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.

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.

2.2 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]

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]

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]

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.

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

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.

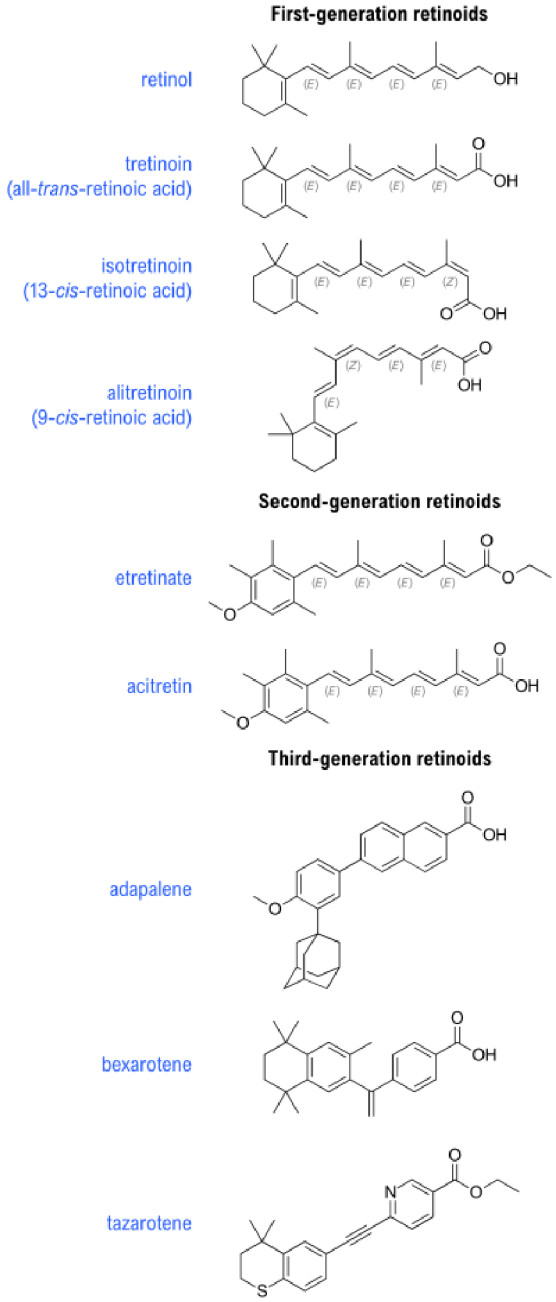

Fig. 6: Retinoid classes and their representants.[61]

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]

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]

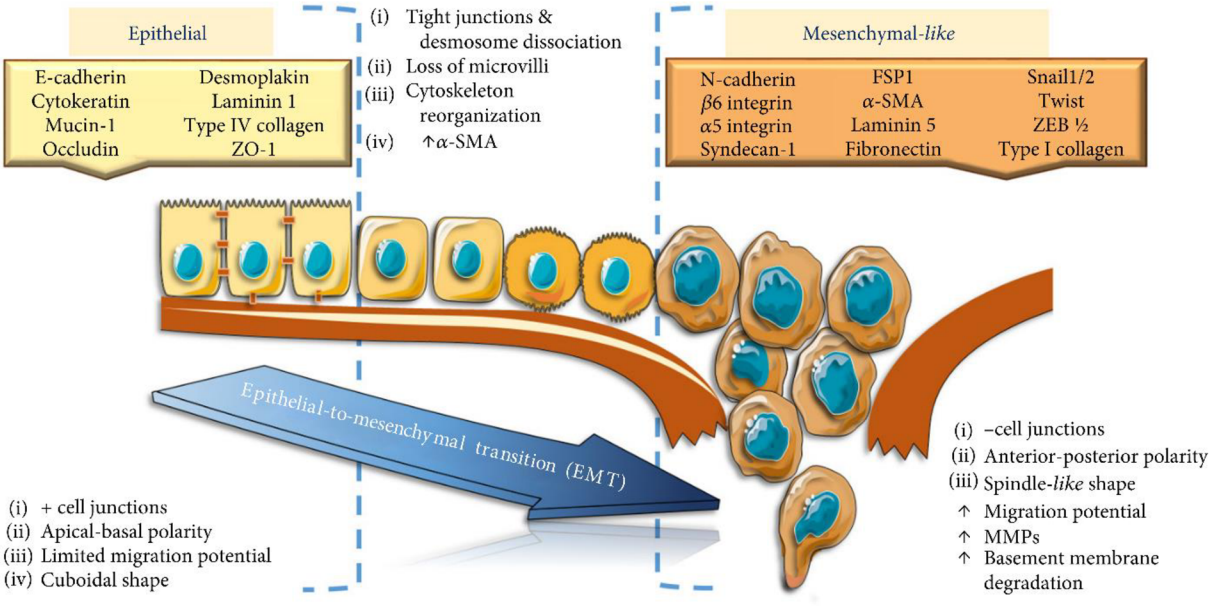

Fig. 8: The process of EMT with respective biomarkers for epithelial and mesenchymal cells[85]

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

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.

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

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.

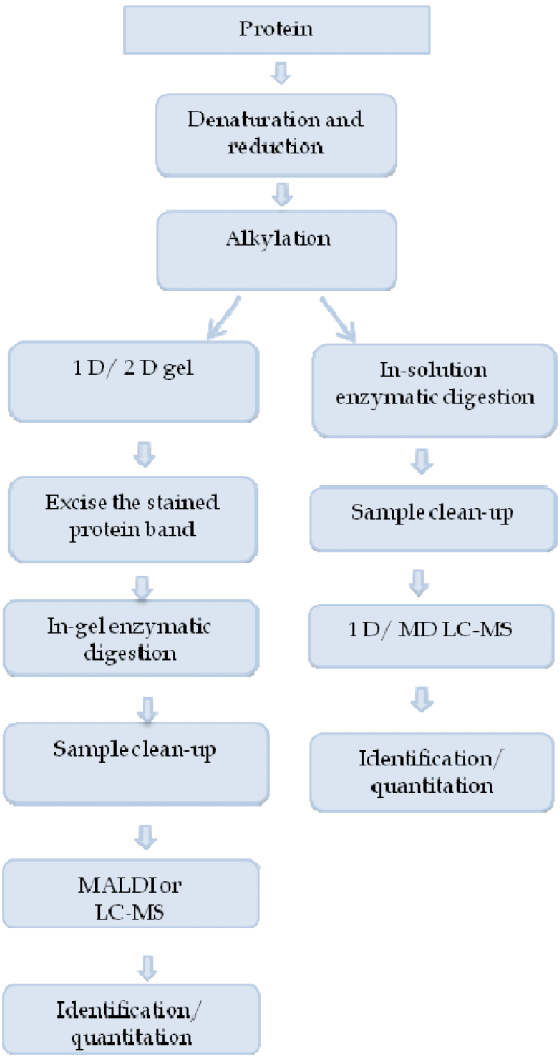

Fig. 11: Workflows of in-gel(left) and in-solution(right) digestion and subsequent LC-MS analysis on a protein sample[121]

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.

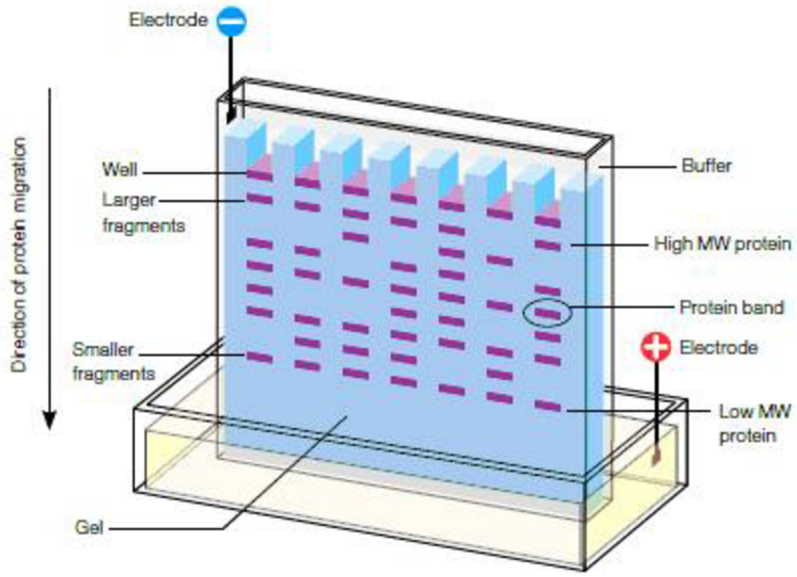

Fig. 12: Gel electrophoresis apparatus, the application of samples onto the gel.[122]

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.

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.

4.6.2 Isolation of glycosylated proteins

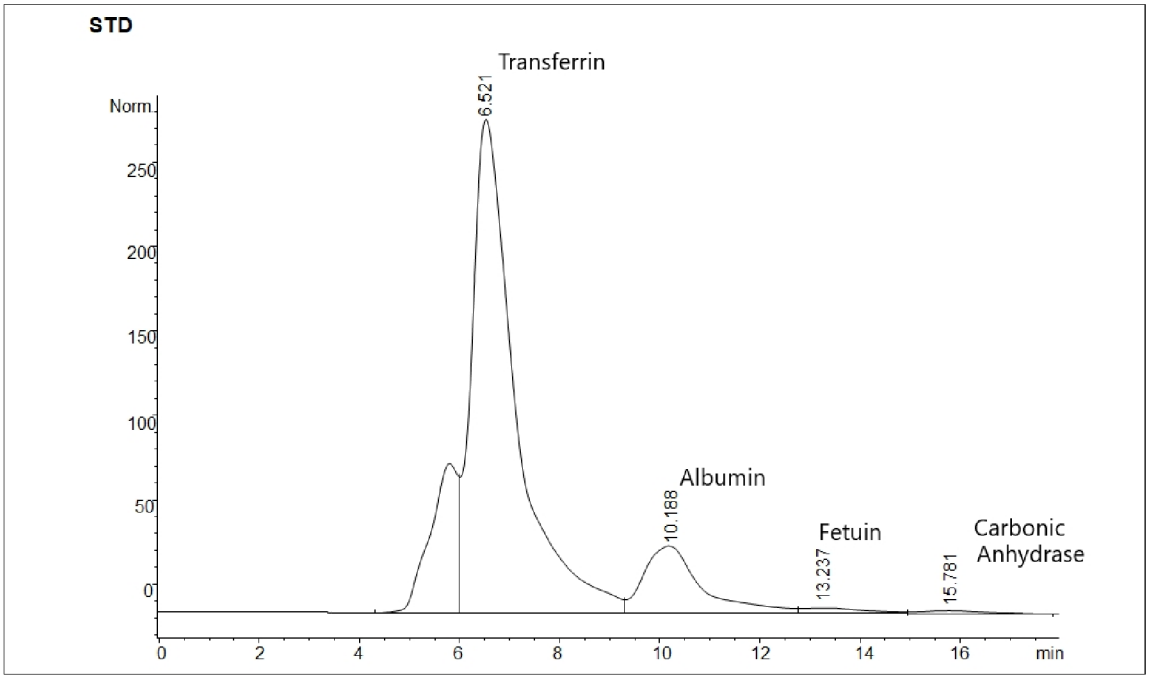

Fig. 13: Separation of the standard used, including compounds with known masses for reference

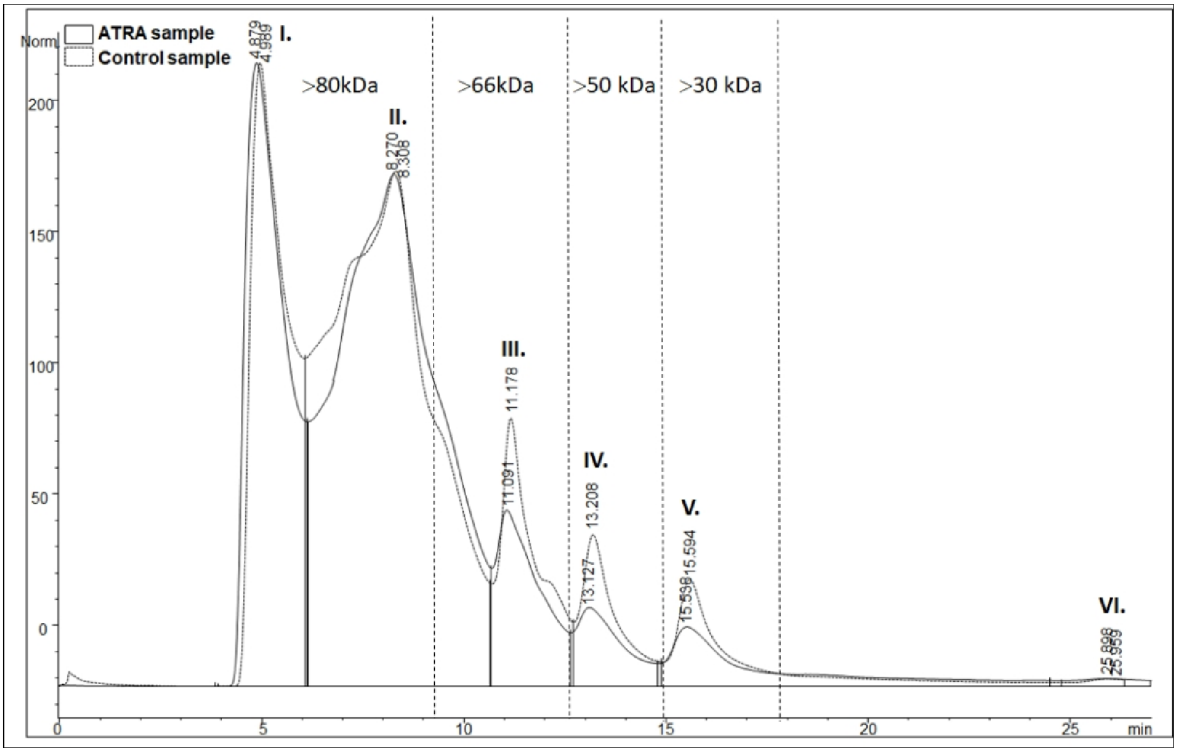

Fig. 14: Separation of ATRA and control samples with respective protein masses

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.

	TBB4A_HUMAN	50010	Tubulin beta-4A chain	Phosphoserine	

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.

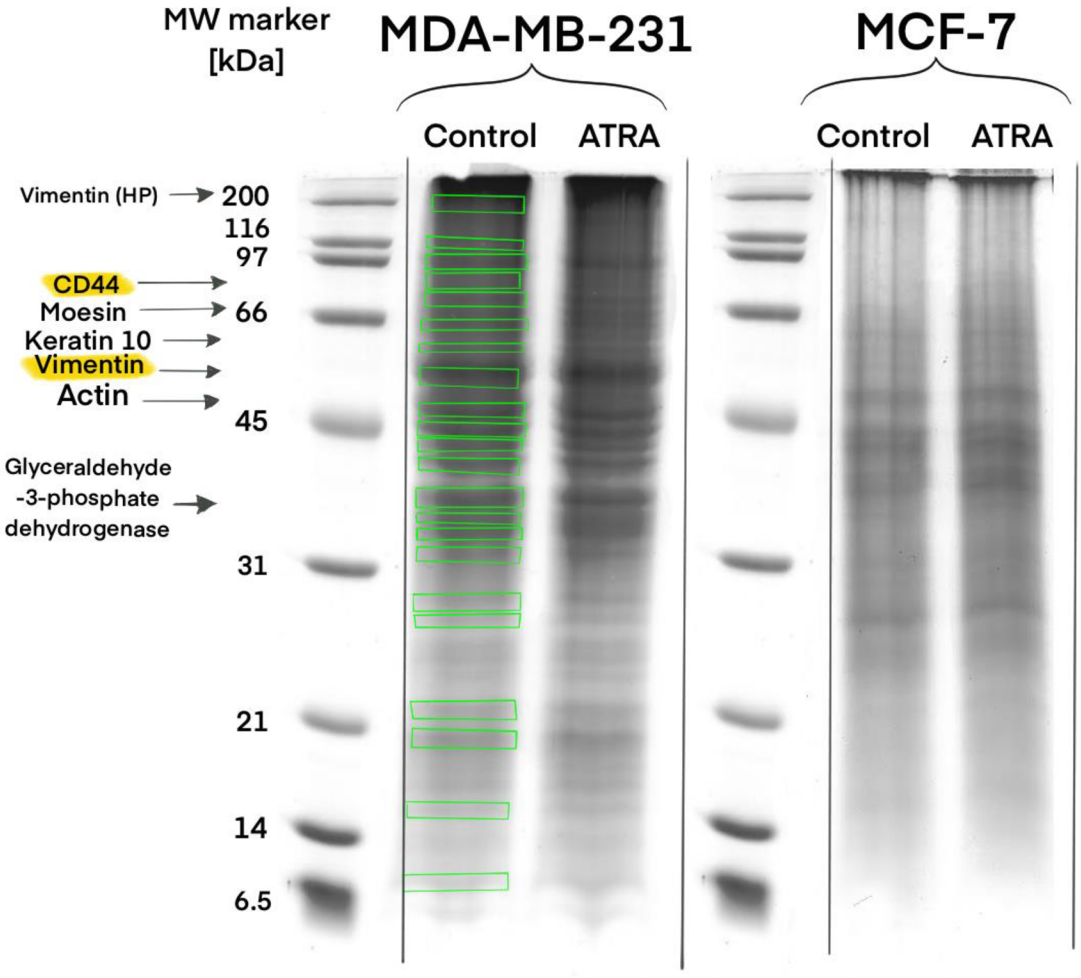

Fig. 15: Basic overview of the separated proteins using 1D SDS gel electrophoresis on whole cell lysates. The anticipated proteins CD44 and Vimentin may be spotted. Vimentin has a mass of 53681 Da, while CD44 has a mass of 82009 Da

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.

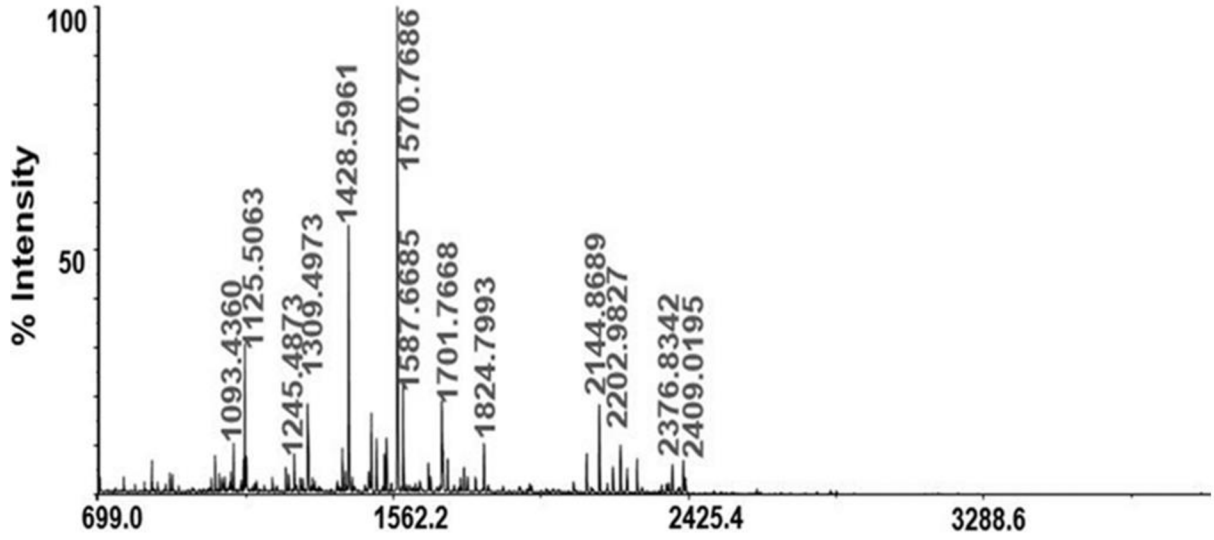

Fig. 16: The chosen MS spectrum of identified vimentin found in MDA-MB-231

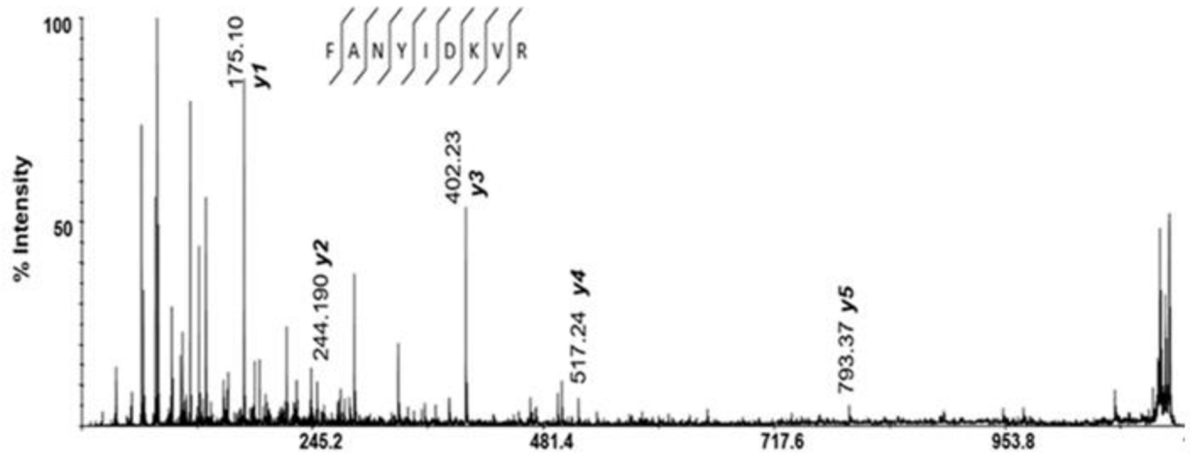

Fig. 17: An example of fragmentation MS/MS spectrum at 1125 m/z corresponding to Vimentin

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.

Y		unnamed protein product	51727	B4DRW1
Y		peptidyl-prolyl cis-trans isomerase FKBP4	51772	Q02790
Y		cytokeratin 8	53529	P05787
	Y			

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.

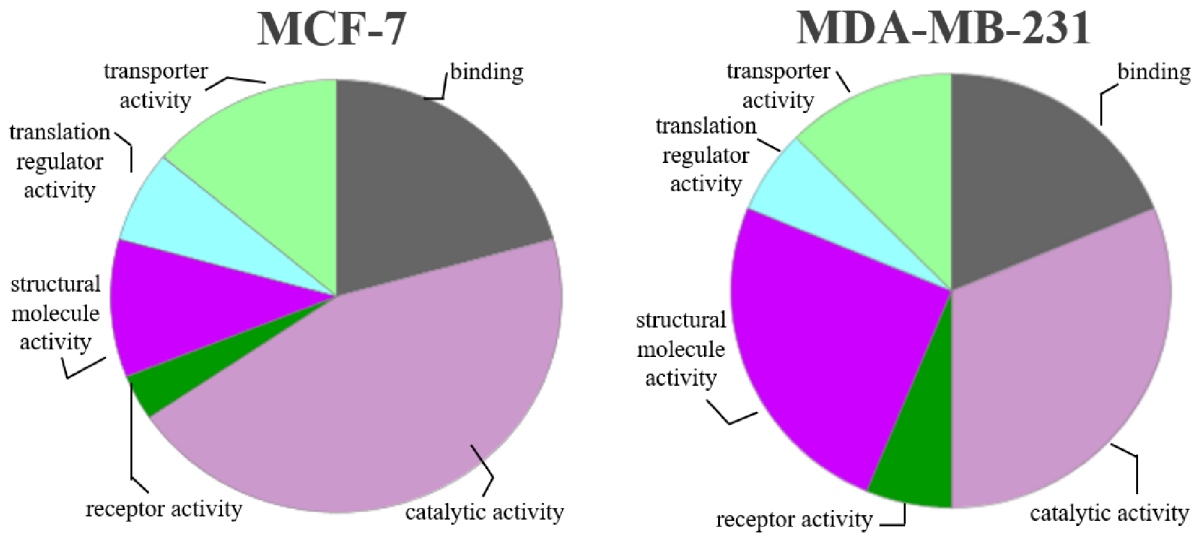

Fig. 18: Proteins categorization based on major biological, molecular functions and their potential association networks with known and predicted protein¯protein interactions

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]

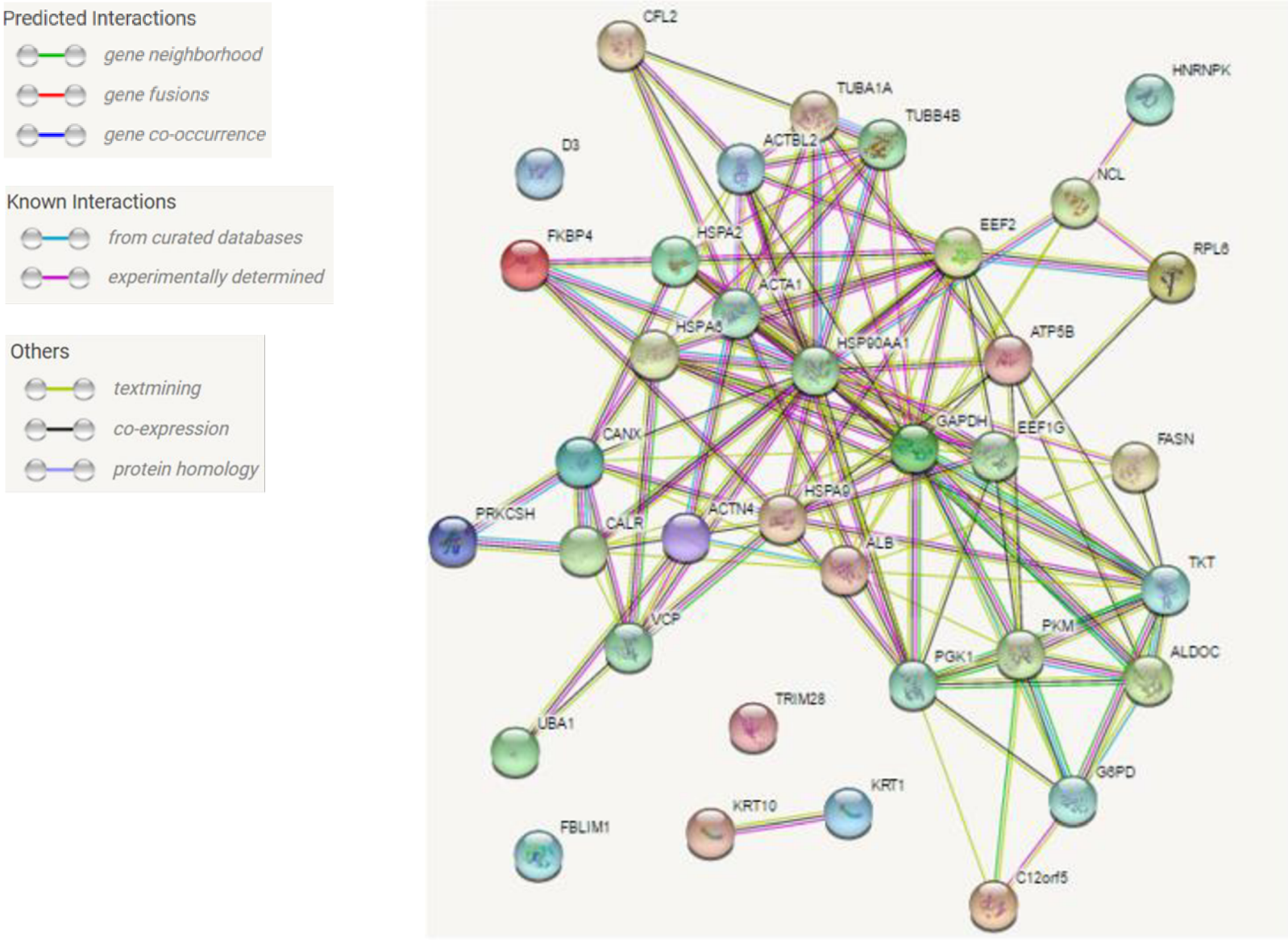

Fig. 19: An overview of protein-protein interactions in MCF-7, generated from STRING database

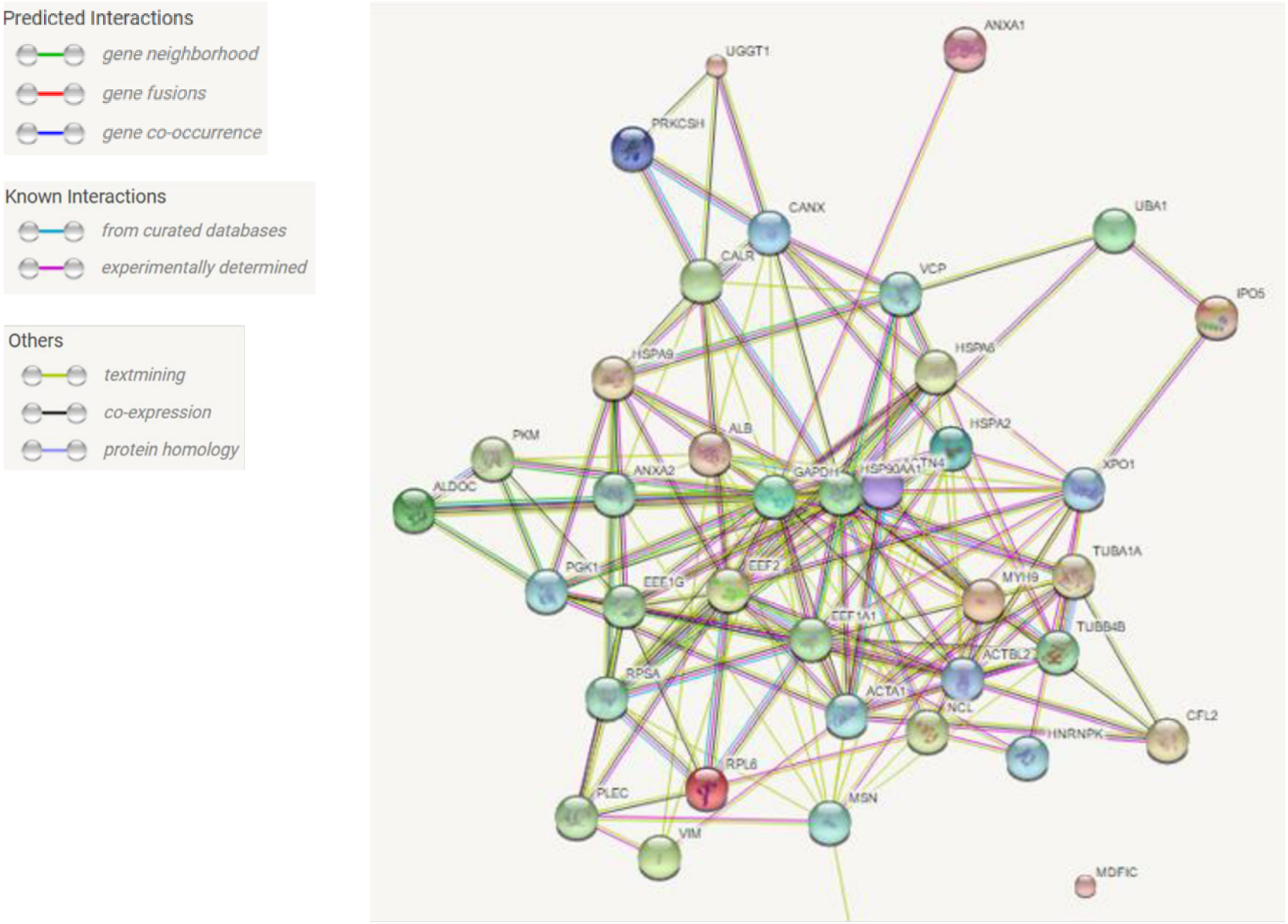

Fig. 20: An overview of protein-protein interactions in MDA-MB-321, generated from STRING database

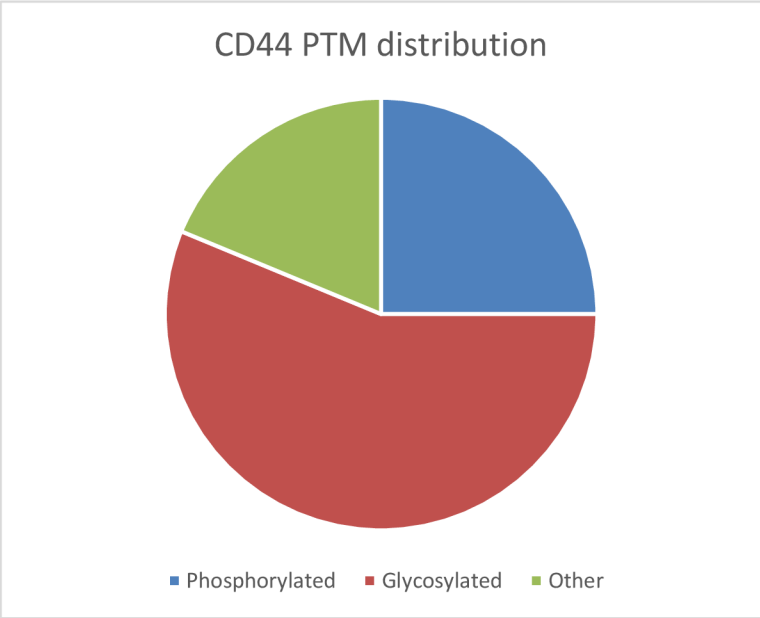

Fig. 25: CD44 PTM distribution

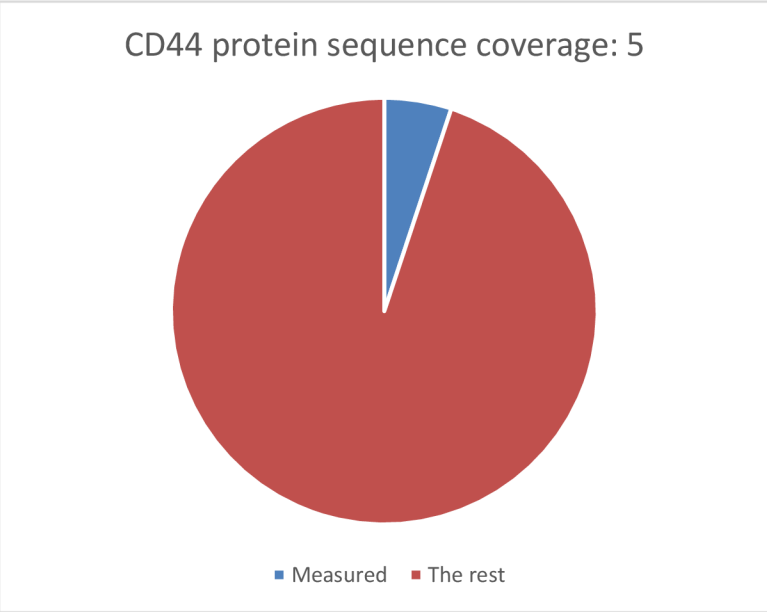

Fig. 26: CD44 protein sequence coverage

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).

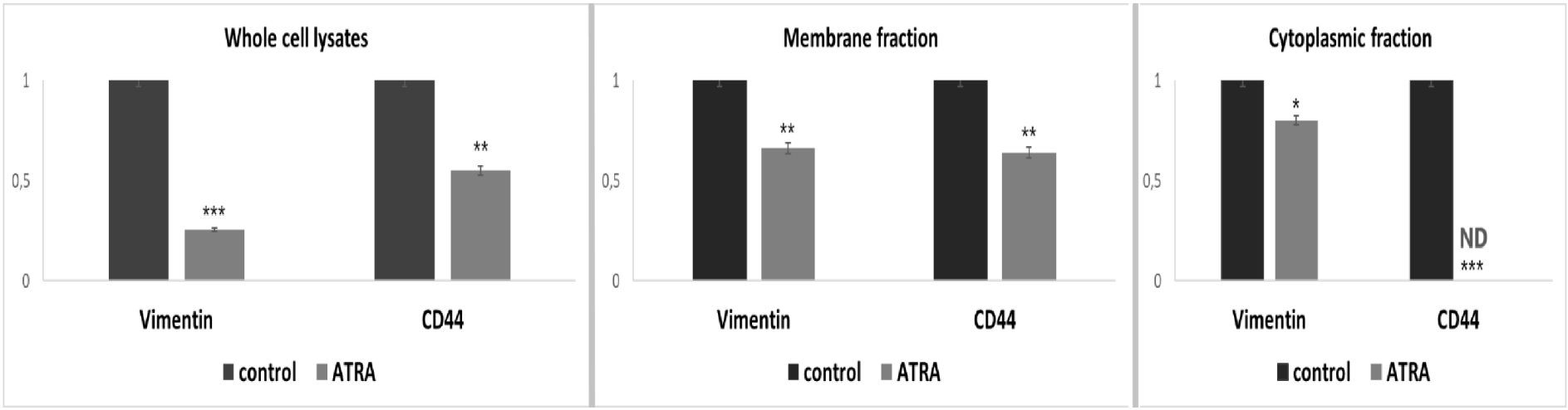

Fig. 29: Graphical expression of vimentin and CD44 proteins based on iTRAQ relative quantification assay. The ratio of each individual protein is expressed as an average value obtained from all identified labelled peptides corresponding to the given protein. The value of the control sample is set to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

[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

[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

[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

[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.

[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

[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

[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

[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

[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

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