S. OESTERREICH, N. E. DAVIDSON, et al., 2015. MCF-7 Cells--Changing the Course of Breast Cancer Research and Care for 45 Years: An Update. *JNCI Journal of the National Cancer Institute* [online]. 107(7), djv073-djv073 [cit. 2022-05-10]. ISSN 0027-8874. Available at: doi:10.1093/jnci/djv073

[22] LEVENSON, Anait S, Kristen M SVOBODA, Hau C KWAAN, et al., 1998. Agonist activity of antiestrogen-receptor complexes to regulate urokinase plasminogen activator (uPA) and plasminogen activator inhibitor type 1 (PAI-1) endogenous gene expression in breast cancer cells: An Update. *Cancer Letters* [online]. 125(1-2), 215-220 [cit. 2022-05-10]. ISSN 03043835. Available at: doi:10.1016/S0304-3835(97)00516-8

[23] WASINGER, Valerie C., Stuart J. CORDWELL, Anne CERPA-POLJAK, et al., 1995. Progress with gene-product mapping of the Mollicutes: Mycoplasma genitalium. *Electrophoresis*. **16**(1), 1090-1094. DOI: 10.1002/elps.11501601185. ISSN 0173-0835. Available on: http://doi.wiley.com/10.1002/elps.11501601185

[24] LIEBLER, Daniel C., 2002. *Introduction to proteomics: tools for the new biology*. Totowa: Humana Press. ISBN 08-960-3992-7.

[25] MÓTYÁN, János, Ferenc TÓTH a József TŐZSÉR, 2013. Research Applications of Proteolytic Enzymes in Molecular Biology. *Biomolecules*. **3**(4), 923-942. DOI: 10.3390/biom3040923. ISSN 2218-273X. Available on: http://www.mdpi.com/2218-273X/3/4/923

[26] VEJRAŢKA, M. Buněčné kultury. [online]. Available at: http://bioprojekty.lf1.cuni.cz/3381/sylaby-prednasek/textova-verzeprednasek/bunecne-kultury-vejrazka.pdf

[27] SMITHIES, O., Stuart J. CORDWELL, Anne CERPA-POLJAK, et al., 1955. Zone electrophoresis in starch gels: group variations in the serum proteins of normal human adults. *Biochemical Journal*. **61**(4), 629-641. DOI: 10.1042/bj0610629. ISSN 0306-3283. Available on: https://portlandpress.com/biochemj/article/61/4/629/52105/Zone-electrophoresis-in-starch-gels-group

[28] KLOUDA, Pavel, 2003. *Moderní analytické metody*. 2., upr. a dopl. vyd. Ostrava: Pavel Klouda. ISBN 80-863-6907-2.

[29] CELIS, J., BRAVO R. *Two-dimensional gel electrophoresis of proteins: methods and applications*. New York: Academic Press, 1984, xvi, 487 p. ISBN 01-216-4720-X.

[30] REYNOLDS, J. A. a C. TANFORD, 1970. Binding of Dodecyl Sulfate to Proteins at High Binding Ratios. Possible Implications for the State of Proteins in Biological Membranes. *Proceedings of the National Academy of Sciences*. **66**(3), 1002-1007. DOI: 10.1073/pnas.66.3.1002. ISSN 0027-8424. Available on: http://www.pnas.org/cgi/doi/10.1073/pnas.66.3.1002

[31] DE ST. GROTH, S.Fazekas, R.G. WEBSTER a A. DATYNER, 1963. Two new staining procedures for quantitative estimation of proteins on electrophoretic strips. *Biochimica et Biophysica Acta*. **71**(3), 377-391. DOI: 10.1016/0006-3002(63)91092-8. ISSN 00063002. Available on: https://linkinghub.elsevier.com/retrieve/pii/0006300263910928

[32] WILSON, Curtis M., R.G. WEBSTER a A. DATYNER, 1979. Studies and critique of amido black 10B, Coomassie blue R, and fast green FCF as stains for proteins after polyacrylamide gel electrophoresis. *Analytical Biochemistry*. **96**(2), 263-278. DOI: 10.1016/0003-2697(79)90581-5. ISSN 00032697. Available on: https://linkinghub.elsevier.com/retrieve/pii/0003269779905815

[33] WAGNER, K., K. RACAITYTE, K.K. UNGER, et al., 2000. Protein mapping by two-dimensional high performance liquid chromatography: the move to mixtures. *Journal of Chromatography A*. **893**(2), 293-305. DOI: 10.1016/S0021-9673(00)00736-6. ISSN 00219673. Available on: https://linkinghub.elsevier.com/retrieve/pii/S0021967300007366

[34] PENG, Junmin, Steven P. GYGI a J. P. SALZMANN, 2001. Proteomics: the move to mixtures. *Journal of Mass Spectrometry*. **36**(10), 1083-1091. DOI: 10.1002/jms.229. ISSN 1076-5174. Available on: http://doi.wiley.com/10.1002/jms.229

[35] JANČÁŘOVÁ, I., JANČÁŘ L. *Analytická chemie*. Brno: Mendelova zemědělská a lesnická univerzita, 2003, 195 s. ISBN 978-80-7157-647-12008

[36] CHERVET, J. P., M. URSEM a J. P. SALZMANN, 1996. Instrumental Requirements for Nanoscale Liquid Chromatography. *Analytical Chemistry*. **68**(9), 1507-1512. DOI: 10.1021/ac9508964. ISSN 0003-2700. Available on: https://pubs.acs.org/doi/10.1021/ac9508964

[37] LEVINE ROULHAC, Petra, Carol HANEY BALL, K.K. UNGER, et al., 2009. Multidimensional Techniques in Protein Separations for Neuroproteomics: the move to mixtures. *Neuroproteomics*. CRC Press, 2009-10-26, **893**(2), 25-49. Frontiers in Neuroscience. DOI: 10.1201/9781420076264.ch3. ISBN 978-1-4200-7625-7. ISSN 00219673. Available on: http://www.crcnetbase.com/doi/abs/10.1201/9781420076264.ch3

[38] *Amino-acids, peptides, and proteins: Volume 1*, 1969. Cambridge: Royal Society of Chemistry. ISBN 18-475-5725-2.

[39] GARCIA N DUA, F. Využití proteomiky pro monitorování potenciálních biomarkerů rakoviny prsu. Brno: Vysoké učení technické v Brně, Fakulta chemická, 2014. 48 s. Vedoucí bakalářské práce Ing. Dana Flodrová, Ph.D.

[40] GYGI, Steven P., Beate RIST, Scott A. GERBER, et al., 1999. Quantitative analysis of complex protein mixtures using isotope-coded affinity tags: the move to mixtures. *Nature Biotechnology*. CRC Press, 2009-10-26, **17**(10), 994-999. Frontiers in Neuroscience. DOI: 10.1038/13690. ISBN 978-1-4200-7625-7. ISSN 1087-0156. Available on: http://www.nature.com/articles/nbt1099_994

[41] NORENA-CARO, Daniel, 2017. In: *Wikipedia* [online]. [cit. 2020-07-10]. Available at: https://upload.wikimedia.org/wikipedia/en/7/71/Liquid_chromatography_MS_spectrum_3D_analysis.png

[42] KARAS, Michael, Matthias GLUCKMANN a Jurgen SCHAFER, 2000. Ionization in matrix-assisted laser desorption/ionization: singly charged molecular ions are the lucky survivors. *Journal of Mass Spectrometry*.

[43] HILLENKAMP, Franz, Michael KARAS, Ronald C. BEAVIS a Brian T. CHAIT, 1991. Matrix-assisted laser desorption/ionization mass spectrometry of biopolymers. *Analytical Chemistry*. **63**(24), 1193A-1203A. DOI: 10.1021/ac00024a002. ISSN 0003-2700. Available on: https://pubs.acs.org/doi/abs/10.1021/ac00024a002

[44] KARAS, Michael, Ralf KRÜGER, Ronald C. BEAVIS a Brian T. CHAIT, 2003. Ion Formation in MALDI: The Cluster Ionization Mechanism. *Chemical Reviews*. **103**(2), 427-440. DOI: 10.1021/cr010376a. ISSN 0009-2665. Available on: https://pubs.acs.org/doi/10.1021/cr010376a

[45] XIAN, Feng, Christopher L. HENDRICKSON, Alan G. MARSHALL a Brian T. CHAIT, 2012. High Resolution Mass Spectrometry: The Cluster Ionization Mechanism. *Analytical Chemistry*. **84**(2), 708-719. DOI: 10.1021/ac203191t. ISSN 0003-2700. Available on: https://pubs.acs.org/doi/10.1021/ac203191t

[46] CHAIT, Brian T., Christopher L. HENDRICKSON, Alan G. MARSHALL a Brian T. CHAIT, 2011. Mass Spectrometry in the Postgenomic Era: The Cluster Ionization Mechanism. *Annual Review of Biochemistry*. **80**(1), 239-246. DOI: 10.1146/annurev-biochem-110810-095744. ISSN 0066-4154. Available on: http://www.annualreviews.org/doi/10.1146/annurev-biochem-110810-095744

[47] TRAUGER, Sunia A., William WEBB, Gary SIUZDAK a Brian T. CHAIT, 2002. Peptide and protein analysis with mass spectrometry: The Cluster Ionization Mechanism. *Spectroscopy*. **16**(1), 15-28. DOI: 10.1155/2002/320152. ISSN 0712-4813. Available on: http://www.hindawi.com/journals/jspec/2002/320152/abs/

[48] KELLNER, Stefanie. (2015). Re: What are the differences between LC/MS and LC/MS/MS ?. Retrieved from: https://www.researchgate.net/post/What_are_the_differences_between_LC_MS_and_LC_MS_MS/54ca8bd7d4c118941b8b467b/citation/download.

[49] FENSELAU, Catherine, Yulin N. HUANG, Jason N. MARCHESE, et al., 2007. A review of quantitative methods for proteomic studies: the move to mixtures. *Journal of Chromatography B*. **855**(1), 14-20. DOI: 10.1016/j.jchromb.2006.10.071. ISSN 15700232. Available on: https://linkinghub.elsevier.com/retrieve/pii/S157002320600883X

[50] ROSS, Philip L., Yulin N. HUANG, Jason N. MARCHESE, et al., 2004. Use of a lectin affinity selector in the search for unusual glycosylation in proteomics: the move to mixtures. *Journal of Chromatography B*. **3**(12), 405-418. DOI: 10.1074/mcp.M400129-MCP200. ISSN 1535-9476. Available on: https://linkinghub.elsevier.com/retrieve/pii/S1570023202006712

[51] LOVRIĆ, Josip, 2011. *Introducing proteomics: from concepts to sample separation, mass spectrometry and data analysis*. Chichester: Wiley-Blackwell. ISBN 978-0-470-03524-5.

[52] FLODROVA, Dana. *Gel associated quantification based on protein changes profiling breast cancer cell lines after retinoid treatments* [PP presentation]. Brno: Academy of Science, Department of analytical chemistry, 2017.

[53] LIEBLER, D. C. Proteomics in cancer research. Nashville, TN: Wiley-Liss, c2005, xii, 201 p. ISBN 04-714-4475-8.

[54] LOEB, Lawrence A. A mutator phenotype in cancer. *Cancer research*, 2001, 61.8: 3230-3239.

[55] SIEBER, Oliver M., Karl HEINIMANN a Ian P. M. TOMLINSON, 2003. Genomic instability — the engine of tumorigenesis? *Nature Reviews Cancer*. **3**(9), 701-708. DOI: 10.1038/nrc1170. ISSN 1474-175X. Available on: http://www.nature.com/articles/nrc1170

[56] FILIOU, Michaela D., Christoph W. TURCK a Daniel MARTINS-DE-SOUZA, 2011. Quantitative proteomics for investigating psychiatric disorders. *PROTEOMICS - Clinical Applications*. **5**(1-2), 38-49. DOI: 10.1002/prca.201000060. ISSN 18628346. Available on: http://doi.wiley.com/10.1002/prca.201000060

[57] KOOMEN, John M., Eric B. HAURA, Gerold BEPLER, et al., 2008. Proteomic Contributions to Personalized Cancer Care. *PROTEOMICS - Clinical Applications*. **7**(10), 1780-1794. DOI: 10.1074/mcp.R800002-MCP200. ISSN 1535-9476. Available on: http://www.mcponline.org/lookup/doi/10.1074/mcp.R800002-MCP200

[58] BRTKO, J., J. THALHAMER a József TŐZSÉR, 2003. Renaissance of the Biologically Active Vitamin A Derivatives: Established and Novel Directed Therapies for Cancer and Chemoprevention. *Current Pharmaceutical Design*. **9**(25), 2067-2077. DOI: 10.2174/1381612033454144. ISSN 13816128. Available on: http://www.mdpi.com/2218-273X/3/4/923

[59] SUN, S.-Y., R. LOTAN a József TŐZSÉR, 1998. Retinoids as chemopreventive and therapeutic agents: Established and Novel Directed Therapies for Cancer and Chemoprevention. *Drugs of the Future*. **23**(6), 2067-2077. DOI: 10.1358/dof.1998.023.06.858360. ISSN 0377-8282. Available on: http://www.mdpi.com/2218-273X/3/4/923

[60] BIENOVÁ, M., KUČEROVÁ R. Retinoidy v dermatologii. Dermatologie pro praxi. 2008; 2(4): 171-174. Available at: http://www.dermatologiepropraxi.cz/pdfs/der/2008/04/03.pdf

[61] https://en.wikipedia.org/wiki/Retinoid. [online] Link to image: https://upload.wikimedia.org/wikipedia/commons/thumb/c/c6/Retinoids_skeletal_formulae.svg/400px-Retinoids_skeletal_formulae.svg.png

[62] VAHLQUIST, A., 1999. What Are Natural Retinoids? *Dermatology*. **199**(1), 3-11. DOI: 10.1159/000051371. ISSN 1018-8665. Available on: https://www.karger.com/Article/FullText/51371

[63] SUN, Shi-Yong a Reuben LOTAN, 2002. Retinoids and their receptors in cancer development and chemoprevention. *Critical Reviews in Oncology/Hematology* [online]. 41(1), 41-55 [cit. 2022-05-11]. ISSN 10408428. Available at: doi:10.1016/S1040-8428(01)00144-5

[64] HINDS, Pamela S., Ralph J. VOGEL a Laura CLARKE-STEFFEN, 1997. The Possibilities and Pitfalls of Doing a Secondary Analysis of a Qualitative Data Set. *Qualitative Health Research* [online]. 7(3), 408-424 [cit. 2022-05-11]. ISSN 1049-7323. Available at: doi:10.1177/104973239700700306

[65] BIENOVÁ, M., KUČEROVÁ R. Retinoidy v dermatologii. *Dermatologie pro praxi*. 2008; 2(4): 171-174. Available at: http://www.dermatologiepropraxi.cz/pdfs/der/2008/04/03.pdf

[66] KLENER, P. *Základy klinické onkologie*. 1. vyd. Praha: Galén, c2011, 96 s. ISBN 978807-2627-165.

[67] SIMMS, W. A ROSS P. S. Vitamin A physiology and its application as a biomarker of contaminant-related toxicity in marine mammals: a review. *Toxicology and Industrial Health*. 2000-08-01, vol. 16, 7-8, s. 291-302. DOI: 10.1177/074823370001600706.

[68] ADÈS, Lionel, Agnes GUERCI, Emmanuel RAFFOUX, et al., 2010. Very long-term outcome of acute promyelocytic leukemia after treatment with all-trans retinoic acid and chemotherapy: the European APL Group experience. *Blood*. **115**(9), 1690-1696. DOI: 10.1182/blood-2009-07-233387. ISSN 0006-4971. Available on: https://ashpublications.org/blood/article/115/9/1690/27215/Very-longterm-outcome-of-acute-promyelocytic

[69] LIPPMAN, S M, S E BENNER a W K HONG, 1994. Retinoid Chemoprevention Studies in Upper Aerodigestive Tract and Lung Carcinogenesis. *Cancer Res*. (54), 2025-2028.

[70] KRUPITZA, Georg, Wolfgang HULLA, Hanna HARANT, et al., 1995. Retinoic acid induced death of ovarian carcinoma cells correlates with c-myc stimulation. *International Journal of Cancer*. **61**(5), 649-657. DOI: 10.1002/ijc.2910610511. ISSN 00207136. Available on: http://doi.wiley.com/10.1002/ijc.2910610511

[71] WU, Shujian, Anne DONIGAN, Chris D. PLATSOUCAS, et al., 1997. All-trans-Retinoic Acid Blocks Cell Cycle Progression of Human Ovarian Adenocarcinoma Cells at Late G1. *Experimental Cell Research*. **232**(2), 277-286. DOI: 10.1006/excr.1997.3495. ISSN 00144827. Available on: https://linkinghub.elsevier.com/retrieve/pii/S0014482797934959

[72] LENKO, Vladimir, Lucia BIALESOVA, Dana MACEJOVA, et al., 2013. The relationship between renal cell carcinoma and nuclear retinoid/rexinoid receptors. *Biomedical Papers*. **157**(4), 316-324. DOI: 10.5507/bp.2013.060. ISSN 12138118. Available on: http://biomed.papers.upol.cz/doi/10.5507/bp.2013.060.html

[73] HONG, Waun Ki a Reuben LOTAN, 1993. *Retinoids in Oncology*. CRC Press. ISBN 0824790480.

[74] BIALEŠOVÁ, Lucia, Július BRTKO, Vladimír LENKO a Dana MACEJOVÁ, 2013. Nuclear receptors – target molecules for isoflavones in cancer chemoprevention. *General physiology and biophysics* [online]. 32(04), 467-478 [cit. 2022-05-11]. ISSN 1338-4325. Available at: doi:10.4149/gpb_2013064

[75] GARATTINI, Enrico, Marco BOLIS, Silvio Ken GARATTINI, et al., 2014. Retinoids and breast cancer: From basic studies to the clinic and back again. *Cancer Treatment Reviews* [online]. 40(6), 739-749 [cit. 2022-05-11]. ISSN 03057372. Available at: doi:10.1016/j.ctrv.2014.01.001

[76] LEDZEWITZ, Urszula a Heinz SCHATTLER, 2012. *Mathematical Methods and Models in Biomedicine*. Springer Science & Business Media. ISBN 1461441773.

[77] GINESTIER, Christophe, Julien WICINSKI, Nathalie CERVERA, et al., 2014. Retinoid signaling regulates breast cancer stem cell differentiation. *Cell Cycle*. 8(20), 3297-3302. DOI: 10.4161/cc.8.20.9761. ISSN 1538-4101. Available also on: http://www.tandfonline.com/doi/abs/10.4161/cc.8.20.9761

[78] FLODROVA, D., D. BENKOVSKA, D. MACEJOVA, L. BIALESOVA, J. BOBALOVA a J. BRTKO, 2013. Effects of retinoic acid isomers on proteomic pattern in human breast cancer MCF-7 cell line. *Endocrine Regulations*. **47**(04), 205-209. DOI: 10.4149/endo_2013_04_205. ISSN 1336-0329.

[79] FLODROVA, D., D. BENKOVSKA, D. MACEJOVA, L. BIALESOVA, J. BOBALOVA a J. BRTKO, 2013. Effects of retinoic acid isomers on proteomic pattern in human breast cancer MCF-7 cell line. *Endocrine Regulations*. **47**(04), 205-209. DOI: 10.4149/endo_2013_04_205. ISSN 1336-0329.

[80] *Journal of Clinical Oncology*, 2007. 25. ISSN 0732-183X. Available on: http://ascopubs.org/doi/10.1200/JCO.2007.14.2364

[81] MARYÁŠ, Josef, Jakub FAKTOR, Monika DVOŘÁKOVÁ, Iva STRUHÁROVÁ, Peter GRELL a Pavel BOUCHAL, 2014. Proteomics in investigation of cancer metastasis: Functional and clinical consequences and methodological challenges. *PROTEOMICS*. **14**(4-5), 426-440. DOI: 10.1002/pmic.201300264. ISSN 16159853. Available on: http://doi.wiley.com/10.1002/pmic.201300264

[82] SOWA, Terumasa, Toshi MENJU, Makoto SONOBE, et al., 2015. Association between epithelial-mesenchymal transition and cancer stemness and their effect on the prognosis of lung adenocarcinoma: the role of cadherins in metastasis. *Cancer Medicine*. **4**(12), 1853-1862. DOI: 10.1002/cam4.556. ISSN 20457634. Available on: http://doi.wiley.com/10.1002/cam4.556

[83] ZHAO, Shujie, Chen CHEN, Katherine CHANG, Anand KARNAD, Jaishree JAGIRDAR, Addanki P. KUMAR a James W. FREEMAN, 2016. CD44 Expression Level and Isoform Contributes to Pancreatic Cancer Cell Plasticity, Invasiveness, and Response to Therapy: cancer stem cells in metastasis. *Clinical Cancer Research*. **22**(22), 5592-5604. DOI: 10.1158/1078-0432.CCR-15-3115. ISSN 1078-0432. Available on: http://clincancerres.aacrjournals.org/lookup/doi/10.1158/1078-0432.CCR-15-3115

[84] PREDNÁ, Nikola. *Proteomic approach for the study of cancer cell line profiles*. Brno, 2020. Available at: https://www.vutbr.cz/studenti/zav-prace/detail/124246. Bachelor's thesis. Brno University of Technology, Faculty of Chemistry, Institute of Food Science and Biotechnology. Head of thesis Ing. Dana Strouhalová, Ph.D.

[85] NEAGU, Monica, Carolina CONSTANTIN, Marinela BOSTAN, Constantin CARUNTU, Simona Rebeca IGNAT, Sorina DINESCU a Marieta COSTACHE, 2019. Proteomic Technology "Lens" for Epithelial-Mesenchymal Transition Process Identification in Oncology. *Analytical Cellular Pathology* [online]. 2019, 1-17 [cit. 2022-05-12]. ISSN 2210-7177. Available at: doi:10.1155/2019/3565970

[86] BEAVON, I.R.G, 2000. The E-cadherin–catenin complex in tumour metastasis. *European Journal of Cance*r. **36**(13), 1607-1620. DOI: 10.1016/S0959-8049(00)00158-1. ISSN 09598049. Available on: https://linkinghub.elsevier.com/retrieve/pii/S0959804900001581

[87] KALLURI, Raghu, Robert A. WEINBERG, Alexander R A ANDERSON a Dirk DRASDO, 2009. The basics of epithelial-mesenchymal transition: the role of cadherins in metastasis. *Journal of Clinical Investigation*. **119**(6), 1420-1428. DOI: 10.1172/JCI39104. ISSN 0021-9738. Available on: http://www.jci.org/articles/view/39104

[88] IVASKA, Johanna, Robert A. WEINBERG, Alexander R A ANDERSON a Dirk DRASDO, 2014. Vimentin: the role of cadherins in metastasis. *Small GTPases*. **2**(1), 51-53. DOI: 10.4161/sgtp.2.1.15114. ISSN 2154-1248. Available on: http://www.tandfonline.com/doi/abs/10.4161/sgtp.2.1.15114

[89] VUORILUOTO, K, H HAUGEN, S KIVILUOTO, et al., 2011. Vimentin regulates EMT induction by Slug and oncogenic H-Ras and migration by governing Axl expression in breast cancer: the role of cadherins in metastasis. *Oncogene*. **30**(12), 1436-1448. DOI: 10.1038/onc.2010.509. ISSN 0950-9232. Available on: http://www.nature.com/articles/onc2010509

[90] SATELLI, Arun, Shulin LI, S KIVILUOTO, et al., 2011. Vimentin in cancer and its potential as a molecular target for cancer therapy: the role of cadherins in metastasis. *Cellular and Molecular Life Sciences*. **68**(18), 3033-3046. DOI: 10.1007/s00018-011-0735-1. ISSN 1420-682X. Available on: http://link.springer.com/10.1007/s00018-011-0735-1

[91] DERYCKE, Lara D.M., Marc E. BRACKE, Alexander R A ANDERSON a Dirk DRASDO, 2004. N-cadherin in the spotlight of cell-cell adhesion, differentiation, embryogenesis, invasion and signalling: the role of cadherins in metastasis. *The International Journal of Developmental Biology*. **48**(5-6), 463-476. DOI: 10.1387/ijdb.041793ld. ISSN 0214-6282. Available on: http://www.intjdevbiol.com/paper.php?doi=041793ld

[92] RAMIS-CONDE, Ignacio, Mark A J CHAPLAIN, Alexander R A ANDERSON a Dirk DRASDO, 2009. Multi-scale modelling of cancer cell intravasation: the role of cadherins in metastasis. *Physical Biology*. **6**(1). DOI: 10.1088/1478-3975/6/1/016008. ISSN 1478-3975. Available on: https://iopscience.iop.org/article/10.1088/1478-3975/6/1/016008

[93] GREENBURG, G, E D HAY, S KIVILUOTO, et al., 1982. Epithelia suspended in collagen gels can lose polarity and express characteristics of migrating mesenchymal cells: the role of cadherins in metastasis. *The Journal of Cell Biology*. **95**(1), 333-339. DOI: 10.1083/jcb.95.1.333. ISSN 0021-9525. Available on: https://rupress.org/jcb/article/95/1/333/19781/Epithelia-suspended-in-collagen-gels-can-lose

[94] THIERY, Jean Paul, Robert A. WEINBERG, S KIVILUOTO, et al., 2002. Epithelial–mesenchymal transitions in tumour progression: the role of cadherins in metastasis. *Nature Reviews Cancer*. **2**(6), 442-454. DOI: 10.1038/nrc822. ISSN 1474-175X. Available on: http://www.nature.com/articles/nrc822

[95] GREENBURG, G, E D HAY, S KIVILUOTO, et al., 1982. Epithelia suspended in collagen gels can lose polarity and express characteristics of migrating mesenchymal cells: the role of cadherins in metastasis. *The Journal of Cell Biology*. **95**(1), 333-339. DOI: 10.1083/jcb.95.1.333. ISSN 0021-9525. Available on: https://rupress.org/jcb/article/95/1/333/19781/Epithelia-suspended-in-collagen-gels-can-lose

[96] FRANKE, W. W., E. SCHMID, M. OSBORN, et al., 1978. Different intermediate-sized filaments distinguished by immunofluorescence microscopy: the role of cadherins in metastasis. *Proceedings of the National Academy of Sciences*. **75**(10), 5034-5038. DOI: 10.1073/pnas.75.10.5034. ISSN 0027-8424. Available on: http://www.pnas.org/cgi/doi/10.1073/pnas.75.10.5034

[97] ZHAO, Shujie, Chen CHEN, Katherine CHANG, Anand KARNAD, Jaishree JAGIRDAR, Addanki P. KUMAR a James W. FREEMAN, 2016. CD44 Expression Level and Isoform Contributes to Pancreatic Cancer Cell Plasticity, Invasiveness, and Response to Therapy: cancer stem cells in metastasis. *Clinical Cancer Research*. **22**(22), 5592-5604. DOI: 10.1158/1078-0432.CCR-15-3115. ISSN 1078-0432. Available on: http://clincancerres.aacrjournals.org/lookup/doi/10.1158/1078-0432.CCR-15-3115

[98]Epithelial-Mesenchymal Transition (EMT)and the Involvement of Pho Family Small G-proteins. *Cytoskeleton NEWS: The Protein Experts*. Epithelial-Mesenchymal Transition (EMT) News. EMT Publications, s. 2.

[99] LI, Feng, Benjamin TIEDE, Joan MASSAGUÉ a Yibin KANG, 2007. Beyond tumorigenesis: cancer stem cells in metastasis. *Cell Research*. **17**(1), 3-14. DOI: 10.1038/sj.cr.7310118. ISSN 1001-0602. Available on: http://www.nature.com/articles/7310118

[100] EIBL, Robert H., Torsten PIETSCH, Jürgen MOLL, et al., 1995. Expression of variant CD44 epitopes in human astrocytic brain tumors: cancer stem cells in metastasis. *Journal of Neuro-Oncology*. **26**(3), 165-170. DOI: 10.1007/BF01052619. ISSN 0167-594X. Available on: http://link.springer.com/10.1007/BF01052619

[101] NAOR, David, Shlomo NEDVETZKI, Itshak GOLAN, Lora MELNIK a Yoram FAITELSON, 2008. CD44 in Cancer: cancer stem cells in metastasis. *Critical Reviews in Clinical Laboratory Sciences*. **39**(6), 527-579. DOI: 10.1080/10408360290795574. ISSN 1040-8363. Available on: http://www.tandfonline.com/doi/full/10.1080/10408360290795574

[102] LE BRAS, Grégoire F., Gillian L. ALLISON, Nicole F. RICHARDS, Shazia S. ANSARI, M. Kay WASHINGTON, Claudia D. ANDL a Adam I. MARCUS, 2011. CD44 Upregulation in E-Cadherin-Negative Esophageal Cancers Results in Cell Invasion: cancer stem cells in metastasis. *PLoS ONE*. **6**(11), 1303-1313. DOI: 10.1371/journal.pone.0027063. ISSN 1932-6203. Available on: https://dx.plos.org/10.1371/journal.pone.0027063

[103] PREDNÁ, Nikola. *Proteomic approach for the study of cancer cell line profiles*. Brno, 2020. Available at: https://www.vutbr.cz/studenti/zav-prace/detail/124246. Bachelor's thesis. Brno University of Technology, Faculty of Chemistry, Institute of Food Science and Biotechnology. Head of thesis Ing. Dana Strouhalová, Ph.D.

[104] STROUHALOVA, D, D MACEJOVA, M LASTOVICKOVA, J BRTKO a J BOBALOVA, 2020. CD44 and vimentin, markers involved with epithelial-mesenchymal transition: A proteomic analysis of sequential proteins extraction of triple-negative breast cancer cells after treatment with all-trans retinoic acid. *General Physiology and Biophysics*.

[105] REYA, Tannishtha, Sean J. MORRISON, Michael F. CLARKE, et al., 2001. Stem cells, cancer, and cancer stem cells: From basic studies to the clinic and back again. Nature [online]. 414(6859), 105-111 [cit. 2022-05-11]. ISSN 0028-0836. Available at: doi:10.1038/35102167

[106] KAPLAN, Rosandra N., Rebecca D. RIBA, Stergios ZACHAROULIS, et al., 2005. Stem cells, cancer, and cancer stem cells: From basic studies to the clinic and back again. *Nature* [online]. 438(7069), 105-111 [cit. 2022-05-11]. ISSN 0028-0836. Available at: doi:10.1038/nature04186

[107] VARKI, A, RD CUMMINGS a et al., 1999. Chapter 47: Glycosylation Changes in Cancer. *Essentials of Glycobiology*. 3rd Edition. Cold Spring Harbor.

[108] KAILEMIA, Muchena J., Dayoung PARK, Carlito B. LEBRILLA, et al., 2017. Glycans and glycoproteins as specific biomarkers for cancer: the role of cadherins in metastasis. *Analytical and Bioanalytical Chemistry*. **409**(2), 395-410. DOI: 10.1007/s00216-016-9880-6. ISSN 1618-2642. Available on: http://link.springer.com/10.1007/s00216-016-9880-6

[109] GÖTTE, Martin, Ilona KOVALSZKY, Carlito B. LEBRILLA, et al., 2018. Extracellular matrix functions in lung cancer: the role of cadherins in metastasis. *Matrix Biology*. **73**(2), 105-121. DOI: 10.1016/j.matbio.2018.02.018. ISSN 0945053X. Available on: https://linkinghub.elsevier.com/retrieve/pii/S0945053X17303293

[110] VARKI, Ajit a Reiji KANNAGI, 2017. *Essentials of Glycobiology: Glycosylation Changes in Cancer*. 3. New York: Cold Spring Harbor.

[111] NIKITOVIC, Dragana, Aikaterini BERDIAKI, Ioanna SPYRIDAKI, Theodoros KRASANAKIS, Aristidis TSATSAKIS a George N. TZANAKAKIS, 2018. Proteoglycans—Biomarkers and Targets in Cancer Therapy. *Frontiers in Endocrinology*. **9**. DOI: 10.3389/fendo.2018.00069. ISSN 1664-2392. Available on: http://journal.frontiersin.org/article/10.3389/fendo.2018.00069/full

[112] CHEN, I-Hsuan, Liang XUE, Chuan-Chih HSU, et al., 2017. Phosphoproteins in extracellular vesicles as candidate markers for breast cancer. *Proceedings of the National Academy of Sciences*. **114**(12), 3175-3180. DOI: 10.1073/pnas.1618088114. ISSN 0027-8424. Available on: http://www.pnas.org/lookup/doi/10.1073/pnas.1618088114

[113] LARSEN, Martin R., Morten B. TRELLE, Tine E. THINGHOLM a Ole N. JENSEN, 2006. Analysis of posttranslational modifications of proteins by tandem mass spectrometry. *BioTechniques*. **40**(6), 790-798. DOI: 10.2144/000112201. ISSN 0736-6205. Available on: https://www.future-science.com/doi/10.2144/000112201

[114] VIRÁG, Dávid, Borbála DALMADI-KISS, Károly VÉKEY, László DRAHOS, Imre KLEBOVICH, István ANTAL a Krisztina LUDÁNYI, 2020. Current Trends in the Analysis of Post-translational Modifications. *Chromatographia*. **83**(1), 1-10. DOI: 10.1007/s10337-019-03796-9. ISSN 0009-5893. Available on: http://link.springer.com/10.1007/s10337-019-03796-9

[115] KAILEMIA, Muchena J., Dayoung PARK, Carlito B. LEBRILLA, László DRAHOS, Imre KLEBOVICH, István ANTAL a Krisztina LUDÁNYI, 2017. Glycans and glycoproteins as specific biomarkers for cancer. *Analytical and Bioanalytical Chemistry*. **409**(2), 395-410. DOI: 10.1007/s00216-016-9880-6. ISSN 1618-2642. Available on: http://link.springer.com/10.1007/s00216-016-9880-6

[116] *Characterization of Protein Therapeutics using Mass Spectrometry*, 2013. Springer US. Guodong Chen. ISBN 978-1-4419-7862-2.

[117] XIONG, Li, Fred E REGNIER a J. P. SALZMANN, 2002. Use of a lectin affinity selector in the search for unusual glycosylation in proteomics: the move to mixtures. *Journal of Chromatography B*. **782**(1-2), 405-418. DOI: 10.1016/S1570-0232(02)00671-2. ISSN 15700232. Available on: https://linkinghub.elsevier.com/retrieve/pii/S1570023202006712

[118] YOUSEF, George M a Serge JOTHY, 2014. *Molecular Testing in Cancer*. Springer Science & Business. ISBN 9781489980502.

[119] SANO, Akira a Hiroshi NAKAMURA, 2004. Chemo-affinity of Titania for the Column-switching HPLC Analysis of Phosphopeptides. *Analytical Sciences*. **20**(3), 565-566. DOI: 10.2116/analsci.20.565. ISSN 0910-6340. Available on: http://joi.jlc.jst.go.jp/JST.JSTAGE/analsci/20.565?from=CrossRef

[120] SANO, Akira a Hiroshi NAKAMURA, 2007. Evaluation of Titanium and Titanium Oxides as Chemo-affinity Sorbents for the Selective Enrichment of Organic Phosphates. *Analytical Sciences*. **23**(11), 1285-1289. DOI: 10.2116/analsci.23.1285. ISSN 0910-6340. Available on: http://joi.jlc.jst.go.jp/JST.JSTAGE/analsci/23.1285?from=CrossRef

[121] KOLSRUD, Hanne, Helle MALEROD, Steven RAY, Leon REUBSAET, Elsa LUNDANES a Tyge GREIBROKK, 2012. A Critical Review of Trypsin Digestion for LC-MS Based Proteomics. *Integrative Proteomics*. InTech, 2012-02-24. DOI: 10.5772/29326. ISBN 978-953-51-0070-6. Available on: http://www.intechopen.com/books/integrative-proteomics/a-critical-review-of-trypsin-digestion-for-lc-ms-based-proteomics

[122] *Protein Electrophoresis Methods* [online], 2022. Germany: Bio-Rad Laboratories [cit. 2022-05-11]. Available at: https://www.bio-rad.com/de-de/applications-technologies/protein-electrophoresis-methods?ID=LUSOW4GRI

[123] CARR, David. *A Guide to the Analysis and Purification of Proteins and Peptides by Reversed-Phase HPLC*. Advanced Chromatography Technologies, s. 64.

[124] DYČKA, Filip. *Identifikace a charakterizace proteinŢ pomocí hmotnostní spektrometrie* [online]. Brno, 2012 [cit. 2022-05-11]. Available from: https://is.muni.cz/th/kz5f3/. Doctoral theses, Dissertations. Masaryk University, Faculty of Science. Thesis supervisor Janette BOBÁĽOVÁ.

[cxxv] FLODROVA, Dana, Lucia TOPOROVA, Dana MACEJOVA, Marketa LASTOVICKOVA, Julius BRTKO a Janette BOBALOVA, 2016. A comparative study of protein patterns of human estrogen

receptor positive (MCF-7) and negative (MDA-MB-231) breast cancer cell lines. *General physiology and biophysics* [online]. 35(03), 387-392 [cit. 2022-05-11]. ISSN 1338-4325. Available at: doi:10.4149/gpb_2016009

[126] STROUHALOVA, Dana, Dana MACEJOVA, Marketa LASTOVICKOVA, Julius BRTKO, Janette BOBALOVA, Sorina DINESCU a Marieta COSTACHE, 2020. CD44 and vimentin, markers involved with epithelial-mesenchymal transition: A proteomic analysis of sequential proteins extraction of triple-negative breast cancer cells after treatment with all-trans retinoic acid. *General physiology and biophysics* [online]. 39(04), 399-405 [cit. 2022-05-12]. ISSN 1338-4325. Available at: doi:10.4149/gpb_2020026

[127] STROUHALOVA, Dana, Dana MACEJOVA, Marketa LASTOVICKOVA, Julius BRTKO, Janette BOBALOVA, Sorina DINESCU a Marieta COSTACHE, 2020. CD44 and vimentin, markers involved with epithelial-mesenchymal transition: A proteomic analysis of sequential proteins extraction of triple-negative breast cancer cells after treatment with all-trans retinoic acid. *General physiology and biophysics* [online]. 39(04), 399-405 [cit. 2022-05-12]. ISSN 1338-4325. Available at: doi:10.4149/gpb_2020026

[128] ZHOU, Guang-Biao, Ji ZHANG, Zhen-Yi WANG, et al., 2007. Treatment of acute promyelocytic leukaemia with all-trans retinoic acid and arsenic trioxide: a paradigm of synergistic molecular targeting therapy. *Philosophical Transactions of the Royal Society B: Biological Sciences*. **362**(1482), 959-971. DOI: 10.1098/rstb.2007.2026. ISSN 0962-8436. Available on: https://royalsocietypublishing.org/doi/10.1098/rstb.2007.2026

[129] ARUFFO, Alejandro, Ivan STAMENKOVIC, Michael MELNICK, Charles B. UNDERHILL a Brian SEED, 1990. CD44 is the principal cell surface receptor for hyaluronate: cancer stem cells in metastasis. *Cell*. **61**(7), 1303-1313. DOI: 10.1016/0092-8674(90)90694-A. ISSN 00928674. Available on: https://linkinghub.elsevier.com/retrieve/pii/009286749090694A

[130] KAROUSOU, Evgenia, Suniti MISRA, Shibnath GHATAK, et al., 2017. Roles and targeting of the HAS/hyaluronan/CD44 molecular system in cancer. *Matrix Biology*. **59**, 3-22. DOI: 10.1016/j.matbio.2016.10.001. ISSN 0945053X. Available on: https://linkinghub.elsevier.com/retrieve/pii/S0945053X16301706