PALACKÝ UNIVERSITY

FACULTY OF MEDICINE AND DENTISTRY

Program DSP: Oncology

CANCER PROTEOMICS IN CLINICAL AND EXPERIMENTAL ONCOLOGY

Tomáš Oždian, M.Sc.

Supervising department:

Institute of Molecular and Translational Medicine, Faculty of Medicine and Dentistry, Palacký University and University Hospital in Olomouc

Supervisor:

Petr Džubák, MD, PhD

Olomouc 2016

I hereby declare that this thesis has been written solely by myself and that all the sources used in this thesis are cited and included in the references part. The research was carried out in the frame of the Institute of Molecular and Translational Medicine, Faculty of Medicine and Dentistry, Palacký University Olomouc.

In Olomouc   ………………………

September 2016   Tomáš Oždian, M.Sc.

Acknowledgments

I would like to thank to my supervisor, Petr Džubák, M.D., Ph.D. for his valuable advices, willingness and help. My thanks belong also to Marián Hajdúch, M.D., Ph.D. for facilitating the laboratory work at the Institute of Molecular and Translational Medicine, with the great advantage of state of the art equipment and guidance throughout the research.

I would like to thank to all my colleagues from IMTM Olomouc, especially Lakshman Varanasi, who helped me with writing of articles and dissertation. My appreciation goes to my colleagues from the Department of Clinical and Molecular Pathology, Faculty of Medicine and Dentistry, Palacký University, Olomouc for great collaboration.

Last, but not at least I would like to thank my family and friends for their patience and motivation during all time of my study.

Olomouc   ……………………

September 2016   Tomáš Oždian, M.Sc.

Bibliografická identifikace

Jméno a příjmení autora	Tomáš Oždian
Název práce	Nádorová proteomika v klinické a experimenální onkologii
Typ práce	Dizertační
Pracoviště	Ústav molekulární a translační medicíny, Lékařská fakulta, Univerzita Palackého v Olomouci
Vedoucí práce	MUDr. Petr Džubák, Ph.D
Rok obhajoby práce	2016
Klíčová slova	Proteomika, hmotnostní spektrometrie, cisplatina, karboplatina, oxaliplatina, asporin,
Jazyk	Anglický

Abstrakt

Tato práce je zaměřena na proteomiku a její metody využívané při výzkumu léčiv a biomarkerů. V teoretické části jsou shrnuty hlavní body pracovních a analytických postupů moderní proteomiky. Experimentální část je rozdělena mezi dva projekty, které společně využívají většinu metod popsaných v teoretické části. První část diskutuje projekt proteomického profilování nádorových buněčných linií pro zjištění jejich odpovědi na léčivo. K tomu bylo využito více hmotnostně spektrometrických přístupů Proteomická analýza v tomto projektu odhalila, že hlavní buněčnou odpovědí na ošetření oxaliplatinou je nukleolární stres. Druhým projektem je přehled kvantitativních MS metod použitých k analýze hladiny potencionálního biomarkeru, asporinu. Specifická citlivost většiny metod na asporin byla velmi nízká. Důvody pro tuto nízkou citlivost jsou v této práci také diskutovány, nejpravděpodobnější je rezistence asporinu k trypsinové digesci.

Bibliographical identification:

Author's name and surname	Tomáš Oždian
Title	Cancer proteomics in clinical and experimental oncology
Type of thesis	Dissertation
Department	Institute of Molecular and Translational Medicine, Faculty of Medicine and Dentistry, Palacký University Olomouc
Supervisor	Petr Džubák, MD, PhD
The year of presentation	2016
Keywords	Proteomics, mass spectrometry, cisplatin, carboplatin, oxaliplatin, asporin
Language	English

Abstract — The thesis is focused into proteomic methods and their application in drug and biomarker research. Theoretical part reviews all major aspects of modern proteomic workflow and analysis. Experimental part is split into two parts covering together the large part of methods discussed in theoretical introduction. The first part discusses project of drug response determination by cancer cell line proteomic profiling. A suitability of different MS approaches for this kind of analysis is discussed here as well as successful determination of nucleolar stress as main response to oxaliplatin treatment. The second project of this thesis is an overview of MS quantification approaches in analysis of asporin protein. This protein showed itself to be particularly hard to analyse with reasonable sensitivity. The reasons of such low sensitivity specific to asporin are discussed here as well. The most probable reason is a resistance of asporin to trypsin digestion.

TABLE OF CONTENTS

1 INTRODUCTION

Proteomics is currently well-established method of bioscience. Basically, proteomics allows studying proteins in detail or in large scale identification experiments. Both approaches have their substantial merits. Knowledge of detailed protein properties is essential in deciphering protein function, activity or in description of protein inhibition by various drugs. Second approach, large scale analysis relies on broad knowledge of proteins, based both on experimental evidence of protein or genome sequencing and translation. This approach allows us to determine changes in single proteins, but as well to describe protein groups involved in particular processes. Those approaches are depending on each other. To be able to perform large scale analysis, we have to have enough information about particular proteins. And to study protein in detail, we have to have information about protein itself, about its localization and function in the cell.

Proteomics thus needs precise methods to separate, identify, sequence and describe proteins. Protein separation relies mainly on precipitation, affinity-based, electrophoretic or chromatographic methods. Separated proteins could be identified by specific antibody, by enzyme assay or by mass spectrometry. Protein sequence is important mainly for mass spectrometry (MS) identification and molecular modeling of protein structure. It could be determined by Edman sequencing or by MS approaches like de-novo sequencing. Further description of proteins could be done by elucidation of structure by X-ray crystallography, role in the cell by knock-out/overexpression experiments, description of binding factors by affinity purification and so on. Proteomics is thus very broad and complex method and it is difficult to cover all mentioned methods together. In this thesis, first mentioned methods – protein separation, identification and in limited extent as well sequencing will be discussed. All those steps are covered in current "shotgun" proteomics approach focusing on identification of as broad part of the sample proteome as possible.

This thesis is focused on human proteomics. Two projects using MS and antibody based methods for qualitative and quantitative protein analysis will be discussed in the experimental part.

1.1 Protein separation methods

Proteins in the cell or in serum represent a mixture which is complicated to analyze without any separation. There is a big number of proteins – about 10,000 different proteins in the cell and very broad concentration range – seven orders of magnitude in the cell and even ten to twelve orders of magnitude in the plasma [1]. The newest mass spectrometers are available to analyze samples up to six orders of magnitude [2] of concentration range. A good separation method is thus needed to separate and concentrate proteins of interest. This will reduce this concentration range and will increase probability of successful protein detection. Since each of methods discussed further has different drawbacks, it is beneficial to combine multiple separation methods for further extension of protein detectability. The most important separation methods in proteomics are electrophoresis and chromatography. Those methods are applied most often, but in certain cases it is useful to use other methods, like a protein precipitation or affinity chromatography.

1.1.1 Electrophoretic methods

Protein electrophoresis is a set of diverse methods based on protein mobility in electric field. It varies from starch or paper electrophoresis, gel electrophoresis used in clinical biochemistry to a set of currently used scientific electrophoretic methods. Basic electrophoretic methods used in current science are mainly gel based like native electrophoresis [3], isoelectric focusing [4], capillary electrophoresis [5] or most widely used denaturation electrophoresis using sodium dodecyl sulphate and polyacrylamide gel (SDS-PAGE) [6].

1.1.1.1 SDS-PAGE

SDS-PAGE electrophoresis (Figure 1) is currently the most often used electrophoretic method which allows protein separation according to molecular weight. Electrophoretic migration is in principle dependent on protein mass, charge and shape. In SDS-PAGE, charge and shape are reduced to uniform level by protein denaturation - proteins are denatured by SDS and heating prior gel loading and both gel and electrode buffer contains SDS [6]. This step destroys secondary structure of protein. Proteins are thus reduced to detergent coated linear form with uniform SDS to protein ratio about 1.4 to 1. Denaturation reduces different charges and shapes and makes electrophoretic separation dependent on protein molecular weight only [7]. Molecular weight determination accuracy is further increased by introducing two-gel system. This system consists from less dense stacking gel

and more dense separation gel. Those gels are prepared using buffers containing Tris-HCl and glycine with different pH. This pH difference together with stacking/separation gel boundary has key effect to protein band width and thus mass accuracy. pH of electrode buffer is 8.3 and glycine is negatively charged in this condition. When electric current is applied, glycine travels together with chloride ions into stacking gel, where pH is 6.8. At this pH, glycine becomes uncharged and starts to migrate slowly. Chloride anions, on the other side, hold their charge and migrate still with the same speed. Glycine and chloride creates this way a thin boundary of migrating ions in the stacking gel. Speed of protein migration is in the middle between chlorides and glycine and proteins are forced to migrate in this thin boundary. Situation will change in the moment, when proteins reach separation gel with pH 8.8. Glycine becomes anion again and boundary with chlorides is abolished. Proteins will further focus on the border of separation gel just because of higher density and bigger resistance of the separation gel. Low density of stacking gel allows proteins to migrate to this boundary more less as a single band and in separation gel starts separation according to molecular weight. Experimental set-up of SDS-PAGE result in sharp, well separated protein bands (Figure 1), which is one of the reasons of its wide popularity.
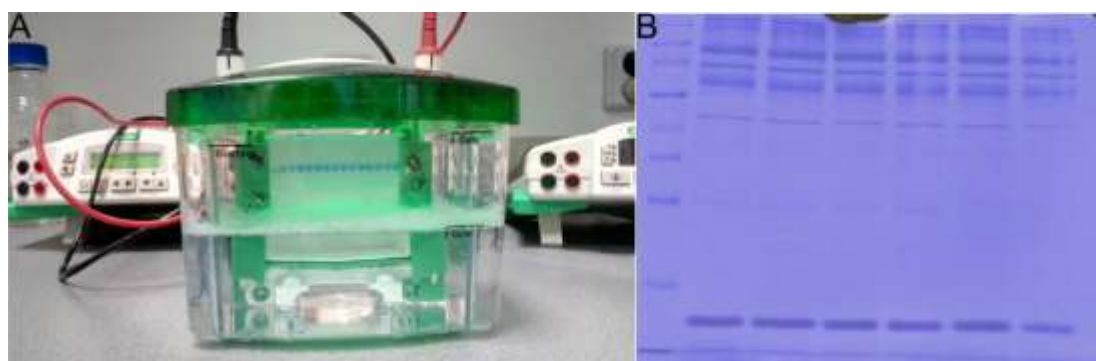
Figure 1: SDS-PAGE electrophoresis – its typical instrumentation (A) and representative Coomassie stained gel (B).

The main advantage of SDS-PAGE lies in its versatility. There is a huge selection of gel densities suitable to separation of proteins of different sizes even with possibility to purchase commercially available gradient gels ideal for separation of broad range of protein molecular sizes. Separated gels can be visualized by a variety of stainings, e.g. by silver or Coomassie staining. Proteins from gels can be specifically identified by transfer to a membrane and detection by specific antibody or by excission from gel, digestion and identification by MS.

1.1.1.2 Native electrophoresis

SDS-PAGE is not the only existing electrophoretic method. There are other gel-based electrophoretic approaches. The first, native electrophoresis is the closest to original electrophoresis setups. Proteins migrate in electric field in their native conformation without any modifications as in first electrophoretic experiments. This bring several benefits, proteins are after separation still in their original shape and in active form. Since there is no denaturation step in native electrophoresis, protein-protein interactions remains preserved and bound proteins are separated together. On the other side native electrophoresis is not much suitable for analytical preparation, since protein mobility is dependent not only on its molecular weight, but as well on protein shape or net charge. Although those factors are expressed in an equation [8], native electrophoresis cannot beat SDS-PAGE in determination of protein molecular mass. Native electrophoresis and its variants are used in protein purification when proteins in active form or protein-protein complex are needed.

1.1.1.3 Isoelectric focusing

Isoelectric focusing (IEF), is a variant of electrophoresis separating proteins according to their relative charge in pH gradient. Relative charge of protein is determined by its isoelectric point. This point is a pH value, where protein has no charge. When a protein is in acidic conditions, it bears positive charge and migrates to cathode. Conversely, protein in basic conditions bears negative charge and migrates to anode. If such a migration is done in pH gradient, protein will stop its migration at pH corresponding to its isoelectric point. pH gradient is made using ampholytes, which are compounds with polyamino and polycarboxy moieties, which have both acidic and basic properties. Based on composition of particular ampholyte, it will buffer pH at its own isoelectric point [9]. Ampholyte mixtures are very complex and thus are offered mainly commercially (e.g. Ampholine, PharmaLyte, Immobiline or Servalyt). Another choice in IEF analysis is pH gradients immobilized on gel strips (e.g. produced by Bio-Rad, GE Life sciences or Thermo Fischer Scientific).

1.1.1.4 Two dimensional electrophoresis

Above-mentioned approaches have limited resolution only. The worst resolution has the native electrophoresis, SDS-PAGE and gel-strips based IEF are better choices for protein separation. However, their resolution is still not high enough to cover whole protein distribution mentioned in introduction of chapter 1.1. The one way how to improve resolution of both methods is to utilize orthogonality of SDS-PAGE and IEF. This approach

has been introduced by O'Farrell in 1975 [10]. 2D electrophoresis (2DE; Figure 2) offers a possibility to separate up to 5000 proteins; however it suffers from low reproducibility caused by gel-to-gel variations. To overcome this, two main approaches are used. The first is to use multiple fluorescent stains and analyze all samples in one gel. This approach is called 2D difference gel electrophoresis (2DIGE, Figure 2) [11]. Second approach consists in software processing of scanned gels to obtain image overlay. There are number of companies focusing on this issue, for example Bio-Rad, Decodon, Non-Linear etc.
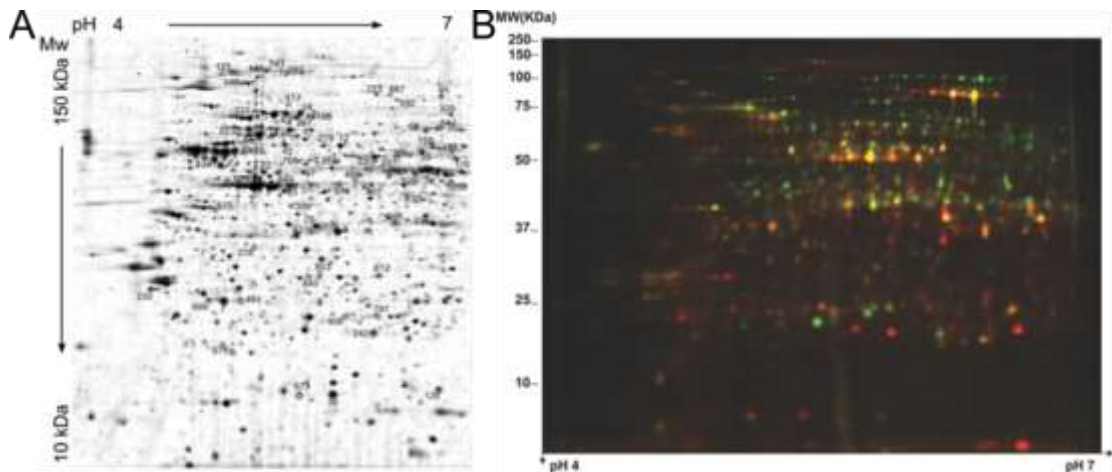
Figure 2: 2D Electrophoretic methods: classic 2D electrophoresis (A) and 2D DIGE (B). 2DE and DIGE figures were adapted from [12,13].

The strength of gel based methods is separation of intact proteins and ability to distinguish protein isoforms. Disadvantages are well-known and they are low dynamic range, poor resolution and gel to gel variations of 2DE gels. DIGE can overcome the problem of gel reproducibility but the dynamic range remains low. Frequent identification of differentially expressed highly abundant proteins (enolase1, heat shock protein or vimentin) is another restriction. We should be very careful with interpretation of identification of these proteins frequently involved in stress response or housekeeping processes [14].

1.1.1.5 Capillary electrophoresis

The most advanced electrophoretic method is capillary electrophoresis. Electrophoretic separation here occurs in a capillary, which offers similar properties as a capillary liquid chromatography (discussed below). Those properties are good resolution and high sensitivity. Capillary separation brings a slightly different principle of separation, which involves behavior of liquid in capillaries. Capillaries can be also filled with functional sorbents. Both issues are discussed in detail elsewhere in the literature, e.g. in specialized books. Capillary electrophoresis can be coupled to MS similar way to liquid chromatography

(LC). In this case, it is necessary to replace regular buffers to volatile. This set-up is used in proteomics as an alternative to 2-DE with MS detection or LC-MS [15]. Capillary electrophoresis is a less popular method in proteomics compared to LC-MS (595 references for "capillary electrophoresis proteomics" and 10413 references for "liquid chromatography proteomics" in PubMed up to July 2016).

1.1.2 Chromatographic methods

Another method of separation frequently used in protein analysis is chromatography. It is a method based on different analyte balance between affinities to stationary or mobile phase. Briefly, all analytes are dissolved in mobile phase. This mobile phase flows through porous stationary phase. Analytes begin to separate according their balance between affinity to stationary phase and solubility in mobile phase. Thus analytes with lower solubility in mobile phase and higher affinity to stationary phase will be more retained compared to analyte with high mobile phase solubility. The key for this type of analysis is a careful selection of stationary phase affinity mechanism and corresponding mobile phase. Best separation is obtained if, for example, stationary phase is hydrophobic and mobile phase is hydrophilic or vice versa.

Chromatographic separation of proteins is very important and versatile method for both preparative and analytical separations. Preparative separations are used in protein purification processes, where great care is taken not only to achieve natural conformation or protein activity, but to get good resolution, recovery, throughput or reproducibility of chromatography as well. The stationary phase for preparative experiments should be also stable, easy to maintain and inexpensive. Preparative chromatography uses mostly principles of affinity chromatography, immunoaffinity chromatography or gel filtration chromatography. Analytical liquid chromatography became a key step in protein / peptide separations in current proteomics [16].

The main difference between preparative and analytical chromatography is in scale and purpose of separation. Preparative chromatography is focused on obtaining bigger amount of target compound, whereas analytical chromatography is usually performed in much smaller amounts enough to prove presence and quantity of analyte. This can be illustrated at most often used analytical chromatography method - high performance liquid chromatography (HPLC). Classic HPLC uses columns with diameter in millimeters, most

often 4.6 or 2.1mm and flow rates less than milliliter per minute. Since its development, HPLC became the method of choice for analytical sample separation in both chemistry and life sciences.

1.1.2.1 HPLC in proteomics

Proteomic HPLC analysis brings several important advantages, such as superior separation power, sensitivity or easy connection to modern identification methods such as mass spectrometry. At the same time HPLC in proteomics is challenging because of two main reasons. The first and main drawback in proteomic analysis is a high concentration range already discussed in chapter 1.1. Lot of biologically very important proteins is present in very low concentrations and good sensitivity of the analysis is thus essential. Sensitivity of HPLC is, besides other, dependent on the inverse of the square of the two radii of the columns [17]. Higher HPLC sensitivity can be thus achieved with lowering of inner column diameter. Decrease of column diameter lead to increase of backpressure and thus is necessary to reduce mobile phase flow. Such reduce in flow don't have any influence on sensitivity of HPLC system, mainly with electrospray ionized MS detection [17]. Decrease of diameter and flow related to regular HPLC (flow in hundreds µl/ml to ml/min, column diameter in mm) led to development of capillary HPLC (flow in µl/ml, column diameter in hundreds of µm) or nano-HPLC (flow in hundreds of nl/min, column diameter in µm) used in current proteomics (Figure 3). Miniaturization of HPLC in proteomics also led to development of HPLC chips, which reduce dead volumes of regular instrumentation and which implement all important parts including sample loading, fluidics, column and electrospray ion source [18].
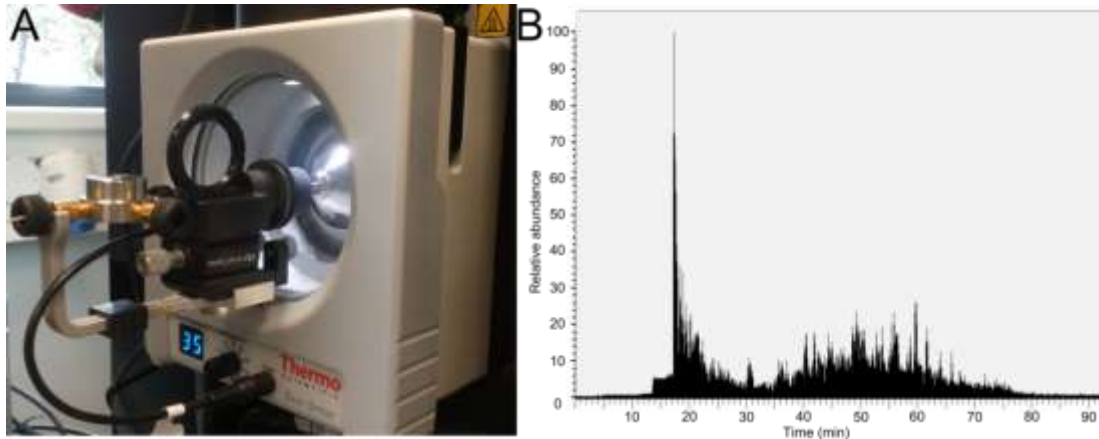
Figure 3: Modification of HPLC specific for proteomics: Nanocolumn (75μm x 15cm, 3μm, 100 Å pore size) packed together with nanospray (A) and total ion chromatogram of typical proteomic separation (first 100 minutes from 160 minutes run time shown; B).

The second main challenge of proteomic HPLC is sample complexity. Protein or peptide fraction of sample is relatively homogenous polymeric mixture differing only in sequence of 21 basic amino acids. Homogeneity of sample is further supported by tryptic digestion, when peptides have average length 8.4 amino acids [19]. Typical proteomic sample is thus very homogenous in composition, but very rich in single compounds – in this case digested peptides. For example in [19], authors were able to identify 27,822 peptides using trypsin alone and 92,095 peptides using multiple proteases. For successful separation of such sample is beneficial to employ several orthogonal separation mechanisms, for example electrophoresis or chromatography based on reverse phase, ion exchange or hydrophobic interaction retention mechanisms. From those, reverse phase based chromatography allows the easiest coupling to MS detection. Reverse phase HPLC is thus most often used method in LC-MS based proteomics. Separation power of reverse phase HPLC in proteomics can be increased by already discussed pre-separation by orthogonal method [16] or by using long gradients (Figure 3) [20]. Orthogonal methods in protein HPLC such as cation or anion exchange based or hydrophilic interaction liquid chromatography are usually used prior reverse phase separation.
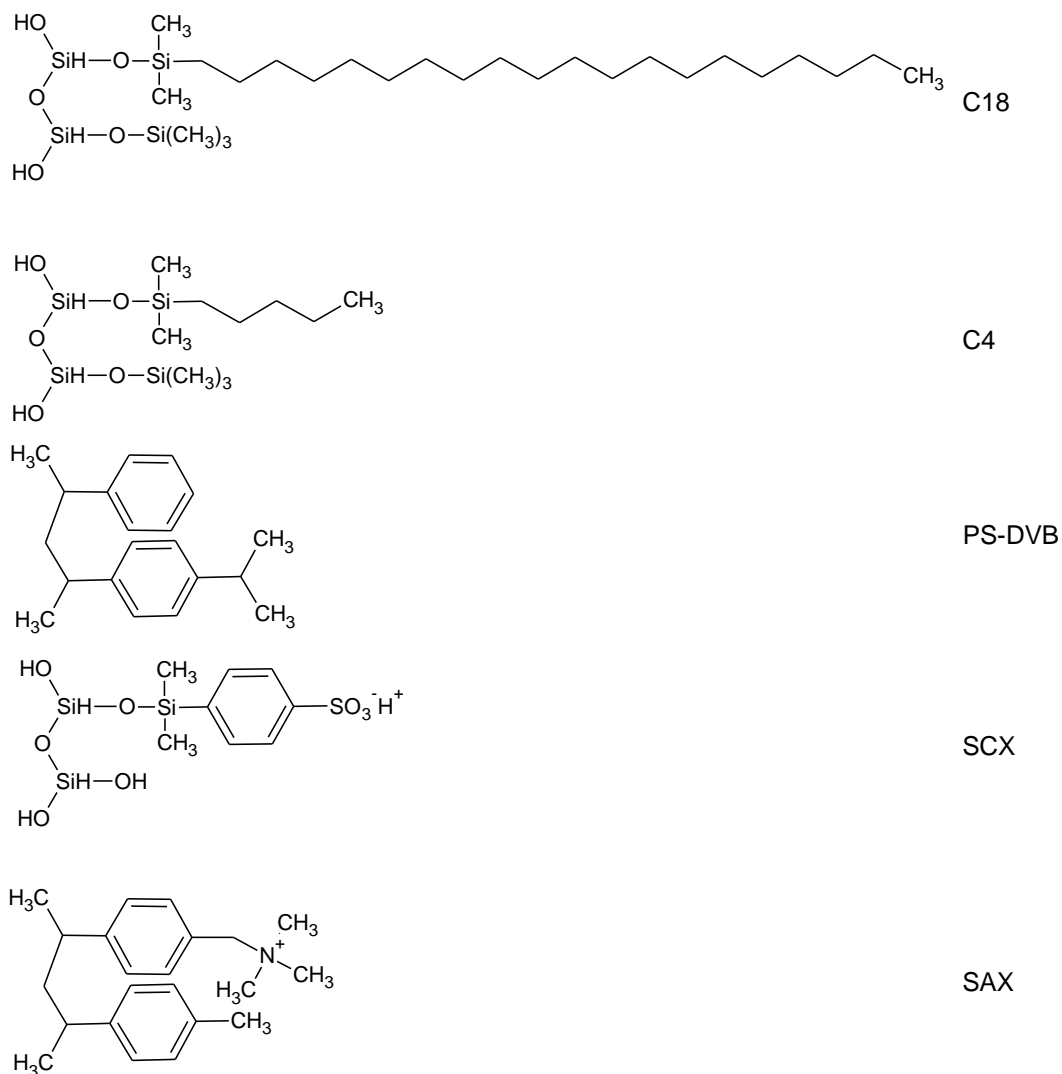
1.1.2.2 Reverse phase HPLC

Reverse phase (RP) separation model is not limited to proteomics. In fact, it has become a very widespread method and when is in current literature referred to HPLC or LC-MS, it is usually thought RP-HPLC or RP-LC-MS [16]. Principle of this method is separation of analytes based on their hydrophobicity between hydrophilic mobile phase and hydrophobic

stationary phase (Figure 4). The simplest hydrophilic solvent is water and elution strength of mobile phase is raised by addition of more hydrophobic organic solvents. The possibility to use water based buffers, which are cheap and environmentally friendly is together with versatility of RP separation one of keys to RP popularity. The most common organic phases are acetonitrile and methanol, although other organic solvents with different elution strength can be used as well. In proteomics, acetonitrile is preferred prior to methanol, since water-acetonitrile solutions have a lower back-pressure, lower UV cut-off and higher elution strength [21]. Both solvents are compatible with MS detection. The other important property of proteomic RP LC-MS analysis is a buffer selection. Buffers for proteomic analysis have to be volatile to be compatible with MS. According to isoelectric properties of peptides, it is beneficial to acidify mobile phase to enable uniform positive charge of analyzed peptides and thus enhance their MS ionization. Formic, acetic and trifluoroacetic (TFA) acids are used for such purpose.

Stationary phase in RP is made hydrophobic in various ways. The most common reverse phase resin is silica chemically coated with octadecyl aliphatic residues (C18 resin, Figure 4). C18 is strongly hydrophobic and it is well suitable for peptide analysis. Resins with octyl (C8) or butyl (C4, Figure 4) residues are less hydrophobic and better e.g. for analysis of intact proteins because less hydrophobicity doesn't cause extensive denaturation [22]. Coated silica represents classic resin used in RP-HPLC. As an alternative to classic bead filled column, polymeric or monolithic RP stationary phases offer a higher throughput and lower backpressures [23]. Methacrylate and polystyrene-divinylbenzene are the most popular polymer sorbents for RP-HPLC. Monolithic columns are made by polymerization of silanes or above-mentioned polymers *in situ* [23]. Both particle-based and monolithic resins have their disadvantages. Particle – based resins have lower surface area and thus retention, selectivity and therefore resolution are limited. Monolithic resins are casted directly into column – this may result to problems with reproducibility of pore structures. This results to high batch-to-batch variability of analytical performance. Furthermore, monolithic columns have generally weak mechanical stability. There is an additional issue with polymer monoliths, i.e., the potential swelling problems in the presence of solvents. A solution to those issues is core-shell sorbents. Core shell sorbents contain solid non-porous cores with porous shells. This allows joining advantages of both previous resins as is high resolution or lower backpressure. The biggest limitation of core shell particles is their complicated preparation [24].

HO
    SiH—O—Si—CH₃
O   CH₃
    SiH—O—Si(CH₃)₃
HO

C18

HO
    SiH—O—Si—CH₃
O   CH₃
    SiH—O—Si(CH₃)₃
HO

C4

H₃C

PS-DVB

HO
    SiH—O—Si—SO₃⁻H⁺
O
    SiH—OH
HO

SCX

H₃C
    N⁺

SAX

Figure 4: Common type of HPLC sorbents: octadecyl (C18) and butyl (C4) bound silica and PS-DVB (polystyrene-divinylbenzene) are sorbents used for RP-HPLC, sulfonium bound silica is for strong cation exchange (SCX) and quaternary ammonium for strong anion exchange (SAX) separation.

1.1.2.3 Alternative HPLC separation modes

Separation of proteins or peptides based on their hydrophobicity is only one of possibilities. Another possibility is their separation based on ion exchange. Cation exchange chromatography or strong cation exchange chromatography (SCX, Figure 4) can be used for that purpose. They are based on protein positive charge at low pH, which have a strong affinity to catex resin (Figure 4). Elution of bound proteins or peptides is done either by pH change (and protein/peptide charge loss) or by ionic strength increase by strong cations which have better affinity to resin. Both pH change and ionic strength increase are done with increase of mobile phase gradient. SCX could be used as stand-alone method for protein purification as well as pre-fractionation method adding next dimension to RP-LC-MS

analysis [25]. In this case, elution strength has to be done by volatile salts compatible with MS. Volatile salts has to be used in case of direct connection of SCX with RP as in MudPIT (Multidimensional Protein Identification Technology) protein identification strategy. Unfortunately, direct coupling of two modes of chromatography is rather complicated as each mode uses different solvent for increasing of elution strength [25]. SCX is also beneficial for phosphopeptide or phosphoprotein separation or selective enrichment of post-translational modifications (PTMs) [26]. SCX prefractionation is powerful method itself and provides a better results than pre-fractionation by SDS-PAGE or isoelectric focusing [27].

SCX is in literature preferred method for orthogonal two-dimensional LC separations. However, some other methods were tested with good orthogonality to RP as well. Those methods are strong anion exchange (SAX), hydrophilic interaction liquid chromatography (HILIC) or Electrostatic repulsion-hydrophilic interaction chromatography (ERLIC). SAX and HILIC are used as alternative to SCX with good orthogonality to RP-HPLC [28]. Those methods can be alternatively used for separation of charged PTMs, like phosphorylation [29], glycosylation [30] or protein nitration [31].

1.1.2.4 Detection in proteomic HPLC

HPLC is a powerful separation model, but there is a need for peak detection in eluent. The most often used HPLC detectors in proteomics are UV spectrophotometer or mass spectrometer. Spectrophotometer using UV lamp is non-destructive method and is less sensitive to salts than MS. It is thus beneficial in SCX/SAX fractionation experiments or as orthogonal detection method to MS. UV spectrophotometry is used mostly to monitor LC efficiency. On the other hand mass spectrometry offers higher sensitivity and ability to identify analyzed peptide. Mass spectrometry will be discussed later in this thesis in its own chapter.

1.1.3 Affinity purification

Another protein separation procedure is selective enrichment of protein (or peptide from digest) based on specific affinity to resin. This affinity is provided by physical or chemical bond of target analyte to this resin. Such bond can have various forms depending on nature of analyte and one analyte can be affinity purified by employing its different properties as will be shown later in this chapter. Affinity purification – sometimes termed affinity chromatography as well – is thus a very variable method allowing isolation of broad range

of proteins. It plays an important role e.g. in drug – target elucidation, phosphoproteome enrichment or in separation of recombinant proteins. This approach is widely used and contains lot of variations depending on particular experimental setup. It is not in the capacity of this thesis to describe all variations of affinity chromatography and this topic is reviewed elsewhere in literature as well as in our review (Appendix A). Basic principle of affinity purification method will be demonstrated on one of possible applications, enrichment of phosphoproteome. This is a very important application of affinity chromatography, because phosphorylation of proteins has important role in cellular regulation and in the same time phosphopeptides have worse ionization properties in MS due negative charge of phosphate. During regular proteomic experiment, phosphopeptides are detected with lower efficiency than non-phosphorylated peptides and their enrichment is thus necessary for any phosphoproteomic study.

1.1.3.1 Enrichment of phoshorylated proteins – an example of affinity purification

There are several ways how affinity purification of phosphorylated proteins or peptides can be done. The first decision in such an experiment is if there will be purified protein or peptide bearing phosphorylation. The enrichment of whole phosphorylated protein has definite benefits – e.g. more peptides from this protein increases probability of proper identification or localization of phosphorylation. The main drawback of this approach is a necessity to isolate protein from complex matrix and poor efficiency, when up to 80% of protein is lost during sample processing. This limits phosphoprotein enrichment only to abundant proteins [32]. Enrichment of phosphorylated peptide only is more often used, as it overcomes complicated approach of protein purification. After such enrichment, only phosphorylated fraction of peptides remains and proteins are identified with those peptides only. This could be complication, because some phosphorylation domains occur in multiple proteins and it is impossible to determine which of those proteins was originally phosphorylated [32].

Second decision in enrichment of phosphorylated proteome is a mechanism of enrichment. There are several modes of chromatography, which allows selective separation of phosphoproteome such as SCX, SAX or HILIC (Figure 5) [32]. Affinity enrichment can offer better selectivity than above mentioned HPLC modes. There are three main principles of enrichment methods based on affinity to phosphorylation. Phosphorylated proteins or

peptides could be enriched based on antibody specificity, affinity of phosphate group to metals or by chemical modification of phosphorylation [32].

1.1.3.2 Immunoaffinity enrichment

There are two main applications of antibodies in protein phosphorylation studies. Antibodies may be specific to phosphorylated amino acid, which could be serine or threonine, but also amino acids, where phosphorylation is less often, like a tyrosine. Those antibodies could be used in detection of phosphorylated proteins, e.g. in western blot or for immunoprecipitation of phosphorylated proteome. The main disadvantage of such an approach is in specificity of antibodies. Antibodies are only rarely specific only to phosphorylated amino acid, but often to surrounding peptide sequence as well. This reduces versatility of this method, but on the other side is possible to produce antibodies specific to particular protein phosphorylation space [32].

1.1.3.3 Metal-based affinity enrichment

Second affinity approach is based on affinity of phosphate to metals. The phosphate group, which has intrinsic negative charge is attracted to metal cations and retained with good efficiency. In this approach are used two main methods - immobilized metal affinity chromatography (IMAC) or metal oxide affinity chromatography (MOAC; Figure 5). IMAC is based on immobilization of metal cations to matrix. Such a matrix is in most applications iminodiacetic acid or nitrilotriacetic acid and typical metals for IMAC are Fe^{3+} or Ga^{3+}. The main advantage of IMAC is lot of available experimentally validated protocols originating in long tradition of this method [32]. MOAC affinity chromatography is based on simpler matrix than IMAC. Instead of immobilization of metal ions, MOAC relies on insoluble metal oxides, like titanium dioxide or aluminium oxide. In MOAC, titanium dioxide is preferentially used for phoshopeptide and aluminium oxide for phosphoprotein enrichment [32]. Those methods are used mainly for separation of phosphorylated peptides. Although they are widely used, both IMAC and MOAC aren't absolutely specific to phoshorylation. Acidic peptides with free carboxyl group are retained at IMAC and MOAC with good affinity as well [33].
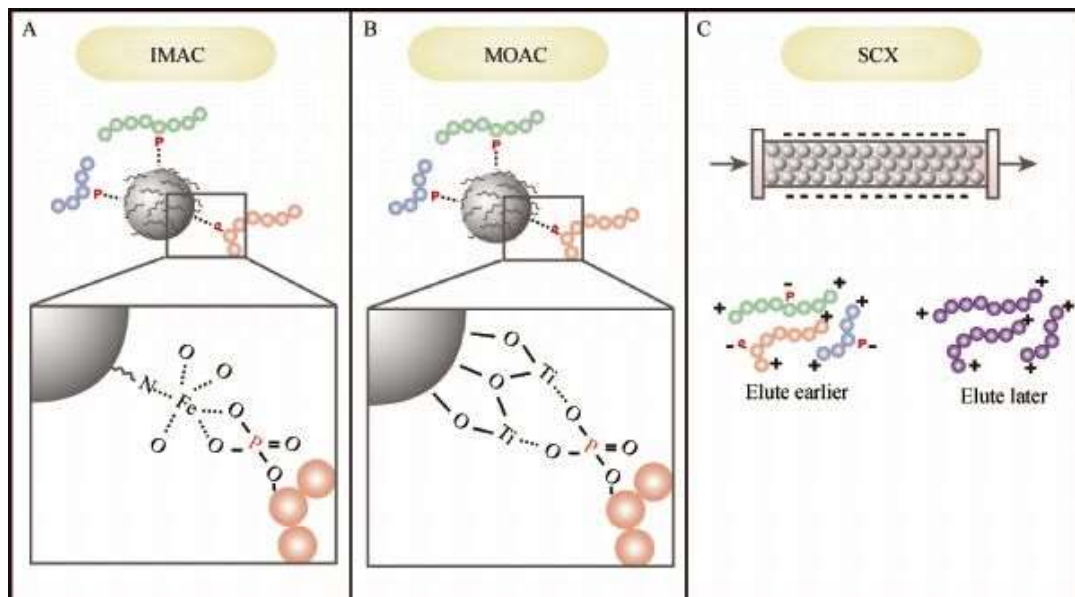
Figure 5: Principles of phosphoprotein enrichment methods: immobilized metal affinity chromatography (A), metal oxide affinity chromatography (B) and strong cation exchange chromatography (C). Figure adapted from [34].

The last method used in affinity enrichment of phosphopeptides is a chemical modification of phosphate group with resulting anchoring to beads. Compared to previous two approaches, this method allows superior specificity, however it complicates sample preparation with additional steps. Efficiency of different chemical modification has been also discussed [32].

Here, one of many applications of affinity chromatography was briefly discussed. Since it is not main aim of this thesis, I've picked phosphoproteome enrichment as a good and well characterized affinity approach important in current proteomics. Other affinity based method and their connection to mass spectrometry are reviewed by Rylová et al. (Appendix A).

1.1.4 Precipitation

Protein precipitation is one of basic method for protein separation. It relies on changing properties of protein's aqueous solvation layer. This layer masks protein's charge by layers of counter ions and prevents single proteins from direct contact by repulsive forces. If aqueous solvation layer is weakened, protein structure becomes more relaxed with exhibiting hydrophobic core and protein repulsion is weakened. Interaction of protein hydrophobic cores leads to formation of non-soluble multi-protein complexes and thus to protein precipitation. We can distinguish two basic protein precipitation approaches. The

first is a total protein precipitation, which is beneficial for isolation of total protein from matrix. This can be done by mixing sample with larger volume of organic solvent, like an ethanol or acetone or precipitating proteins with trichloroacetic acid (TCA). Organic solvents weaken power of aqueous and help with hydrophobic core exhibition whereas TCA precipitates proteins by lowering pH of solution. Three chlorines in structure of TCA also increases precipitation potential [35]. Both organic solvent and TCA precipitation have similar efficiency [36]. Another approach for protein precipitation allows even separation of different proteins by increasing concentration of salt. This method consists of adding inert salt, most commonly ammonium sulphate [37]. Ammonium acetate is kosmotrope, which means that it stabilizes water structure and reduces amount of water in protein solvation layer. When an optimal salt concentration is reached, protein of interest will precipitate. This kind of protein precipitation is reversible and allows purifying protein in native conformation with preserved activity.

1.2 Protein identification methods

In a previous section, protein separation methods were discussed. Such a separation needs to be visualized to check proper efficiency. The simplest visualization is in case of precipitation, where we see protein pellet after centrifugation. Electrophoretic and chromatographic methods can be visualized e. g. by staining SDS-PAGE gel by Coomassie or silver, or joining UV detector after chromatographic column. However, those approaches are non-specific and visualize all proteins. In this section, different strategies of protein identification will be discussed.

Proteins could be identified in various ways, however only three main are most versatile and universal. The first method relies on immune system property to recognize alien molecules in the organism. Alien molecules are recognized by variable chains of immunoglobulins. Variable chains contain a place called paratope, which is specific to certain place of target molecule, an epitope. Using antibody specific to certain protein allows identifying and quantifying of this protein in mixture. Methods based on antibody detection are simple and don't require expensive laboratory equipment. On the other side, highly specific and sensitive antibodies are necessary for those methods. Antibodies itself or as a part of diagnostic kits have higher price as well.

Whereas antibody based protein identification is a clever use of biology, second main approach, mass spectrometry, originates in analytical chemistry. MS was originally used to determine molecular weights of elements and its isotopes. Later it become widespread in analysis of organic compounds and with introduction of soft ionization techniques even for proteins, peptides, RNA and DNA. Although MS identification of proteins is dependent on complex and expensive instruments, it has increasing popularity due to its speed, sensitivity and relatively easy identification of complex protein mixtures.

The last method of protein identification lies in direct sequencing of protein or peptide. This can be done either chemically by Edman sequencing or by fragmentation analysis in mass spectrometer. Direct sequencing is method of choice if we are working with sample with little or no proteomic evidence. Without this evidence, there is no antibody or previously known sequence. Direct sequencing is thus the most straightforward approach in such conditions.

1.2.1 Antibody based identification methods

Antibody based protein detection is older than MS or gene sequencing. In the beginning of immunology, this method suffered from low specificity, which was solved by introduction of monoclonal antibodies [38]. Monoclonal antibodies are immunoglobulins specific to only one epitope. The production of antibodies of such purity was allowed by isolation and immortalization of single leukocytes. Advent of monoclonal antibodies allowed two strategies of protein identification. First strategy is to raise an antibody against specific protein. The main benefit is that antibody and its target protein are well described. When a target protein is not known yet, a second strategy consisting from immunization of animal with protein mixture and then producing a panel of monoclonal antibodies could be used. Screening of this panel could be tested e. g. in biomarker studies. Such a screening was responsible for discovery of part of biomarkers used in current medical practice [39]. An example of this could be carbohydrate antigen 19-9 (CA 19-9). This marker was discovered in 1981 by Koprowski et al. [40] as the antibody best distinguishing colon cancers from other bowel diseases or healthy volunteers.

Target protein- antibody bond has to be visualized. The common visualization methods are conjugation of antibody with enzyme, radioactive or fluorescent tags (Figure 6). There are two enzymes used for such labelling – alkaline phosphatase and horseradish peroxidase. Visualization is then done with specific enzyme substrate which will change in color or became luminescent by enzymatic reaction. Radioactive tag visualization is done mainly with iodine isotopes, which are gamma emitters with reasonable half-life [41]. In the case of fluorescent probes is slightly more complicated situation, because a lot of dyes are available and lot of vendors offer antibodies with fluorescent tags. Fluorescent or radioactive tags are used directly, without adding of additional agents.
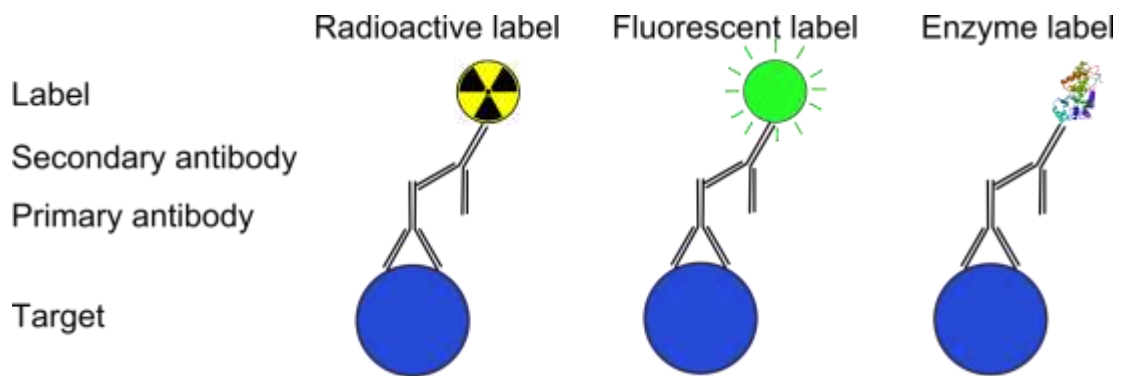
Figure 6: Labelling methods for antibody visualization – radioactive with collection of gamma emissions, fluorescent with collection of emitted light and enzyme labelled with collection of light emitted during enzymatic transformation of light emitting substrate.

Only one labelled antibody is used only very rarely. Usually there is common to use two antibodies – first antibody recognizing the actual antigen and secondary labelled antibody specific to stable chain of this primary antibody. This approach was developed in '70s [42] and brings substantial benefits to the analysis. Secondary antibodies helps promoting the sensitivity of the assay, eliminating the non-specificity for many analytes and thus decreasing the relative standard deviation of the assay or reducing dosage of the antibody [43].

Antibody based methods can be divided into two main branches – to immunoassays and immunomicroscopy. Immunomicroscopy is protein identification and visualization method. Protein is here detected on microscopic slide. Together with information about protein presence, we obtain even information about protein distribution in tissue. However, protein quantification in immunomicroscopy is challenging and requires skilled personnel. Immunoassays are, on the other side, focused mainly on protein quantification. Immunoassays are done usually in liquid and thus information about protein distribution in tissue is lost.

1.2.1.1 Immunoassays

Immunoassays are laboratory assays involving antibody for protein detection and accurate quantification. These assays are divided based on detection system used e. g. to radioimmunoassay or enzyme-linked immunosorbent assay (ELISA; Figure 7) are routinely done for protein quantification in both research and clinical practice. Radioimmunoassay is an older method, determining the level of antigen by level of radioactivity released from antibody conjugated with radioisotope [44]. This method is simple and reliable, and is

clinically used e.g. in analysis of allergens [45] or prostate specific antigen [46]. Radioisotopes however require specific equipment and are subjected to regulations. Therefore different approaches were developed to overcome need of radioisotopes. The most known of these approaches is ELISA. This method became widespread because of its simplicity, accuracy and sensitivity [47]. There is a big offer of commercially available kits for determination of plethora of proteins for both scientific and diagnostic purposes.
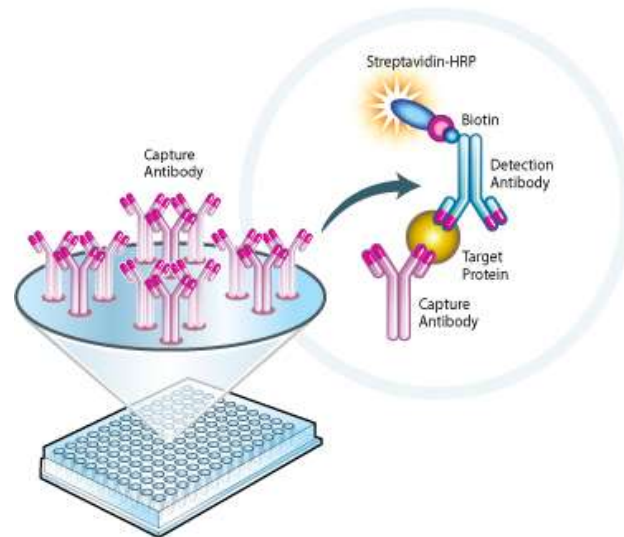
Figure 7: The illustration of enzyme-linked immunosorbent assay principle. Figure adapted from [48].

1.2.1.2 Western blot

Another type if immunoassay, which is as well very widespread and is responsible for discovery of many protein regulations and interactions, is immunoblot or Western blot (WB). Main difference of WB compared to other immunoassays is in previous protein separation by SDS-PAGE or 2DE. WB is thus a method of identification and quantification of gel separated proteins.

Western blot is a simple method relying on transfer of proteins from SDS-PAGE gel to nitrocellulose membrane developed by Towbin in 1979 [49]. The main advantage of WB is that proteins are transferred from space (inside gel) to surface (of membrane; Figure 8). Thus they become more concentrated and better available for antibody detection. WB is the method of choice for comparison of protein expression in various conditions and small amount of samples. Its main advantage is simplicity and low demands to laboratory equipment. The popularity of this method can be demonstrated by fact that original Towbin article has been cited more than 10,000 times and WB has been used in even more articles

without reference. On the other side, weak sides of WB are reproducibility, which is in original WB setup rather poor [50], and protein quantification with low precision. Protein quantification is in original setup claimed only as semi-quantitative [49]. In this chapter, key features of WB will be introduced, such as membrane, antibody detection methods and evaluation of WB.
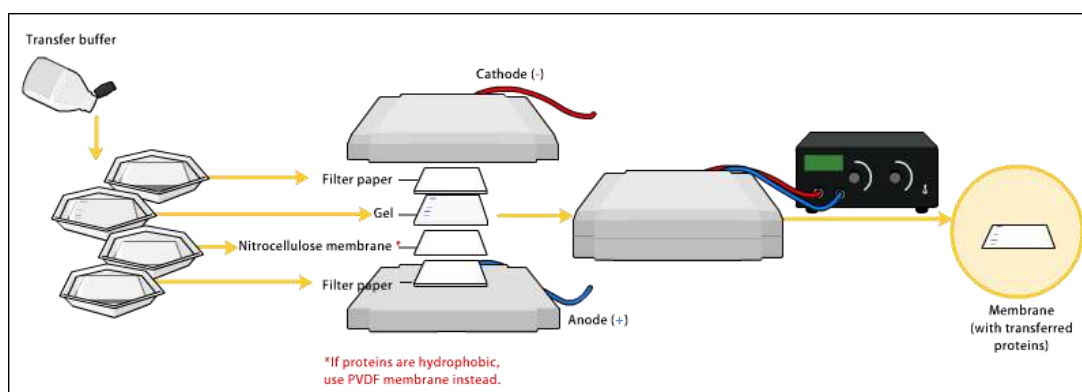
Figure 8: Illustration of protein transfer step in western blot. Figure adapted from [51].

Membrane is the key point for WB. It has to have a good binding capacity to proteins and it cannot allow proteins to migrate through. In original setup, there was used nitrocellulose membrane with 0.45 μm porosity. This membrane is still used. However, there are several options how to increase binding capacity and reduce migration through membrane. The first option is to reduce pore size, usually to 0.2 μm. This membrane retains proteins better than membrane with larger pores. Another possibility is to use another type of membrane. Nylon [52] or polyvinylidene fluoride (PVDF) [53] were originally tested for this purpose and both materials showed better mechanical and binding properties than nitrocellulose. Whereas PVDF is used today in approximately equal level as nitrocellulose, nylon membranes are used preferentially for transfer of nucleic acids [54].

Another important issue in WB is detection. Original system relies on horseradish peroxidase conjugated secondary antibody and recording of chemiluminiscence to X-ray film. This approach is quite sensitive and not requiring advanced equipment. On the other side, photographic detection of chemiluminiscence is labor intensive, difficult to optimize and only limited concentration range can be meaningfully recorded at one exposition time [55]. A solution to limitations of photography was introducing of Charge-Coupled Device (CCD) cameras. Those cameras are recording digital image, thus it is more difficult to obtain oversaturated bands which complicates precise quantification [55]. A solution of other problems common to chemiluminiscence, e. g. unequal distribution of chemiluminiscent

reagent at membrane, is introduction of fluorescently labelled secondary antibodies. This approach brings several advantages as better quantification precision, stable signal even from stored membranes or possibility to detect multiple proteins at one membrane at once [56]. The only drawbacks are higher prices of fluorescent antibodies and autofluorescence of nitrocellulose and PVDF membranes. There is constant development for membranes with lower autofluorescence [57].

Western blotting quantification accuracy is strongly dependent on experimental setup. With classical chemiluminiscence detection and exposition to X-ray films it suspected as only semi-quantitative method [49]. The accuracy grows if CCD camera overcoming saturated peaks is used, however WB in this is still considered as mostly semi-quantitative. A different situation is with using of fluorescent antibodies, which brings western blot really quantitative [58]. Quantitative information of WB is strongly dependent on protein amount loaded to membrane. Variations of protein load are usually corrected by normalization of analyzed values to levels of selected housekeeping protein – ß-actin, tubulin or glyceraldehyde-3-phosphate dehydrogenase. This approach has, however, some limitations, which is different levels of housekeeping proteins among different cells or conditions (e.g. age) or limited accuracy [59]. As an alternative, staining of total protein load was suggested. Initially this type of normalization was done with Ponceau S [60], Sypro Ruby [61] or Coomassie stain [62]. Recently, Bio-Rad introduced a novel technique of total protein normalization called Stain-free [63]. This type of staining consists of covalent bond of stain to proteins after UV activation. This stain is diluted in electrophoretic gel and first imaging step is done after electrophoresis. This allows monitoring efficiency of transfer as well. Stain-free normalization strategy is shown at Figure 9.
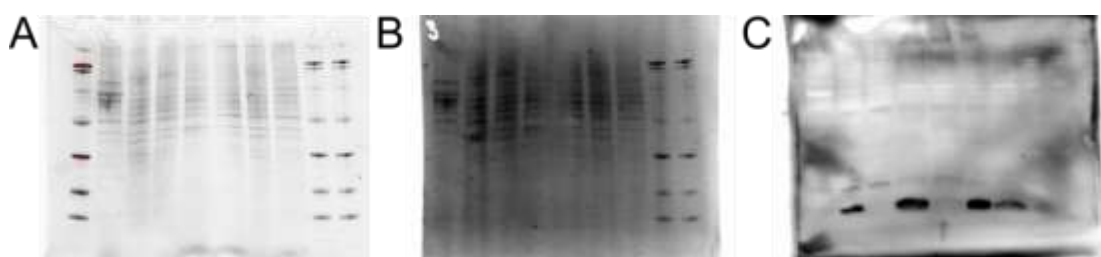
Figure 9: Stain-free strategy for total protein normalization. Stain-free dye is activated in the electrophoretic gel (A) by UV light. This serves as a control of quantitative transfer to membrane (B). Total protein lane is then used as normalization for immunodetection of target protein (C).

1.2.1.3 Immunomicroscopy

Immunomicroscopy allow visualizing protein localization and showing its approximate quantity. Typical representants of this category are immunohistochemistry (IHC) and immunofluorescence microscopy (IF). IHC become popular tool for visualization of protein localization for more than 70 years [64]. IHC is still used in research and routine diagnostics in molecular pathology [65]. Immunofluorescence is special application of IHC. In IF, fluorescently labelled secondary antibodies are used to visualize intracellular localization of primary antibody's target antigen. This approach is particularly beneficial in connection with super-resolution or confocal microscopy [66]. Typical outputs of IHC and IF are shown at Figure 10.
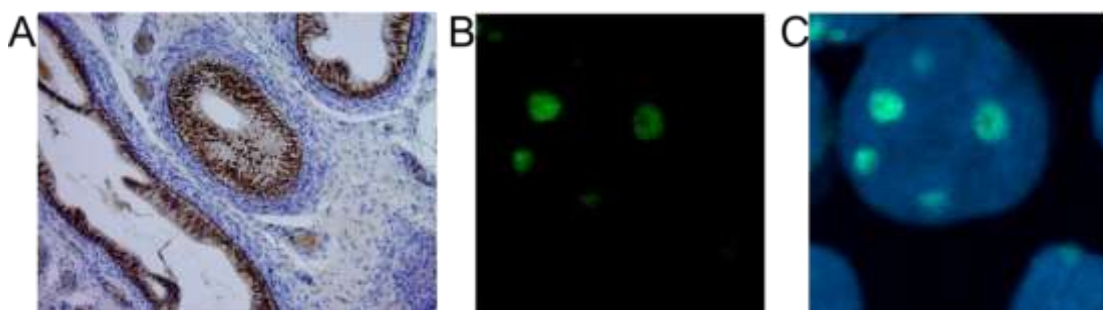
Figure 10: Tissue and cellular analysis by IHC and IF. At IHC picture (A) detection of Rb (4H1) (Cell Signaling) antibody on formalin fixed paraffin embedded human tissue with 100x magnification. IF picture (B) shows intracellular localization of UTP11L protein either alone or merged with DAPI staining of nucleus (C). Detailed description of sample and used antibodies are in Chapter 2.2.1.

1.2.2 Mass spectrometry based methods

Mass spectrometry, especially in combination with capillary or nano-LC, has become a golden standard in proteomics [67]. The key to success of this technology lies in its versatility with easy multiplexing of analyses. LC-MS allows relatively fast and reproducible protein identification and quantification. The output of LC-MS proteomic analysis depends on used instrument, which could be lower resolution instrument like an ion trap mass spectrometer. But currently most expanded approach is the high resolution MS done on precise time-of-flight or Fourier transformation mass spectrometry (ion cyclotron resonance or Orbitrap). High resolution MS is based on obtaining two kinds of information – even MS^2 from the ion trap before Fourier transformation analyzer and accurate mass which simplify peptide identification [68]. The main drawback is that these advanced

instruments are expensive. Optimization of LC/MS method is quite complex and can be taken as drawback as well.

MS can be also used as a stand-alone method, mainly for identification of isolated proteins. Stand-alone MS identification without fragmentation analysis is called Protein mass fingerprinting and is used for identifying proteins separated by 2DE, immunoprecipitation, affinity purification or similar method [69]. Protein mass fingerprinting is simple and fast method, mainly because it lacks long LC separations. On the other side, proteins are identified with lower probability score and identification of more proteins in mixture is complicated.

In this chapter, mass spectrometry and its applications in current proteomics will be briefly reviewed.

1.2.2.1 Principle of MS

Mass spectrometry is method which determines mass of ions in gas phase. Sample of any origin thus must be ionized and transferred into gas phase. This is done in ion source, an entry device of any mass spectrometer. The ions entering to mass spectrometer are in mixture. This mixture is separated according to mass of single compounds in mass analyzer. Single compounds are then detected in ion multiplier transferring ion impact to electric signal.

1.2.2.2 Ion source

Ions source is a device which transfers analytes in various environments to charged ions in gas phase. There are many principles how to do it. Electron ionization (EI) and chemical ionization (CI) are the oldest methods. In case of EI, analyte is already in gas phase (or in vacuum) and is hit by electron from electrode. This leads to ionization. However, this method of ionization is considered as hard because it is leading to extensive fragmentation of analyte [70]. CI is much gentler method, where ionization is done by collision with small molecule, like a methane or ammonia. This leads to less extensive fragmentation and clearer spectra [71]. Those methods need analyte in gas phase and they are used for analysis of organic compounds in connection to gas chromatography.

EI and CI are not usable for bioanalysis from two main reasons. First has been already mentioned, it is need of analyte in gas phase. Second, and main reason, is that those methods transfer amount of energy leading to fragmentation of even low molecular organic compounds. Thus proteins or peptides can't be analyzed because of extensive

fragmentation. This has been solved with introduction of matrix assisted laser desorption/ionization (MALDI) and electrospray (ESI) ion sources by Karas and Fenn [72,73]. Those new ion sources dramatically widened applications of MS. Moreover, they were so important for world science, that Tanaka [74] and Fenn won Nobel prize in chemistry in 2002.

MALDI ion source depends on co-crystallization of analyte with organic compound called matrix (Figure 11). This occurs on steel plate called target. Target is then introduced into evacuated ion source of mass spectrometer. Crystals of analyte and matrix on MALDI target are then shot by laser. Matrix absorbs energy of laser, transfers the energy and charges the analyte. Charged ions are then forced to leave crystals to mass spectrometer by applying current to target. Laser used in MALDI can have ultraviolet or infrared wavelength. For proteomics, UV laser is used almost exclusively. Matrix has to absorb very well in wavelengths of used laser. The most often used matrices for proteomic applications are α-Cyano-4-hydroxycinnamic acid, 2,5-dihydroxybenzoic acid and sinapinic acid. MALDI usually produces singly charged ions, as single proton or electron is transferred by matrix [75].

ESI ion source depends on spraying the charged liquid in stream of hot gas, usually nitrogen. Thus liquid evaporates to the point, where charged ions are very dense and repulsive forces of ions with same polarity will prevail to forces of droplet surface tension. This leads to so called Coulombic explosion and release of ions to gas state. Ions are transported to mass spectrometer by ion optics (Figure 11). ESI could be easily connected to LC-MS separation to enhance analytic power. ESI usually produces multiply charged ions [76].
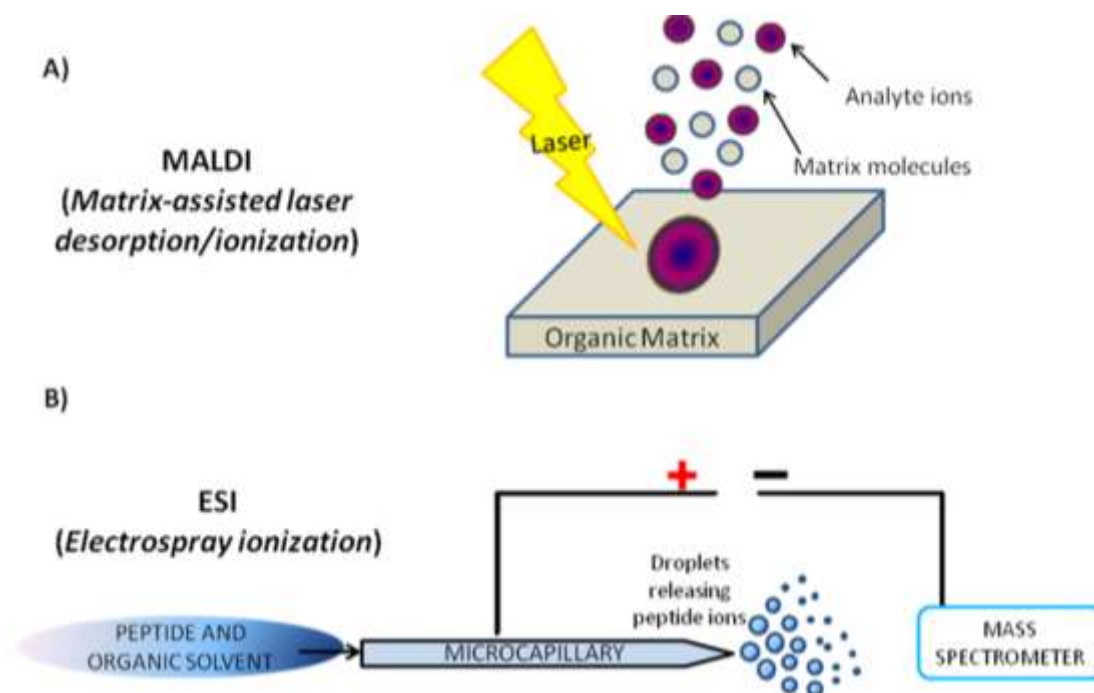
Figure 11: The schematic representation of MALDI (Matrix-assisted laser desorption/ionization; A) and ESI (Electrospray ionization; B) ionization method. Figure adapted from [77].

The MALDI and ESI are the most often used ionization techniques in bioanalytical MS. However, there are different "soft" ion sources as well. Atmospheric pressure chemical ionization (APCI) is used mainly for analysis of less-polar low-molecular weight compounds, e. g. carotenoids [78], analysis of drugs [79] or in lipidomics [80]. Atmospheric pressure photoionization (APPI) takes place as complementary approach to ESI and APCI in analysis of drugs, natural products or in environmental analysis [81]. Although those ionizations techniques are soft, they aren't used in proteomic analysis. The last technique mentioned here is desorption electrospray (DESI). DESI is basically variant of electrospray which allows imaging analyses of lipids [82], drugs and metabolites [83] or proteins [84] as well.

1.2.2.3 Mass analyzer

Mass analyzer is an important part of mass spectrometer in which ions are being separated according their mass. This separation could be done according their physical (time-of-flight) or electro-magnetic properties (e. g. quadrupole, ion trap, magnetic sector etc.). Mass analyzer could be either simple (time-of-flight, ion trap) or hybrid (triple quadrupole, quadrupole-time-of-flight, orbitrap). In proteomics, there is a big interest for using high resolution analyzers like a time-of-flight, ion cyclotron resonance or orbitrap because of sample complexity and accuracy of identification. Some low resolution instruments, like an

ion trap or triple quadrupole are used due to their advantageous properties as well. For example, ion trap allows collecting fragmentation spectra with fast and with good sensitivity and triple quadrupoles are method of choice for MS based quantification. Magnetic or electrostatic sectors aren't usually used in proteomics.

Time-of-flight (TOF) mass analyzer is the simplest of currently used mass analyzers. Basically it is an evacuated tube. Ion will enter on one side and will free fly to another. Because all ions will get the same energy at the ion source, their velocity will be dependent on their mass; ions will thus arrive to detector in different times [85]. Disadvantage of single TOF analyzer is bad resolution, as can be seen e. g. on analysis of big molecules. This disadvantage is caused by spread of energy obtained in ion source and thus ions of the same mass create clouds. This is solved by reflectron, the ion mirror which neutralizes ion clouds and makes ions more focused. It also doubles the length of flight path making resolution even better [86]. The example of molecule analyzed in linear and reflectron mode is shown at Figure 12.
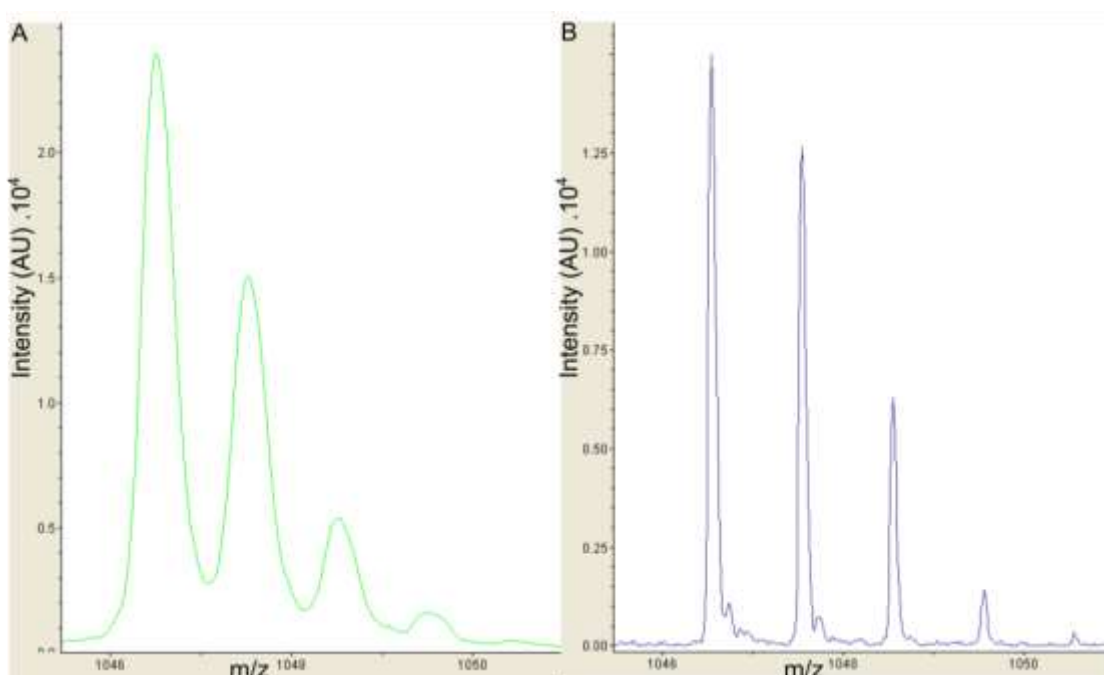
Figure 12: MALDI-TOF mass spectrum of Angiotensin II in linear (A) and reflectron mode (B).

Quadrupole (Q) is just slightly more complicated. It has been developed in 1953 [87] and it consist from four metal rods organized in square. Ion separation is formed by changing current on the rods. Current of the same polarity is inserted always on the diagonal rods. Using frequency based changes of polarity; ions inside the Q are forced to rotation movement. When a rotation frequency of an ion reaches a limit of stability, ion will leave

the Q. This limit is dependent on mass of ion. By changes of rotation frequency, Q can serve as mass analyzer, thus is scanning ion masses, or as mass filter, thus filtering only one mass or mass range through quadrupole. This filtering ability is used in hybrid mass spectrometers and ion optics. Ion optics, which directs the ion flow in hybrid (e. g. orbitrap based) mass spectrometers, is usually based on quadrupoles.

Derivative of quadrupole allowing to not only filter, but also to store ions is called Ion trap [87]. The main differences are two additive rods on the beginning and end of the analyzer. Ions can be thus trapped inside. This is the first analyzer which offers multiple MS. Trapped ions can be hit by gas atoms (collision induced dissociation fragmentation mode, CID) or fluoranthene atoms (electron transfer dissociation fragmentation mode, ETD) leading to ion fragmentation. This can be repeated – one fragment will remain trapped and subjected to further fragmentation. This ease of fragmentation is used with advantage with proteomics, because acquired spectra are obtained fast with good sensitivity. Ion trap can be used either as stand-alone instrument, which is able to identify about 700 of proteins (Appendix B), or as a part of orbitrap based hybrid instrument.

The need of high resolution introduced two mass analyzers based on Fourier transformation (FT). FT is method of transforming signal in time (waves) to frequencies. Mass analyzers based on this principle thus analyze vibrations caused by ions with different masses. This allows achieving of high resolution, however longer analysis times are required for best performance.

The first of commonly used mass analyzers with FT is ion cyclotron resonance introduced in 1974 by Comisarow and Marshall [88]. The principle is recording ion vibrations during their fly in cyclotron. This is an instrument with very powerful resolution, however its main drawbacks is its size and need of cooling superconductive magnet by liquid helium and nitrogen.

The next FT based analyzer is orbitrap. It has been developed in 2000 by Makarov [89]. Ions are analyzed in special orbital ion trap (Figure) where they are separated by rotation and oscillation. Oscillation is then recorded by orbital electrodes and subjected to FT. Resolution of orbitrap is not as high as in ion cyclotron resonance, on the other side orbitrap based mass spectrometer is much less complicated and more compact.

1.2.2.4 Hybrid mass analyzers

Those basic mass analyzers have their limitations, which can be overcome by combining of multiple analyzers in one instrument. Those instruments are then called hybrid instruments. The best known hybrid instrument is triple quadrupole utilizing filtering capability of single quadrupoles. The variant of this instrument is replacing last quadrupole by TOF, which provides better resolution of spectra. The last, unrelated to triple quadrupole and with most complicated architecture, are orbitrap based hybrid instruments. Basic scheme of mentioned instrument is at Figure 13.
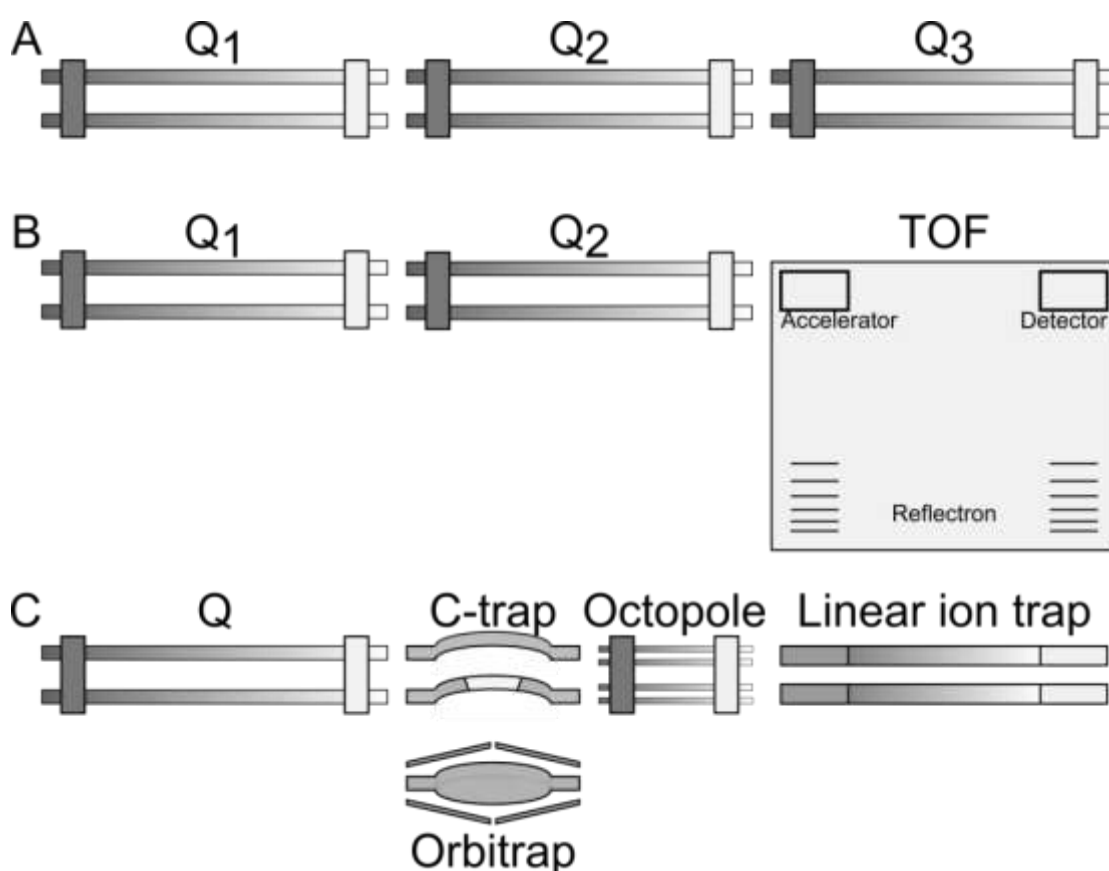
Figure 13: Schema of hybrid mass spectrometers: triple quadrupole (A), quadrupole - TOF (B) and quadrupole - linear ion trap - orbitrap (C). All instruments are coupled to ESI before Q1 (Q). In orbitrap, Q serves as mass filter, C-trap focuses ions for Orbitrap analysis and octopole serves as main crossroad of ions. This crossroad stores and sends ions based on user set protocol.

In triple quadrupole instrument, the first quadrupole can either scan masses of compounds or filter one specific mass. Second quadrupole serves as collision cell, where ions of analyte collide with atoms of gas causing fragmentation of ions. Third quadrupole has the same properties as first one. It scans or filters the ion fragments. The experimental setup, where

first quadrupole is set to filter parent mass of determined compound and third quadrupole to filter only fragment specific to this compound is called selected reaction monitoring (SRM). Precursor and fragment pair is called transition. Monitoring of more fragments from one precursor is then multiple reactions monitoring (MRM). SRM and MRM are more or less synonymous because modern mass spectrometers allow fast and easy monitoring of lot of transitions. SRM/MRM approach is great in selective analysis and quantification of low and middle molecular weight compounds, e. g. in metabolomics or in proteomics. To simplify nomenclature, only MRM abbreviation will be used further in this thesis.

The biggest disadvantage of triple quadrupoles is their low resolution often leading to false positivity of transition. MRM or improving resolution can overcome this issue. The hybrid mass spectrometer with all advantages of triple quadrupole together with high resolution is quadrupole with TOF (Q-TOF). Q-TOF (Figure 11) instrument offers high resolution of TOF mass analyzer without compromising sensitivity or speed of analysis of triple quadrupoles [90]. Q-TOF based instruments are suitable for high resolution MRM (reported as HR-MRM) [91] or sequential window acquisition of all theoretical fragment ion spectra (SWATH) analyses [92]. SWATH analysis is a method of data independent analysis with larger precursor windows (swaths). This method will be further discussed in Label-free techniques chapter.

The hybrid mass spectrometers with most complex architecture are orbitrap based instruments (Figure 11). It usually consists from advanced ion optics, C-trap, orbitrap, linear ion trap and in some instruments quadrupole. This complex architecture is allowing orbitrap to be an ultimate tool for proteomic analysis. Orbitrap allows determining accurate mass of analytes. This process in not fastest, so meanwhile linear ion trap can analyze fragmentation spectra of selected ions. Both information together allows increasing protein identification rate [93]. For example, in cell lysate analyzed in our laboratory by LC-MALDI on Bruker UltrafleXtreme instrument was identified 1500 proteins (unpublished results). The same lysate analyzed on Thermo Velos Pro orbitrap instrument provided 3400 identified proteins (Appendix B).

Combination of multiple mass analyzers allows adding new functions to mass spectrometer or performing more analyses at once. Using this approach, it is possible to build an ultimate instrument for proteomic analysis.

1.2.3 Protein sequencing methods

In typical MS experiment, proteins are identified by probability score based on number of matched peptides. Peptides are searched by matching mass spectra with peptide sequence in database. Those databases have various sizes depending on current knowledge and scientific interest about particular organism. This can be illustrated on Uniprot database, which has more than 1 million entries related to human, but only 262 entries for badger (*Meles meles*). Those numbers are valid to July 2016. Protein sequencing is the approach, which actually fills those databases. Sequencing of proteins can be done by three basic approaches. First one is translating of genome sequence [94]. This is done by reverse transcription of mRNA to cDNA followed by DNA sequencing. Since human genome project is already finished [95], all hypothetical sequences are translated into database and thus they allow faster way of adding knowledge about such proteins via proteomics. If no genome sequence is available, there are still two "wet lab" approaches for determining protein sequence – Edman degradation and MS *de novo* sequencing.

1.2.3.1 Edman degradation

The most often used method of peptide sequencing was published in 1950 by Edman [96]. This method allows sequencing peptides from N-terminus. The principle of method is bond of phenyl isothiocyanate to N-terminal amino acid. Phenylthiolcarbamoyl derivative of amino acid is then cleaved under acidic conditions and separated by HPLC or thin-layer chromatography. Then next amino acid can be processed the same way. This approach can be easily automated, however reasonable sequencing length is about 30 amino acids [97]. Digestion of proteins, routinely used in shotgun proteomics, was originally invented to overcome this limitation of Edman degradation [98]. The main advantage of Edman degradation is that it provides protein sequence independently from databases with almost no errors however with lower throughput. The biggest disadvantage of Edman sequencing is worse sensitivity compared to MS and higher requirements to sample amount. Edman sequencing is considered mainly as historic approach, but as shown on Figure 14, it is still used.

1.2.3.2 Mass spectrometry sequencing

Mass spectrometry replaced Edman sequencing as main proteomic detection method mainly because its higher throughput and sensitivity [98]. MS basically offers multiple analytic modes. The vast majority of those modes are based on comparison of detected masses with databases. Identification analytic mode which is not dependent on good

database is called de novo sequencing [99]. Its principle is to decipher peptide sequence directly from MS/MS spectra without comparison to database. This approach is beneficial for analysis of poorly characterized organisms [100], newly raised antibodies [101] or in study of peptide toxins [102]. Thanks to recent advantages in result processing, this approach can be also used in high throughput manner as an alternative to classic database search [99]. De novo sequencing is used for current MS sequencing tasks and is under constant development. As shown from Figure 14, de novo sequencing is method with raising interest.
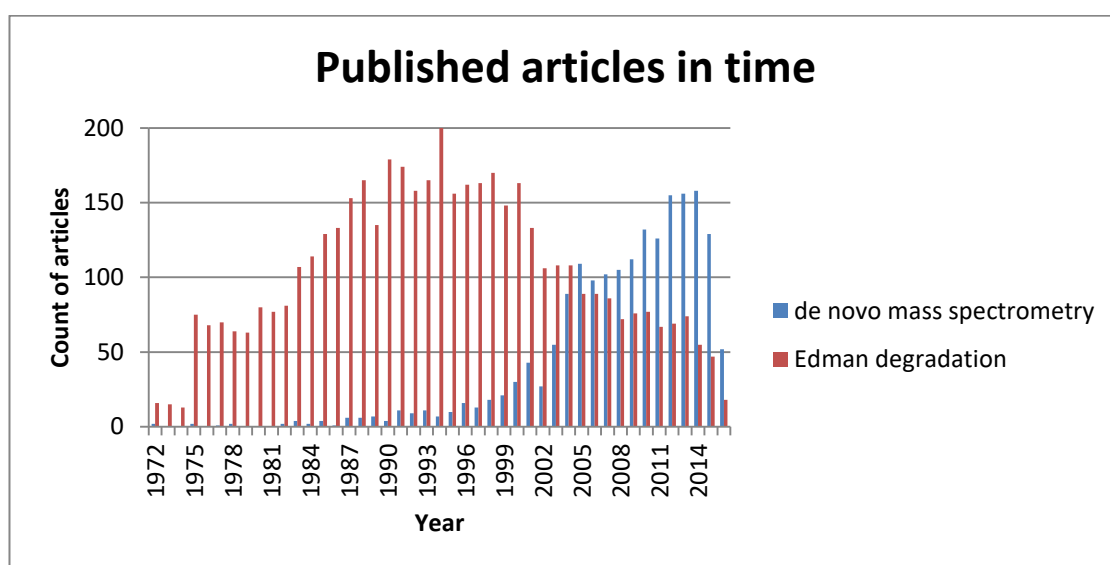
Figure 14: The count of articles in PubMed related to "Edman degradation" and "de novo mass spectrometry" in timescale. Data are relevant to 10. 4. 2016

1.3 Top-down and bottom-up proteomics

There are two basic approaches in large scale proteomic study. Those approaches are called top-down or bottom-up based on direction, in which proteins are studied. Top-down approach is based on separation and analysis of intact proteins. MS analysis of intact proteins allows to precisely determining protein sequence of smaller proteins (less than 50 kDa) with all its PTMs and isoforms (Figure 15). This is achieved by combination of electron capture / transfer dissociation (ECD/ETD) fragmentation mode and high resolution mass spectrometer. High resolution allows having isotopic resolution even for bigger proteins. ECD or ETD is crucial fragmentation method because both modes fragments specifically at peptide bond and they preserve more fragile PTMs. This allows locating position of concrete PTM. However there are some challenges in comparison with more popular

bottom up approach. The main ones are poor protein solubility, insufficient sensitivity of MS in comparison to bottom up and algorithms for MS result processing [103].

The second approach is called bottom-up or shotgun proteomics. Correspondingly with its name, proteins are digested into smaller peptides, which are analyzed (Figure 13). However identified peptides covers complete protein sequence only rarely. Incomplete sequence coverage then resembles quasi-random firing pattern of shotgun. Although digestion of proteins to peptides is favorable for MS analysis, it raises complexity of sample (proteins are digested to multiple peptides). This makes a good separation prior analysis even more necessary. Another challenge is the matching of spectra to peptide and peptide to protein, which is not a trivial task. With respect to all its drawbacks, bottom up approach became the most widely used approach for proteomic MS analysis [104]. In the following paragraphs, single steps crucial in bottom-up proteomics will be introduced.
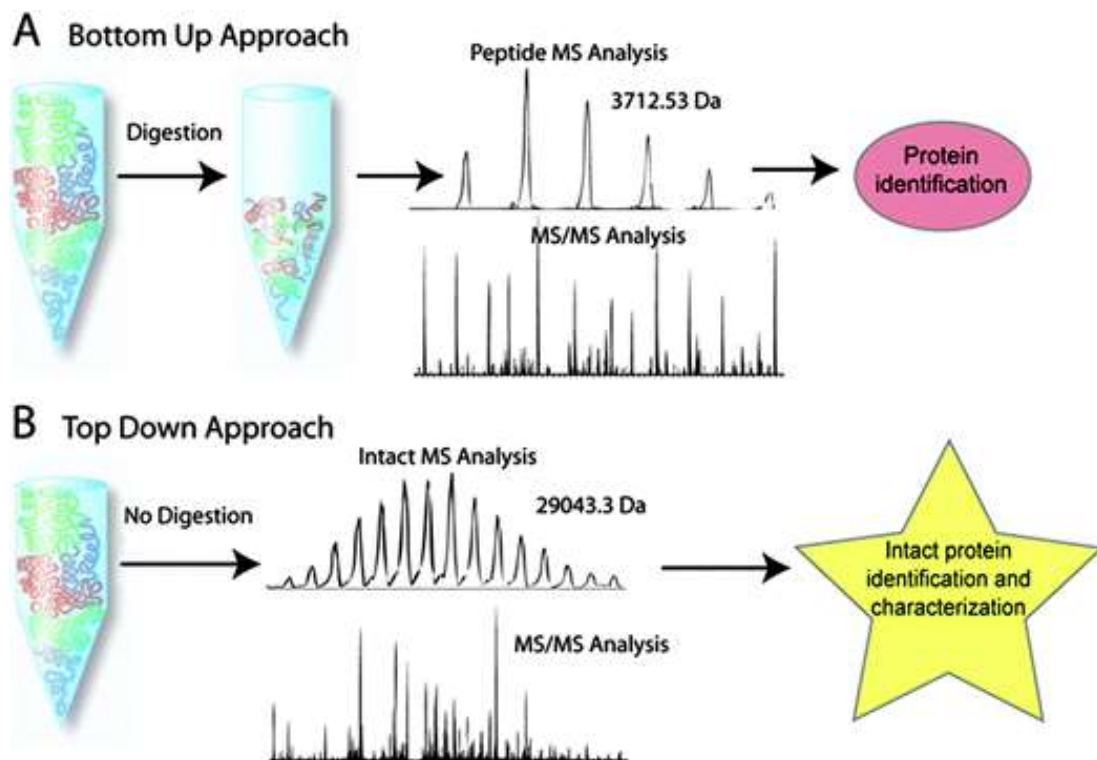
Figure 15: Main proteomic approaches: Bottom up (A) and Top down (B) approaches for MS analysis of proteins. Figure adapted from [105].

1.3.1 Sample preparation and digestion in bottom-up proteomics

Samples are prepared for bottom-up proteomics various ways depending on particular samples. Usual procedure can be briefly described as protein extraction, protein enrichment and digestion. Protein extraction depends on the nature of sample (cells,

39

serum, fresh or formalin fixed paraffin embedded tissue, body fluids etc.). Protein enrichment is optional and is usually performed by protein separation methods discussed in the Chapter 1.1. This step serves for reducing of sample complexity.

1.3.1.1 Sample preparation

Prior to key step of bottom-up approach, protein digestion, there are two common steps in sample preparation: a protein reduction of disulfide bonds and alkylation of free cysteines. This prepares proteins prior digestion and prevents occurring of bounded dipeptides. All mentioned steps are optional and serves for better preparation of sample. Experimental setups for performing digestion are various and dependent on previous processing of sample, for example an in-gel digestion for processing from SDS-PAGE gels or in-solution digestion for free protein solutions. Whereas in-solution digestion is just adding of reagents into solution, in-gel approach takes advantage of one important gel property - proteins are encapsulated in the gel and thus can be washed several times without sample loss. Proteolytic enzyme can later enter the gel freely and digested peptides are no longer retained in the gel and can be easily extracted. These methods suffers from several drawbacks like a difficult automation of in-gel digestion or poor efficiency of detergent removal in in-solution digestion [106].

There is a big effort to overcome such drawbacks together with constant need of improving the sensitivity of proteomic assay. This effort resulted into introduction so called proteomic reactors, which can be in its simplest form a spin filter device [106], reactor built in pipette tip [107], coated magnetic beads [108] or complex microfluidic devices [109]. From these protein reactors, mainly spin filter derived filter aided sample preparation (FASP; Figure 16) protocol [106] gained a wide popularity due to its simplicity, high throughput and acceptable efficiency. FASP basically combines benefits of in-solution and in-gel digestion. Sample is in the beginning in the solution, which is loaded into spin filter. This filter serves as a carrier device and allows sample washing and detergent removal. In the last step, proteolytic enzyme is applied. After digestion, peptides can go through the filter and undigested macromolecular complexes are retained. FASP is two times more efficient than simple in-solution digestion [110].
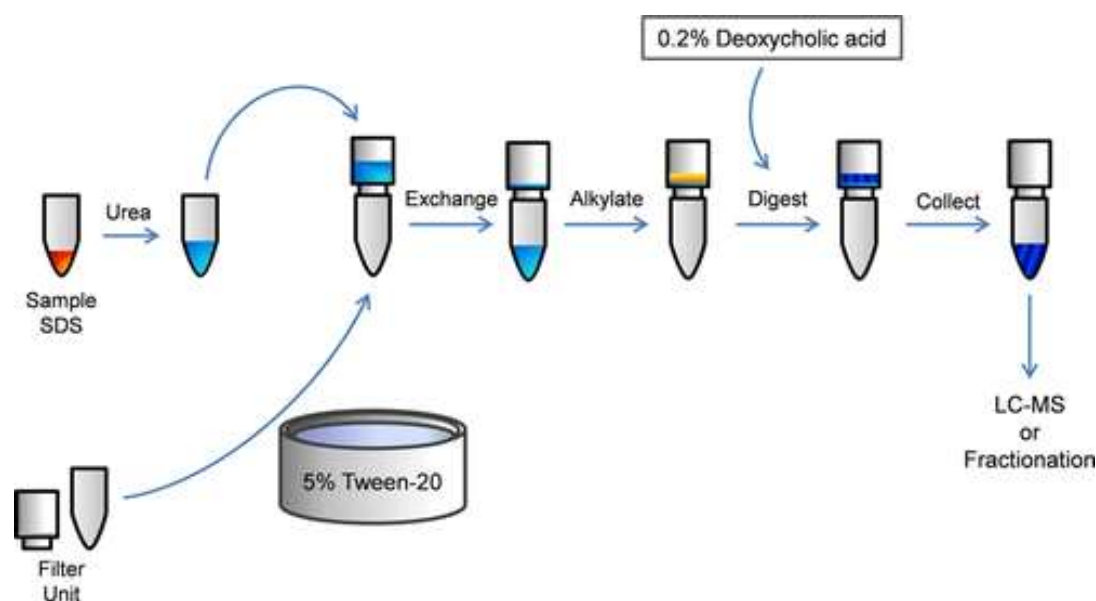
Figure 16: Schema of filter-aided sample preparation (FASP) protocol. Figure adapted from [111].

1.3.1.2 Digestion

The step, which is common in all bottom-up approaches, is digestion by proteolytic enzyme. Although there are several enzymes possible to use [112], only one enzyme is used in most proteomic experiments. This enzyme is trypsin. Beyond its wide popularity is the fact that trypsin is robust, cheap, and specific. Trypsin also generates peptides in a mass range of 500 to 3,000 Da, which is optimal for chromatographic separation and yields peptides that ionize and fragment well due to the presence of a C-terminal lysine or arginine residue that efficiently protonates under acidic conditions [104]. This is beneficial for MS detection of such peptide. Relying on trypsin in bottom-up proteomic experiment has as well its drawbacks. The main one is that majority of tryptic peptides (56% of all generated peptides are ≤ 6 residues) are too small and thus generally not identified by MS, meaning that only a restricted segment of the proteome is covered. Consequently, this limited sequence information makes it often impossible to distinguish between protein isoforms and also to identify all PTMs decorating the proteins [113].

There are several possibilities how to improve tryspin digestion. The first one is to use alternative protease. However, proteases currently used had their own drawbacks reviewed in [113], where the most serious is inability to compare quantitative data between different proteases. Currently the best approach is a combination of multiple proteases. The most often is digestion with lysyl endopeptidase (Lys-C) from

Lysobacter enzymogenesis prior to trypsin digestion [110]. Using of Lys-C, which is specific only to lysine greatly improves trypsin efficiency [110].

The last but not least step in proteomic sample preparation is a desalting of digested peptides. This step is done, because mass spectrometry in general is very prone to salts and polar contaminants of the sample. Desalting step is thus based on separation of salts and polar compounds from less polar peptides. As has been discussed before, a reverse phase resin – mostly C18 – is used for this step almost exclusively. For peptide desalting, there are some common experimental setups, like a sample clean-up before LC-MS or using short pre-column in LC method. Both these approaches have been employed in this thesis (Appendices B and C).

1.3.2 LC-MS analysis

LC-MS analysis is crucial for both top-down and bottom-up proteomic methods. Because LC part of the proteomic analysis has been already discussed in chapter 1.1.2 and principle of MS in chapter 1.2.2, this chapter will focus only to LC-MS acquisition methods specific in proteomics.

In top-down proteomics, there is main approach consisting from already described high resolution MS[1] followed by ECD/ETD fragmentation. In bottom-up, there are three main mass spectrometric approaches – data dependent analysis (DDA), MRM and data independent analysis (DIA) [104].

DDA is the most often used MS approach in proteomics. It consists of two successive steps – first is determination of full MS[1] spectrum. This spectrum is processed and peaks are picked and selected for fragmentation by performing MS[2]. Two kinds of information obtained – the accurate mass of precursor peptide and its fragmentation spectra revealing peptide's sequence – are used separately in succeeding sample processing. Fragmentation spectrum is usually utilized for peptide identification by database or spectral library match (see chapter 1.3.4.2) or by de novo sequencing (chapter 1.2.3.2), but they could serve for quantification in some label experiments discussed in chapter 1.3.3.1 as well. If MS[2] spectra serve mainly for identification and less for quantification, the situation is reverted for parent ion spectra. Quantification of peptides is in most cases calculated from peptide precursor ion and their extracted ion chromatograms. Precursor spectra can be also used for identification by accurate mass and time tag, which is matching mass and retention time to known peptides. The main issue of DDA is undersampling phenomena – this means that

peptides for fragmentation are selected semi-randomly, usually as certain count of most intense ions in the spectrum. The result of undersampling is decreased reproducibility of replicates. Modern MS instruments have undersampling issue reduced, but it is not still eradicated [104] (Appendix B).

MRM principle has been already discussed in chapter 1.2.2.4. It is method of choice for precise quantification of relatively small amount of peptides in multiple samples. Amount of analyzed peptides can be raised by introducing scheduled MRM, what means performing scan of particular transition only in certain retention time. The major drawback of MRM in proteomics is difficult optimization, especially for bigger amounts of transitions [104].

The last approach combining advantages of previous two ones is DIA. DIA is based on fragmentation of wide mass windows making mixed spectra of all precursor and product ions presents in particular mass window. This prevents undersampling and allows extraction of peptide transition ex-post. Peptide search is more demanding than in case of DDA. For example in SWATH analysis (discussed later as quantification approach) is recommended to search peptides against spectral library which is prepared by DDA analysis on the same instrument [104]. Principle of DDA and DIA and their comparison are shown at Figure 17.
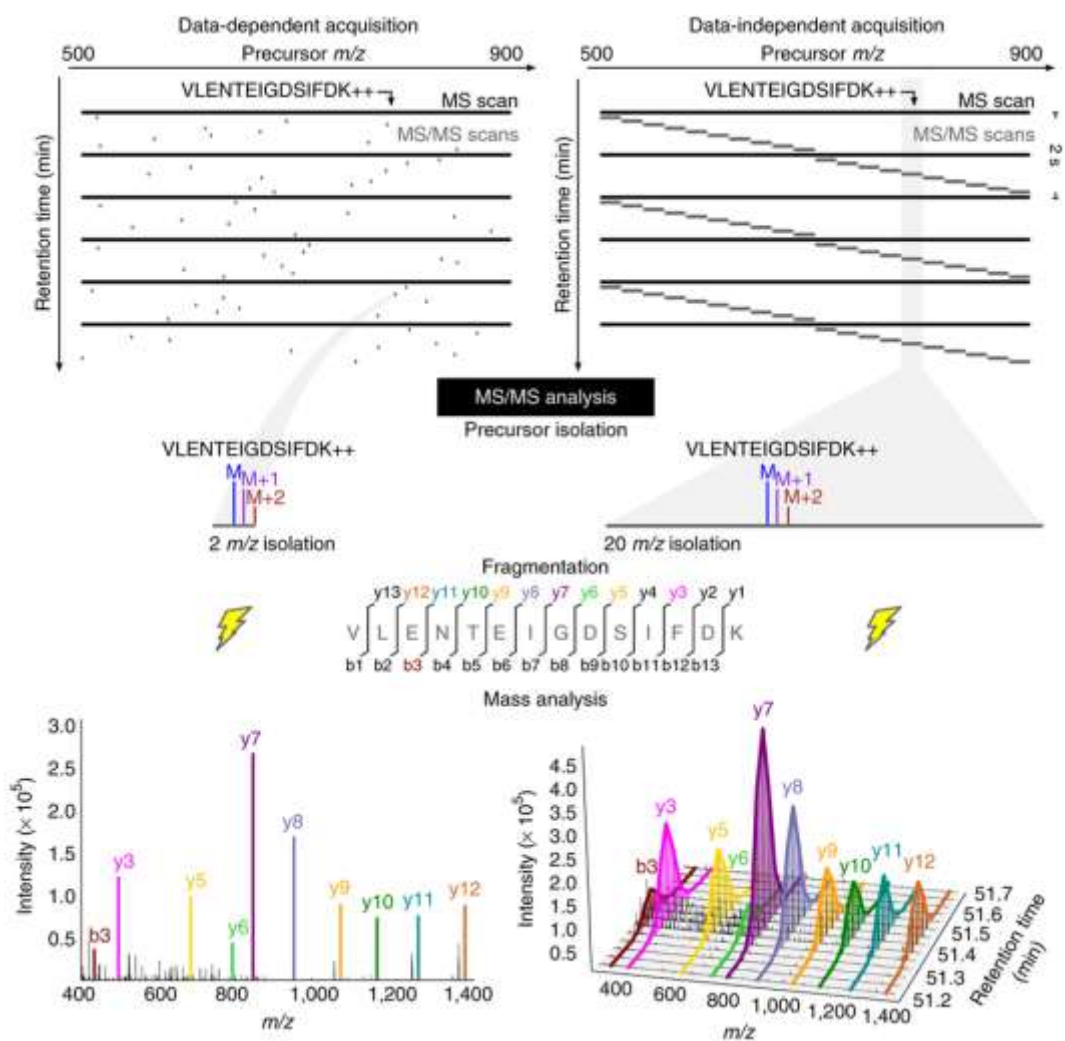
Figure 17: The comparison of data dependent and data independent mass acquisition. Figure adapted from [114]

1.3.3 MS based quantification

Peak area in mass spectrometry is not as straightforwardly dependent on quantity as in HPLC, but also on ionizability of analyzed compound. MS peak can be thus quantified only by comparing to the peak of same compound. There are two basic ways how to quantify in MS. The first is adding an isotope labelled standard, or in case of proteomics, isotope labelled peptide and the second way is a method of calibration curve. Isotopic label can be introduced into sample by several methods discussed below. However sometimes is introducing of isotope label not possible or very difficult. For the calibration curve method, it is absolutely necessary to keep ionizability factor constant. This is possible only with electrospray ionization with stable settings [115]. Those techniques are commonly discussed as label-free quantification. The basic principle of different isotope labelling

methods and their comparison to label-free approach are shown at Figure 18. Both isotope labelling and label-free approaches have been reviewed in Rylová et al. (Appendix A).
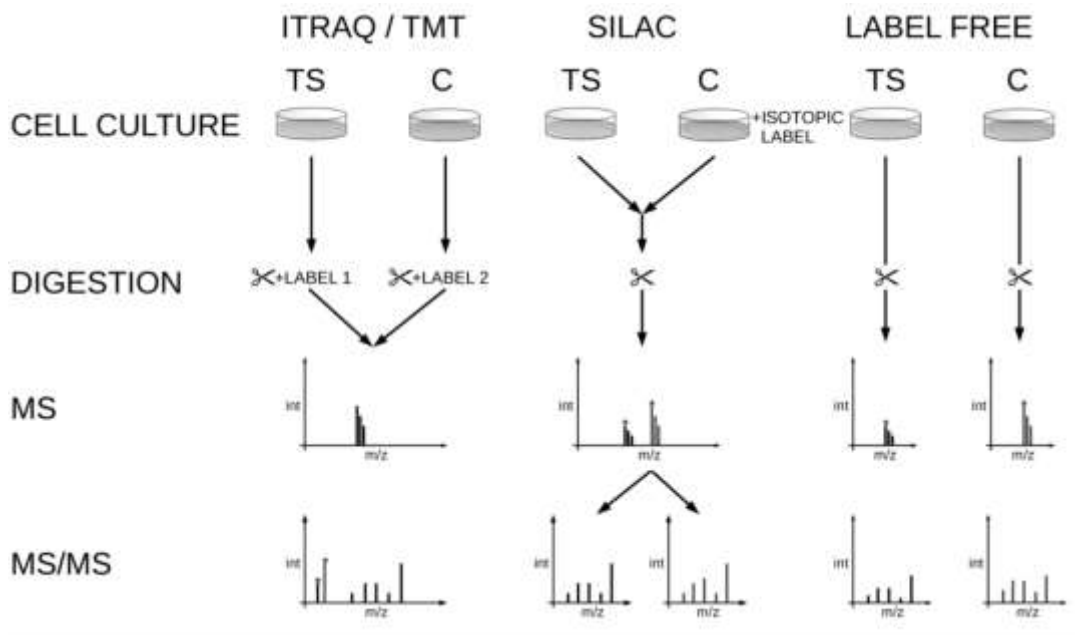
Figure 18: Comparison of different quantification approaches in LC-MS based proteomics. ITRAQ/TMT represents isobaric tagging by chemical modification of the samples; SILAC metabolic labelling and LABEL FREE the quantification approach without any label. Basic steps of proteomic analysis are marked at the figure with marking of the step, where introduction of label occurs. Asterisk marks the stage of quantification event.

1.3.3.1 Isotope labelling methods

There are several quantification protocols differing in the way of introducing isotope tag. The most used methods are currently labelling with isobaric tags or metabolic labelling. Those approaches are used in proteomic experiments focused e. g. to the discovery of biomarkers [116] or molecular targets of possible and established anti-cancer drugs [117].

1.3.3.2 Isobaric labelling

The first category of isotope labelling methods are tandem mass tags (TMT; [118]) and isobaric tags for relative and absolute quantitation (iTRAQ; [119]), which are based on chemical binding of isobaric tag to the N-terminal end of a peptide (Figure 18). The isobaric tag consists from three parts – a reactive part binding to peptide, a reporter ion being released during peptide fragmentation and balance group which balances differences between reporter ions. Isobaric tag based methods allow easy multiplexing with possibility to quantify up to eight samples in case of iTRAQ [120] or ten samples in case of TMT [121].

This multiplexing is in contrast to other labelling methods, which allow quantifying only two or three samples, and is one of main benefits of TMT and iTRAQ. Another advantages are clean spectrum (all labels have the same mass in MS[1]) and possibility to label wide variety of samples including tissues. Disadvantages are risk of co-isolation of more precursors in MS[2] and need of MS[3] in high-plexing experiments. Both TMT and iTRAQ are sold commercially, but there is a cheaper alternative, DiLeu isobaric tagging as well [122]. This approach is freely available with only main drawback – user has to synthetize tags in laboratory. DiLeu works on the same principle as TMT or iTRAQ and allows up 8-plexing [122].

1.3.3.3 Metabolic labelling

Stable isotope labelled amino acid in cell culture (SILAC) is very versatile and frequently used quantification method [123]. This approach is based on culturing cells in media supplemented with isotopically labelled amino acids, typically arginine and leucine. Those two amino acids weren't selected randomly – trypsin cleaves specifically after these amino acids. Regular proteomic sample processing will thus result into peptides bearing one, only occasionally two labelled amino acids. Cultivation of cell cultures with such modified media will result in incorporation of those amino acids into cellular proteins in around five passages, depending on cell type. Those labelled cells are later mixed with its unlabeled counterpart and they are processed together (Figure 18). This approach allows limited multiplexing by introducing different labels into cells. The resulting spectrum is thus more complicated, since every peptide has two or more peaks in MS[1]. Quantification is based on extracted ion chromatogram for each labeled peptide. The major benefit is straightforwardness of labelling and result processing. As will be shown in experimental part, it allows determining drug target response on particular cell line (Chapter 2.2). The drawback of this method is suitability for cell cultures only. There is a stable isotope labelling processes for whole organisms, e. g. mouse, however they are much more complicated [124]. It is not possible to analyze tissue samples by SILAC.

1.3.3.4 Other isotope labelling

 Apart from these main quantification approaches, few other labelling techniques exist. Those methods are: protein digestion in water with heavy oxygen ^{18}O [125], metabolic labelling of cells using heavy nitrogen ^{15}N [126], isotope coded affinity tag (ICAT) [127], isotope coded protein labelling (ICPL) [128], global internal standard technology (GIST) [129]. Labelling with ^{18}O suffers from several disadvantages like incomplete incorporation

of labelled oxygen into digested peptides and not possible multiplexing. However, it is still one of the cheapest and simplest methods for protein quantification in MS [130]. Metabolic labelling with ^{15}N is currently used mainly in the research of simpler organisms such as plants, invertebrates and unicellular organisms [131]. ICAT, ICPL and GIST are based on introducing an isotope labelled tag on peptides. In ICAT, tag is bound on a selected rare amino acid, usually cysteine and affinity purification of only cysteine containing peptides. ICPL tag containing H_4 or D_4 is bound to amine residue of lysine of reduced and alkylated protein or proteins in the complex sample, which allows further manipulation with these proteins. GIST tag labels universally to the primary amino group of all peptides. These methods are not very widely used and basically they offer virtually the same results as SILAC, TMT or iTRAQ.

1.3.3.5 Label-free methods

In some cases, it is difficult or impossible to introduce isotopic label. It is then necessary to introduce quantification approach, which is not isotopic label dependent. This approach depends on quantitative property of ESI. This property is currently the principle of quantitative MS. In analysis of low molecular compounds, MRM is often synonymous with quantitative MS. In discovery proteomics, however, MRM is not often applicable and thus different methods for MS quantification without isotope label had to be developed.

There are two main methods for label-free proteomic MS quantification. First one is similar to common MS quantification approach. It consists of counting peptide's peak area and directly comparing this area to peak area of the same peptide in different run. This approach is called area under curve (AUC) [132]. Second method is based on presumption, that more abundant protein will be identified with more peptides. This method is referred as spectral counting [133]. The main advantages of AUC are linearity in broad concentration range (10 fmol–1000 pmol) and straightforward processing, main drawbacks are high request for LC/MS reproducibility and accuracy. On the other side, accuracy of quantification may be decreased by peptide co-elution as well as by losing peak information during switching between MS and MS^2 scan modes. High resolution mass spectrometers with data independent acquisition mode are recommended for obtaining best results. Spectral counting is less demanding to computing power and can be calculated as protein abundance index during each Mascot search. There are as well a lot of modifications of those two main methods reviewed in [134].

47

Nor AUC or spectral counting is hardly usable in DIA. In DIA, a SWATH acquisition method has been introduced. SWATH is based on fragmenting peptides in larger windows called swaths (Figure 19) [92]. Those swaths have from 10 to 25 Da and resulting spectrum is chimeric MS^2 spectrum of all peptides in particular swath. Quantification lies in between label free and MRM analysis. A workflow for SWATH analysis consists from one data dependent analysis serving for construction of peptide spectral library and several SWATH runs. Particular peptide of interest is then selected as parent mass and mass of its fragments from identification run and quantification information is extracted from SWATH runs and visualized in similar way to MRM transition. The main advantage amount of peptides quantified this way which can be almost unlimited.
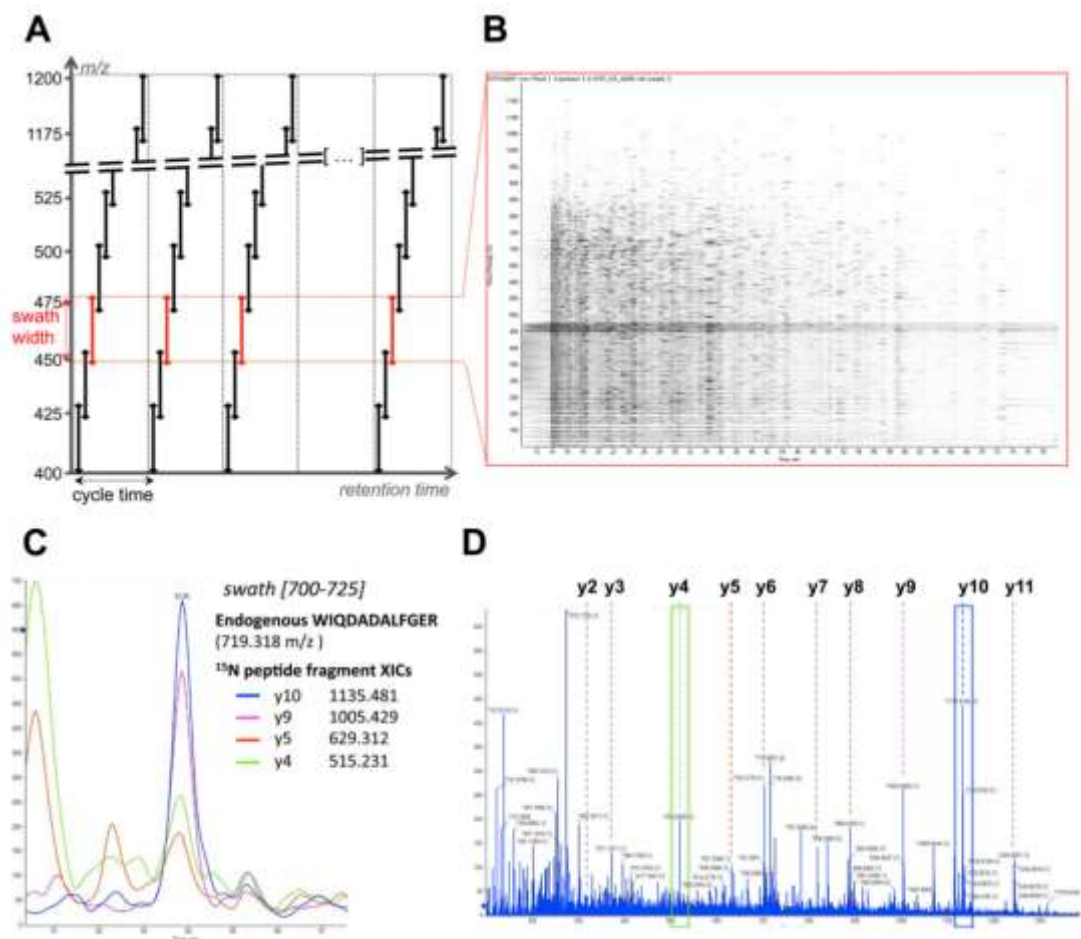
Figure 19: The principle of SWATH data independent acquisition. Mass spectrometer is set to acquire MS/MS spectra in fixed, relatively wide mass windows (A) which leads to cover all mass range (B). Single peptides are evaluated in MRM-similar manner (C) based on single chimeric fragmentation spectra (D). Figure adapted from [135].

1.3.4 Data processing and database search

The most important information in both top-down and bottom-up approach are MS/MS spectra. Those spectra contain sequence information about particular peptides. This information is however coded in mass peaks of single amino acids losses. Peptide sequence has to be deciphered from mass spectra first. There are several approaches for this deciphering, like de-novo sequencing or search in databases or spectral libraries.

1.3.4.1 Database based search

MS/MS spectra in bottom-up proteomics can be evaluated several ways. The most often used approach is database search using specialized algorithm. There are databases available online – the most known are UniProt (http://www.uniprot.org/) or NCBI (https://www.ncbi.nlm.nih.gov/protein/) – but it as well possible to get specialized databases. The most of searching algorithms works with regular FASTA sequence format and both can be done – either download database of choice from internet or to make own. Database search by different algorithms usually works on simple principle – in the first step, mass spectra are deconvoluted – MS^1 and MS^2 peaks are picked and transferred into numeral form. Proteins in database are then digested in-silico and results from MS are compared to this in-silico digest. Peptide, which corresponds with spectrum with highest probability, is chosen as hit. The exact mechanism of database search differ by particular algorithms, typical examples are e.g. Mascot [136], Sequest [137] or Andromeda (part of MaxQuant package [138]).

1.3.4.2 Spectral libraries

Although database search is well established method, it takes a lot of computational power to search. The main drawback of database based approach is repeated search of abundant proteins in multiple samples. To overcome this, spectral libraries, originally used in analysis of organic compounds, were adapted to proteomics. The main advantage of spectral library search is increased speed of analysis. The spectral libraries are instrument specific and have to be filled in similar manner as protein databases. The main difference is that we can fill regular database with DNA transcripts, but spectral library can contain mass spectra corresponding to peptides only. For example, human ion trap library contains 340 357 spectra, whereas chicken database contains just 3125 spectra [139].

1.3.4.3 De-novo sequencing and accurate mass tag search

Besides those two main approaches, there are few other methods for bottom-up proteomics. The first can be de novo sequencing of every spectrum. This approach is computationally even more demanding than database search [139], but it is under constant development [99]. Another can be accurate mass and time tag [104], which is comparison of retention time and accurate mass to already known peptide. Compared to previous approaches, this one is the less sophisticated.

Data analysis for top-down proteomics has still reserves for further development. Currently, there are software tools for protein mass deconvolution from multiply charged peaks. First tools for protein sequence determination from MS^2 spectra have been already introduced even for analysis of more complex samples [103].

1.3.5 Bioinformatic evaluation of results

All proteomic techniques aim to the discovery of a certain number of differentially expressed proteins. Unfortunately, those proteins don't have big information value alone, because proteins are involved in particular cellular pathways, work as a part of multiprotein complex, are regulated or regulate another proteins etc. Similarly to protein IDs or gene symbols, cellular processes have been characterized and referenced in two main databases – Gene ontology (GO) and Kyoto encyclopedia of genes and genomes (KEGG). KEGG serves as a virtual model of the cell and the main purpose of GO is to serve as the unifying vocabulary. Thus the same process and the orthologous protein have the same GO annotation between different species [140,141]. Each protein has multiple GO annotation terms. This count depends on importance of protein in cellular structure and current knowledge about such protein. Such GO annotation terms can be downloaded to each protein by simple tools. One of those tools is Annotation in MaxQuant. However, GO annotation itself doesn't provide much information, because proteins are usually involved in multiple pathways. For such a purpose, more sophisticated algorithms to annotate and visualize genomic, transcriptomic or proteomic data were developed. Commercially available are the Thomson Reuters MetaCore™ program (GeneGo; Figure 20), Elsevier Pathway Studio and IPA (Ingenuity® Systems). Further, several free programs were introduced to help researchers gain more information about identified proteins. To this class of analytical tools belongs for example: the Software tool for researching annotations of proteins (STRAP, [142]), Gene Map Annotator and Pathway Profiller (GeneMAPP, [143]), Pathway Tools [144], on-line tools available from Gene Ontology Web site

(www.geneontology.org, [145]), ProteomeCommons (proteomecommons.org/tools-browse.jsp, [146]) or Database for Annotation, Visualization and Integrated Discovery (DAVID, https://david.ncifcrf.gov/ [147]).
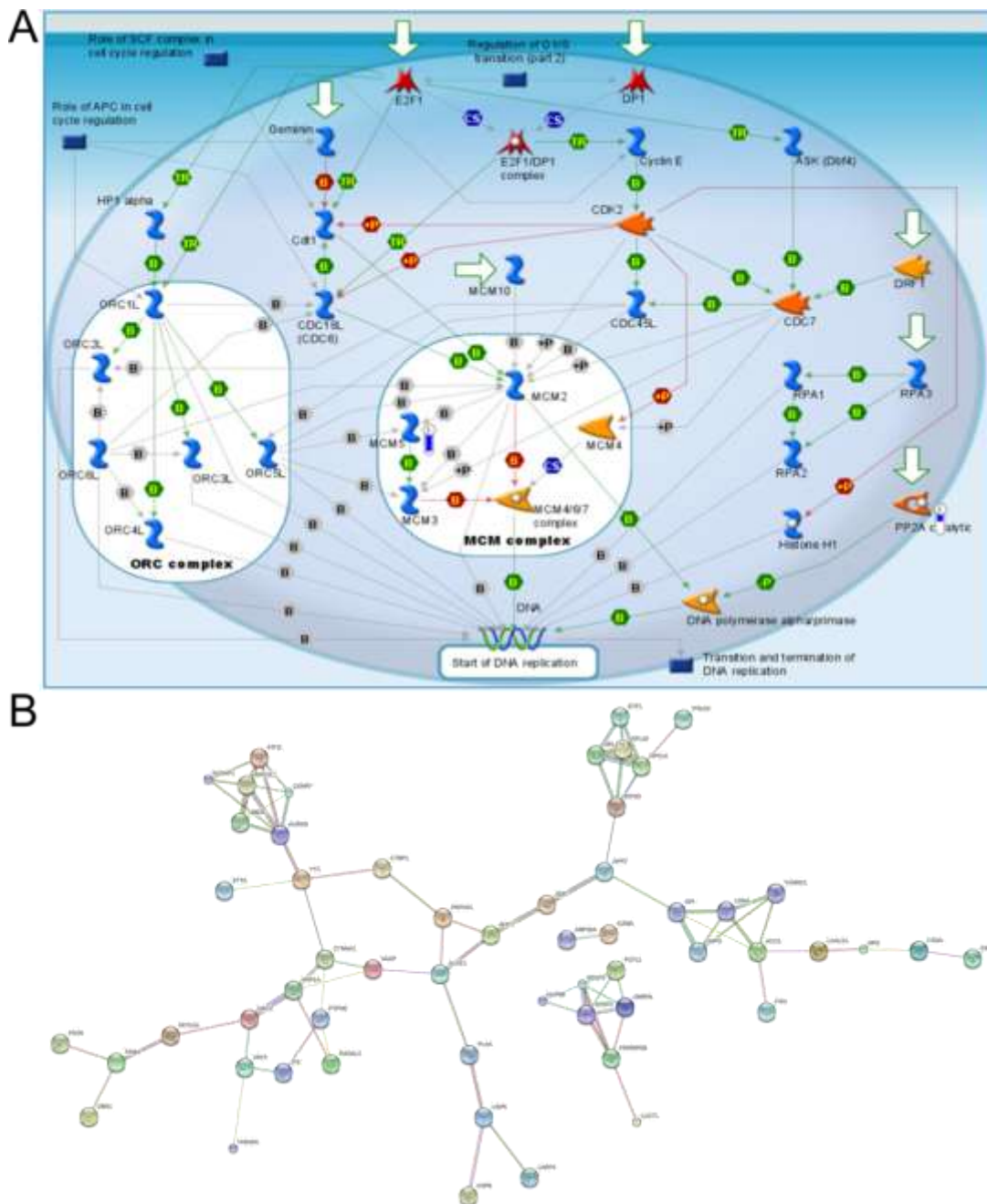
Figure 20: Graphical output from GeneGo annotation (A) and STRING protein-protein interaction (B) analysis.

Protein-protein interactions are aside from cellular processes annotated in GO. These interactions are much wider than participation of protein in single process. Protein-protein interactions are very important part of cellular communication, signal transfer or

modulation of enzyme activity. A good visualization of protein-protein interaction can be a good guide in determining mechanism of cellular response, as can be seen in Chapter 2.2.2.3 in Experimental part. The tools used for visualization of protein – protein interactions are STRING (http://string-db.org/; Figure 20 [148]), BioGrid (http://thebiogrid.org/ [149]), Agile Protein Interaction DataAnalyzer (APID, http://bioinfow.dep.usal.es/apid/index.htm [150]) or human protein–protein interaction prediction database (http://www.compbio.dundee.ac.uk/www-pips/ [151]).

1.4 Other MS-based proteomic methods

1.4.1 Targeted MS methods

Targeted MS methods overcome limitations of both western blot and classical identification approach. They are focused to study particular proteins; however, analysis could be set to quantify a number of proteins, which are no longer suitable to analyze by WB both economically and in terms of time. Targeted MS methods beat identification approach in sensitivity, which are few orders of magnitude higher. There are few basic approaches: classical MRM done at triple quadrupoles, and two variants called high resolution MRM (HR-MRM) and parallel reaction monitoring (PRM). HR-MRM is done at Q-TOF instruments and PRM at Q-Orbitrap instruments. HR-MRM and PRM are similar in a way that both are performed at high resolution instruments and both filter only precursor ions. Instead of filtering product ions, they record full product spectrum and filtering is done during data processing. Targeted MS approach is as well suitable for analysis of post-translational modifications [152]. This approach can be used e.g. for biomarker validation [153], phosphoproteomic assays [154] or as validation of protein expression change in regular proteomic experiments instead of WB [155]. Targeted MS can be for popularization purpose called mass ELISA relating to quantification power, specificity and high-throughput of ELISA experiments.

1.4.2 Non-covalent mass spectrometry

MS can be also used in structural proteomics. Although it lacks the sensitivity and accuracy of NMR and X-ray crystallography, it is much faster and provides sufficient information on protein conformation and ligand binding. There are several methods for studying drug-protein and drug-DNA interactions. Analysis of intact proteins could be performed using both ESI and MALDI ion sources. ESI, and more beneficially nano-ESI in combination with single quadrupole, triple quadrupole or Q-TOF, analyzers are used in the analysis of

biological macromolecules more often. The main reason for using nESI is that it produces a lot of multiply charged ions with high accuracy rather than a few ions produced by MALDI – usually single, double and sometimes even higher charged ions with low resolution caused by the linear TOF analyser. Change in intensity of different charged states also allow to determine changes in conformation [156]. Macromolecule shape determination is elucidated as charge profile of multiply charged protein. This shape determination is allowed because in the duration of a typical MS experiment, a protein is held in a vacuum by hydrogen bonds in its natural conformation. Fragmentation of multimeric protein complexes is also a non-covalent mass spectrometry tool. The methodical part of non-covalent MS is reviewed in more detail in [156]. Non-covalent MS can be used to reveal the protein drug target [157] or to study drug-DNA binding [158].

1.4.3 Surface enhanced laser desorption/ionisation

Surface enhanced laser desorption/ionisation (SELDI) is a variant of MALDI-TOF which is used for differential proteomics [159]. The SELDI is designed to affinity enrich proteins of interest directly on the surface of target and no previous separation is needed. This method also allows construction of protein chips with direct mass spectrometry analysis. The chips are currently commercially available as a Bio-Rad ProteinChip Technology™. However, even with SELDI approach advantages, like on-chip sample preparation and direct mass analysis, whole proteins without digestion are measured in linear time-of-flight analyzer with low accuracy. The SELDI-TOF approach is used preferentially in medicinal biomarker discovery and diagnostics in preference than in drug discovery. SELDI approach has only limited use alone, because it provides only protein mass, but not identification. Protein identification has to be determined by different approaches, for example by using ESI-FTICR to identify differentially expressed proteins [160].

1.5 Clinical proteomics – presence, future and perspectives

The proteomics is currently one of the main scientific methods; however, it just starts getting on clinical use. The oldest and most often clinically used methodology is IHC and western blot used mainly in molecular pathology and microbiology diagnostic. Another widely used method is ELISA or RIA, which is frequently provided as a commercial kit allowing precise and reproducible protein/antigen determination and quantification with low cost instrumentation. In comparison, mass spectrometry based proteomics is not widely disseminated method in clinical proteomics yet from several reasons, which are: expensive instrumentation, complex sample preparation, qualification of operators and

difficult quantification. On the other side, MS offers substantial advantages as a high sensitivity and high specificity in comparison to antibody identification. The MS approach, which offers the best applicability for clinical analysis is an MRM. This approach is targeted (see Chapter 1.4.1) and can be compared with ELISA in terms of sensitivity and specificity. Moreover, MRM analysis can be easily multiplexed for detection of multiple proteins in one analysis, which is hard to obtain in immunoassays.

Targeted MRM approach has been employed in our laboratory to determine hepcidin levels [161] and is currently used as a part of diagnostic panel used for paediatric patients with iron homeostasis dysregulations. Another possible uses of MRM could be in detection of pancreatic cancer biomarker [162] or in prediction of Tamoxifen resistance [163].

However, not only MRM based MS approaches could be used for clinical diagnostics. Regular bottom-up approach can be used as well. Its main benefit is ability to identify and (at least semi-) quantify those unknown proteins. A practical application of this approach could be typization of amyloid tissues [164], also currently used on our institute [165].

The role of MS methods in clinical diagnostics is continuously growing. It is routinely used in metabolomic screening of newborns [166], microbiological identifications [167] or toxicological screening. From those mentioned, microbiological identification on MALDI Biotyper from Bruker is based on our current knowledge a first FDA approved proteomic MS method (Figure 21) [168]. MALDI Biotyper is based on acquiring of unique fingerprint spectra from highly abundant proteins in the bacteria. Those applications show good applicability in clinical practice.
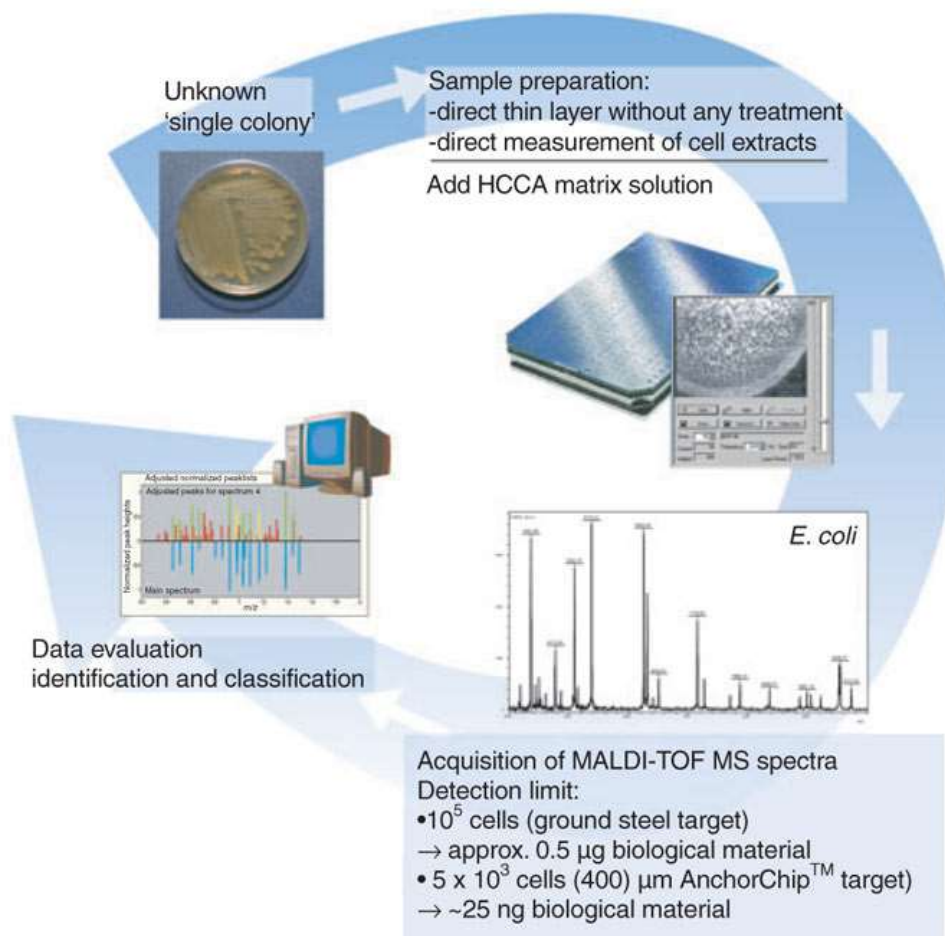
Figure 21: Principle of MALDI Biotyper microbial strains identification strategy. Bacterial colony smear is transferred on MALDI target, mixed with matrix (HCCA; α-Cyano-4-hydroxycinnamic acid) and unique spectrum of high abundant proteins is collected and MS fingerprint is compared with database. Figure adapted from [168]

It is highly expected, the connection of MS and proteomics will follow this trend. The first swallow of this statement is the proteomic MRM analysis, which became known among physicians as "mass spec ELISA" with comparable properties, such as above mentioned specificity and sensitivity. As has been already discussed, MRM has growing importance in clinical proteomics [169]. The role of MS could increase further, correspondingly with growing number of mass spectrometers and clinically validated and FDA approved methods/analytes. Clinical applicability of bottom-up or top-down proteomics will be rather limited to discovery experiments (e.g. for biomarkers) or to typization of simple matrices as before mentioned in the classification and diagnostic of amyloidosis. There are two main reasons for this opinion. The sample processing for bottom-up proteomics is time and work demanding and secondly, and according to me more important reason, is that such a profiling provides long and complex list of proteins, where substantial portion of

identificated proteins are irrelevant for diagnosis. Furthermore, about 80% of collected spectra aren't identified at all (Appendix B). This unidentified spectra are called protemics dark matter [170], which can be visualized by better search for PTMs, mutations or alternatively spliced protein forms [171]. Those tasks couldn't be done with current commercial software setup, but they can be particularly beneficial e.g. in cancer diagnostics before targeted therapy application.

2 EXPERIMENTAL PART

2.1 Aims

The first aim was to study response to anti-cancer drugs by proteomic profiling.

> During this project we were focusing on cellular response to platinum-based anti-cancer drugs on complex proteomic level.

Second aim was to establish quantitative and qualitative analysis of new cancer biomarker asporin in stably transfected *E. Coli* and breast cancer cell cultures.

2.2 Proteomic profiling of anticancer drugs

Our project is typical representative of large scale proteomic analysis in discovery phase. In this kind of experiment are analyzed proteins from biological samples with the aim to identify and quantify as much proteins as possible. This allows detecting proteins which don't have any change in expression together with differentially expressed proteins with/without biological relevance. We have applied this approach to study of drug response of in clinical practice used platinum drugs on cancer cell lines models. For this purpose, we have selected well-growing and highly chemosensitive CCRF-CEM cells derived from acute lymphoblastic leukemia.

The original and first platinum drug on the market is cisplatin (CisPt). This drug is widely used in clinical practice. The main mechanism of action of CisPt is formation of covalent adduct to DNA. Thus the most rapidly dividing cells are most affected. The most significant side effects are nephrotoxicity, nausea and resistance development [172].

To overcome several drawbacks of CisPt, new derivatives were introduced on market. From this second generation platinum drugs, carboplatin (CarboPt) is best known and most widely used [173]. CarboPt is transformed in the cell into the same active form – cis(diamine-diaqua)platinum and forms the same kind of DNA adducts [172]. However it results in cross-resistance of CisPt and CarboPt [174].

This cross-resistance led to development of third generation platinum drugs. The most known one, oxaliplatin (OxaPt), doesn't share cisplatin cross-resistance [175] and moreover, it is effective even to cancers insensitive to CisPt, e.g. colorectal cancer. OxaPt became one of treatment strategies to colorectal cancer and after discovery of its highly synergic effect with 5-fluorouracil became a golden standard of adjuvant therapy in colorectal cancer [176].

Importance of those drugs in the clinic is as well followed by particular scientific interest. The biggest effort is focused into cisplatin, carboplatin and oxaliplatin follows. This interest can be split into several ways – i) into an optimization of treatment regimes, ii) into ways how to prevent or reduce side effects, iii) into a characterization of resistance mechanisms and finally iiii) into description of molecular mechanism of drug action. The mechanism of action of platinum drugs is reviewed e.g. in [172]. However, large scale proteomic analysis of even those deeply studied drugs can reveal a new, currently unknown, target or it can

help in joining currently known mechanisms of response together and further help in designing scientifically driven, rational based, therapeutic protocols.

The second aim in this approach is to compare cellular response between those drugs. Although they are very similar and even forming the same (CisPt and CarboPt) or very similar (OxaPt) DNA adducts, they differ significantly in the spectra of toxicities. For example, the most discussed problem of CisPt is nephrotoxicity [177], in CarboPt it is myelosuppression [178] and for OxaPt peripheral neuropathy [179]. The comparison of cellular responses can suggest mechanisms involved in susceptibility of different tissues to those compounds. This comparison is best when all conditions are the same. To allow this, we have determined constant IC_{50}, which is concentration of drug which will inhibit cellular growth by 50%. The concentration of five times of IC_{50} was used in all following experiments to ensure that enough of drug has been delivered to cell. The next issue is time of treatment. The approaches used usually, which is fixed time point, like a 24 or 48 hours, don't respect the different speed of drug effect and can lead to monitoring of late apoptotic mechanisms only. We have thus decided to monitor time to caspase activation as a hallmark of apoptosis and stop the treatment in the half time of such caspase activation and start of irreversible cell death. This will prevent beginning of apoptosis and it should reveal better image of cellular response.

2.2.1 Material and methods

2.2.1.1 Chemicals and reagents

Oxaliplatin (Eloxatin) was obtained as a 5 mg/ml stock solution from Sanofi. If not mentioned otherwise, all chemicals were obtained from Sigma. Fetal calf serum (FCS) was purchased from Pan Biotech (Germany). The MagicRed assay kit was purchased from Immunochemistry Technologies, USA. RPMI media lacking arginine and lysine was obtained from Biowest, USA. HPLC/MS grade water was produced using a Merck Millipore Milli-Q Direct-8 purifier with an LC-PAK polisher (Merck), and HPLC/MS grade acetonitrile was purchased from J. T. Baker, USA. Proteomic grade trypsin was purchased from Promega. Bensonase and Luminata Forte were obtained from Merck. Pierce 660 nm Protein Assay kits for protein concentration measurements with Ionic Detergent Compatibility Reagent were purchased from ThermoFisher Scientific, USA. Pertex rapid drying medium was used for slide mounting (Histolab, Sweden). The following antibodies were used: Anti-UTP11 (AbCam), Anti-WDR46 (Sigma), Anti-TRF1 (Santa Cruz), Anti-RPS19 (Santa Cruz), Anti-p53 (Cell Signaling), Anti-Fibrilarin (Sigma) and Anti-ß-actin (Sigma), horseradish peroxidase

conjugated anti-rabbit (Sigma), horseradish peroxidase conjugated anti-mouse (Sigma), Alexa-488 anti-goat (Life technologies) and Alexa 488 anti-rabbit (Invitrogen). ELISA kits were purchased from Thermo Fisher (Ferritin) and RnD Systems (Apolipoprotein A1).

2.2.1.2 Cell culture, analysis of IC_{50} and time to apoptosis

T-lymphoblastic leukemia derived cell line (CCRF-CEM) was purchased from American Tissue Culture Collection and was cultivated according recommendation in RPMI media supplemented with 20% of FCS, 100 IU/ml penicillin and 10 µg/ml streptomycin. Inhibition of tumor growth/survival expressed by 50% inhibition constant (IC_{50}) value was determined by MTT assay as described previously [180,181]. Time to apoptosis was determined by Magic Red assay (Immunochemistry Technologies). Briefly, CCRF-CEM cells 10^{6}/ml were treated by 5 x IC_{50} concentration of OxaPt. Assay was performed according to manufacturer instructions and fluorescence was detected with excitation at 590 nm and emission 630 nm wavelength.

2.2.1.3 SILAC labeling and treatment with drugs

All experiments in this study were done in three biological replicates. CCRF-CEM cells were grown in complete RPMI media lacking arginine and lysine and supplied with either $^{12}C_{6}$ (light) or $^{13}C_{6}$ (heavy) labeled arginine and lysine. They were cultured for five passages to ensure that "light" or "heavy" amino acids were completely incorporated. The completeness of labeling was verified by mass spectrometry (data not shown). Light cell lines were diluted to concentration of 10^{6} cells per ml and treated with 5 x IC_{50} concentration of OxaPt for half time to apoptosis induction (concentration of drug and respective treatment time for CisPt 12.6 µM and 2.5 h, for CarboPt 50.7 µM and 3 h and for OxaPt 29.3 µM and 4 h). Following treatment, light and heavy cells were mixed in the same ratio. The mixed cell suspension was washed twice in ice-cold PBS with protease inhibitors and once with ice-cold PBS. Washed cells were lysed in lysis buffer containing 20 mM Tris-HCl, 7 M Urea, 10 mM DTT, 1% Triton X-100 and 0.5% SDS, 2U benzonase and transferred to 1.5 ml microfuge tubes. The cell-lysate was incubated on ice for 5 min; two extra units of benzonase were added and incubated for an additional five minutes. Lysates were cleared at 16000 g at 4°C for 10 minutes and supernatants were transferred into 1.5 ml eppendorf tubes and stored at -80°C.

2.2.1.4 Sample preparation for LC- MS/MS

Cell-lysate was resolved by preparative SDS-PAGE (MiniPrep Cell, BioRad) and the gel was cut into 20 slices representing protein fractions separated by the MW. The gel slices were

subsequently dehydrated by acetonitrile and proteins were reduced with 50mM Tris(2-carboxyethyl)phosphine at 90°C for 10 minutes and alkylated with 50 mM iodoacetamide for 1 hour in the dark. Samples were then washed three times with water and acetonitrile successively, with last 50% acetonitrile wash. Samples were solubilized and trypsinized in trypsin buffer (6.25 ng/µl trypsin, 50 mM 4-ethylmorpholine, 10% v/v acetonitrile, pH 8.3) overnight at 37°C. The supernatant was transferred into new eppendorf tubes and peptides were extracted from gel pieces successively in three steps by 80% acetonitrile with 0.1% TFA, 0.1% TFA in water and with 50% acetonitrile. Extracts were pooled and dried in SpeedVac (Eppendorf). Dried samples were reconstituted in 5 µl of 80% ACN with 0.1% TFA and diluted with 145 µl of 0.1% TFA. Reconstituted samples were purified using a C-18 MacroTrap column (Michrom Bioresources, USA), dried in SpeedVac and resuspended in 20 µl of 5% acetonitrile with 0.1% formic acid for LC-MS/MS analysis.

2.2.1.5 LC/MS

Mass spectrometric analysis was performed in parallel on three mass spectrometers. The first one was electrospray/ion trap (ESI/IT) low resolution instrument. Samples for this analysis were analyzed in nanoHPLC Agilent nano1200 (Agilent Technologies, USA) coupled with Bruker HCTUltra ESI-IT mass spectrometer (Bruker Daltonics, Germany). Analyses were performed with flow 0.4 µl.min^{-1} nonlinear gradient with water with 0.1% FA as a solvent A and ACN with 0.1% FA as a solvent B on Agilent Zorbax 300SB-C18 column with diameters 150 mm x 75 µm and particle size 3.5 µm (Agilent Technologies, Germany).

The second tested approach was using MALDI-TOF mass spectrometry coupled with LC separation. In this case, samples were separated in Agilent Capillary 1200 (Agilent Technologies, USA) coupled with Proteineer fc (Bruker Daltonics, Germany) spotting machine. Analyses were performed with nonlinear gradient with flow 4 µl.min^{-1} using MS grade water with 0.1% TFA as solvent A and ACN with 0.1% TFA as solvent B on Michrom (Michrom Bioresources, USA) Magic C18AQ column with diameters 0.2 x 150 mm and particle size 5 µ and 200 Å inner material size. Fractions were collected to 384 spots, each with 8 seconds step and 1 µl of 1 µg.ml^{-1} α-cyano hydroxycinnamic acid. Mass spectra were acquired on Bruker Autoflex III instrument (Bruker Daltonics, Germany) with *m/z* scale 600 – 3500, reflectron positive mode. Acquired spectra were automatically evaluated by WARP-LC and FlexAnalysis software (Bruker Daltonics, Gemany) and this software also selected a parent masses for subsequent MS/MS acquisition.

The third mass spectrometric approach was using an Orbitrap Elite (Thermo) instrument fitted with a Proxeon Easy-Spray ionization source (nESI-Orbitrap), coupled to an Ultimate 3000 RSLCnano chromatograph. One microliter of sample was loaded on a PepMap 100 (7 5µm x 2 cm, 3 µm, 100 Å pore size) desalting column (Thermo) "in-line" with a PepMap RSLC (75 µm x 15 cm, 3 µm, 100 Å pore size) analytical column (Thermo) heated at 35°C. The peptides were subsequently separated on the analytical column by ramping the organic phase from 5% to 35% during a total run time of 150 minutes. The aqueous and organic mobile phases were, respectively, 0.1% formic acid and acetonitrile, 0.1% formic acid. The FTMS resolution was set to 120,000 and precursor ions were scanned across an m/z range of 300.0- 1950.0. The twenty most intense ions were selected in the linear ion-trap for fragmentation by collision (CID) in the orbitrap. Collision energy of 35 eV was applied throughout.

2.2.1.6 Data analysis

Mass spectra obtained from ESI/IT were analyzed in Proteinscape 2.1.0 573 software (Bruker Daltonics, Germany), where those mass spectra were searched by two algorithms – Mascot (Matrix Science, UK) and Phenyx (Genebio, Switzerland). For Mascot, these settings were used: database NCBI, taxonomy human (Homo sapiens), enzyme trypsin, 2 allowed missed cleavages, modifications: carbamidomethylation of cysteine (fixed), oxidation of methionine (variable) and label $^{13}C_6$ on arginine and lysine (variable), peptide tolerance ± 0.5 Da, MS/MS tolerance ± 0.5 Da, peptide charge 2+ and 3+, monoisotopic mass. One peptide was accepted with mascot score at least 12 and protein was accepted with at least one peptide with mascot score higher than 40. For Phenyx, we have used these settings: database uniprot_sprot, taxonomy homo sapiens, default parent charge 2, 3, 4, scoring model HCTUltra, Trust parent charge medium, modification same as in Mascot, enzyme Trypsin (KR), parent error tolerance 0.5 Da, maximum 2 allowed missed cleavages, cleavage mode half cleaved, conflict resolution none, Turbo on, b-y ions, error tolerance 0.5 Da, coverage 20%. Search results from both algorithms were collected and processed by Proteinscape's utility Protein extractor.

Mass spectra from MALDI-TOF were processed by Proteinscape software and searched by Mascot and Phenyx in similar conditions as in ESI experiment. Differences in search methods were following: Mass charge 1, MS toleration 100 ppm, MS/MS toleration 0.5 Da.

Peak picking and peptide search for nESI-Orbitrap data was done by MaxQuant v. 1.3.0.5 [138] with SwissProt human database downloaded 4. 4. 2013. All fractions data of one

sample were searched together. Carbamidomethylation of cysteines was selected as fixed modification, oxidation of methionines and protein N-term acetylation as variable modifications. Minimal peptide and razor peptide count was set to 1 and peptide length to 6. Mass tolerance of parent ions was set to 20 ppm and tolerance of fragments was set to 0.5 Da. Peptide and protein FDR rate were set to 1%. Arginine and lysine were set as special amino acids with filtering of labeled amino acids. For quantification, razor peptides were selected with discarding unmodified counterparts of peptides, re-quantification and iBAQ were allowed. Search result was processed in Perseus 1.4.0.6, where decoy search results and contaminant items were removed and Significance B test has been performed using normalized ratios of replicates and their respective intensities. Proteins, which had significance score below 0.05 in at least two replicates, were selected for further studies. Pearson coefficients were calculated using log2 ratios of SILAC quantifications of single replicates divided by their average.

Venn diagrams were created in Perseus and Venn Diagram Plotter 1.5 (http://omics.pnl.gov). Pearson coefficients were calculated in Microsoft Excel as a plot of $\log_2$ of replicate L/H ratio vs. $\log_2$ of average L/H ratio from proteins which were quantified in all replicates.

Bioinformatics analysis of the B significant protein list was performed with the Database for annotation, visualization and integrated discovery (DAVID, v6.7, http://david.abcc.ncifcrf.gov/, downloaded 13. 1. 2016) [147]. Results were viewed using DAVID's Functional Annotation Clustering method. Protein-protein interactions were visualized using STRING database (http://string-db.org, downloaded 13. 1. 2016) [148]. STRING graphical output was then divided into clusters. Cluster was selected as close group of at least 7 proteins which binds closely together than with proteins outside of cluster.

2.2.1.7 Immunoblotting

CCRF-CEM cells were-treated by OxaPt as before. Cells were harvested and centrifuged at 90 g for 5 min at 4°C, washed twice in PBS with 5 mM sodium pyrophosphate, 1 mM sodium orthovanadate, 5 mM sodium fluoride and 1 mM PMSF. Cells were lysed with SDS lysis buffer (62.5 mM Tris-HCl, pH 6.8, 10% glycerol, 2% SDS, 1% mercaptoethanol, 0.5% bromophenol blue), heated for 10 minutes at 95°C and incubated with 2 units of benzonase for one hour at room temperature. Protein content was measured using the Pierce 660 nm Protein Assay containing the Ionic Detergent Compatibility Reagent. For SDS-PAGE electrophoresis, a 10 µg of total protein was resolved by 12% SDS-PAGE. Separated proteins

were blotted to a nitrocellulose membrane (0.2 µm pore size, Bio-Rad) using the TransBlot Turbo semi-dry system (Bio-Rad). The membrane was blocked for 1 hour in 5% non-fat dry milk in TBS/T (Tris-buffered saliva with 0.1% Tween-20) and incubated with appropriate antibodies in 5% v/v BSA in TBS/T overnight, at 4°C, with agitation. After incubation, membranes were washed with TBS/T and incubated with peroxidase labeled secondary antibody and visualized by Luminata Forte peroxidase substrate. Chemiluminiscence was collected by the HCD camera (Li-Cor Odyssey FC). Band intensities were normalized to their respective actin band and averages of three biological replicates were processed into bar graphs. Significance of protein change was calculated using t-test and p-value below 0.05.

2.2.1.8 ELISA assay

Cell lysates were prepared the same way as for immunoblotting. The proteins from 10 million of cells were precipitated by extent of ethanol and centrifuged 10 minutes with 2000 g. Pellet was reconstituted in water. The total protein amount was determined using bicinchonic acid assay. The ELISA assays to ferritin and apolipoprotein A1 were performed according to manufacturer's instructions.

2.2.1.9 Nucleoli staining

The CCRF-CEM cell line was treated the same way as for proteomic analysis. After treatment, cells were spinned to microscopy glass by Cytospin centrifuge (Sakura, Japan) at 500 rpm for 5 minutes. Spinned cells were allowed to dry, fixed by 4% paraformaldehyde in PBS and stained by toluidine blue according to Smetana [182]. After staining, slides were mounted into Pertex (Histolab, Sweden). Three replicates were observed on Zeiss Axio microscope with Zeiss AxioCamERc 5s digital camera at 400x magnificence. From each sample, 100 cells nucleoli were counted and classified according to Smetana to compact nucleoli, ring nucleoli and micronucleoli [182].

2.2.1.10   Immunofluorescence

Cells were spun using a Cytospin centrifuge as described above, fixed with 4% paraformaldehyde in PBS for 15 minutes at room temperature, and permeabilized with 0.5% TritonX-100 in PBS for 15 min. The cells were then washed three times in PBS with 0.1% Tween 20 and incubated overnight at 4 °C with primary rabbit or goat antibodies against fibrilarin, DDX56, WDR46, RPS9 and UT11L in 5% FCS. The samples were then washed three times in PBS with 0.1% Tween 20 and incubated with the corresponding anti-rabbit/goat secondary antibodies conjugated with Alexa Fluor 488 in 5% FCS for 1 hour at room temperature. Coverslips were stained with DAPI for nuclei counterstaining and

mounted in mounting medium. The localization of the proteins of interest was examined using confocal spinning disc fluorescent microscopy (Carl-Zeiss Observer Z1) at 1000 x magnification.

2.2.2 Results

2.2.2.1 Analysis of IC_{50} and time to apoptosis

The IC_{50} value of CisPt was determined to 2.52 μM, of CarboPt to 10.14 μM and IC_{50} of OxaPt to 5.9 μM by MTT cytotoxic assay. We have used five-fold of this value (5 x IC_{50}; 12.6 μM, 50.7 μM and 29.3 μM, respectively) for our experiments. Those started with Magic Red time to caspase activation analysis. This time has been set as a time, when fluorescence signal of treated cells is 10% higher than signal of control cells. CCRF-CEM cell line was treated with 5 x IC_{50} of drugs and lysed at half-time to caspase activation - 150 min for CisPt, 180 min for CarboPt and 240 min for OxaPt.

2.2.2.2 Mass spectrometry results

The first obtained information is count of fragmentation spectra. Because identification in regular proteomic experiment is done based on those spectra, it is important to have spectra of both high quantity and quality. We have acquired in average 70283 ± 10137 spectra from ESI-IT, 10862 ± 2052 spectra from this were successfully identified. MALDI-TOF acquired 73547 ± 14990 spectra and identified 3838 ± 812. Using nESI-Orbitrap, 596296 ± 32521 spectra were acquired and 109260 ± 22419 were identified (Figure 22).
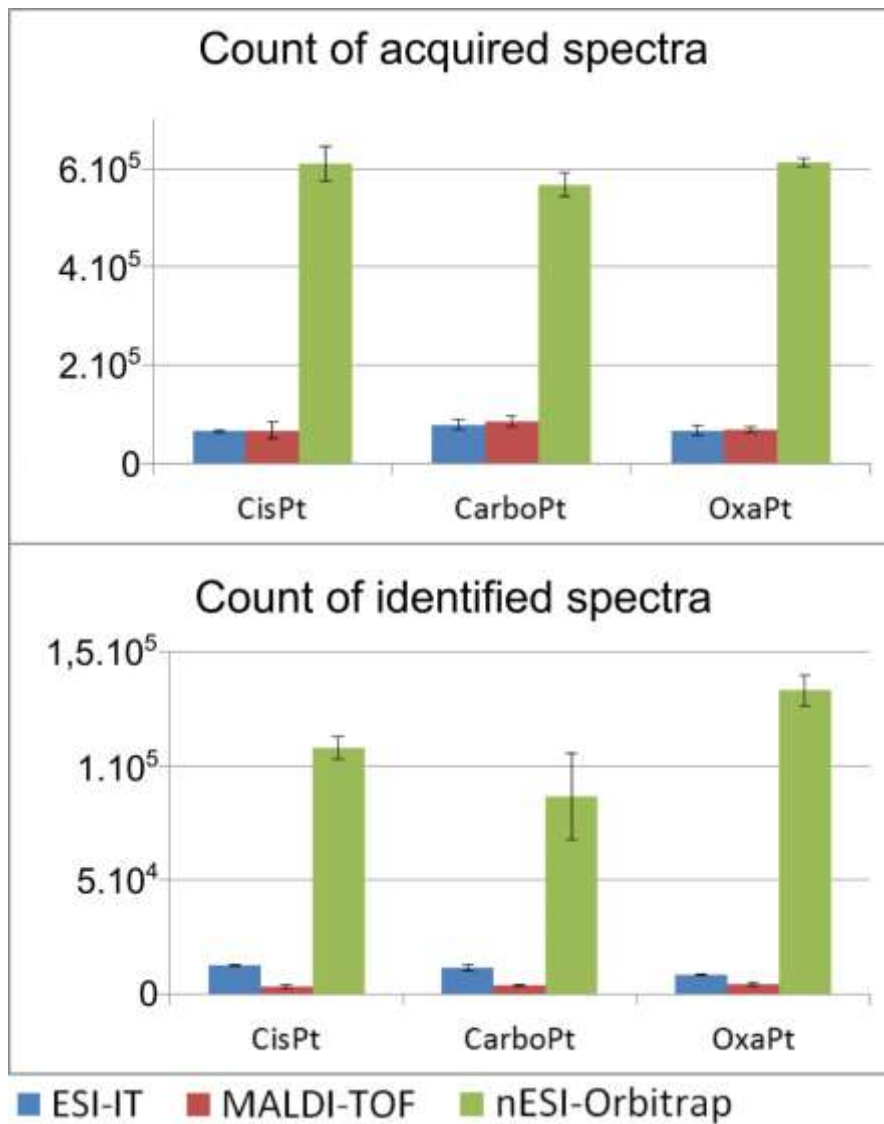
Figure 22: Count of fragmentation spectra acquired and identified using different MS methods.

Next to count of identified peptides, their reproducibility is also a very important parameter. The low reproducibility of large scale proteomic and metabolomic experiments is called undersampling [183] and it is caused by semi-random rules of peak picking prior to MS^2 analysis. In order to determine level of undersampling, we have compared reproducibility of MS method we have used in analysis. In ESI-IT was identified 6177 ± 1218 peptides common in all three biologic replicates, in MALDI-TOF 1778 ± 92 peptides and in nESI-Orbitrap was common 12296 ± 2028 peptides (Figure 23).
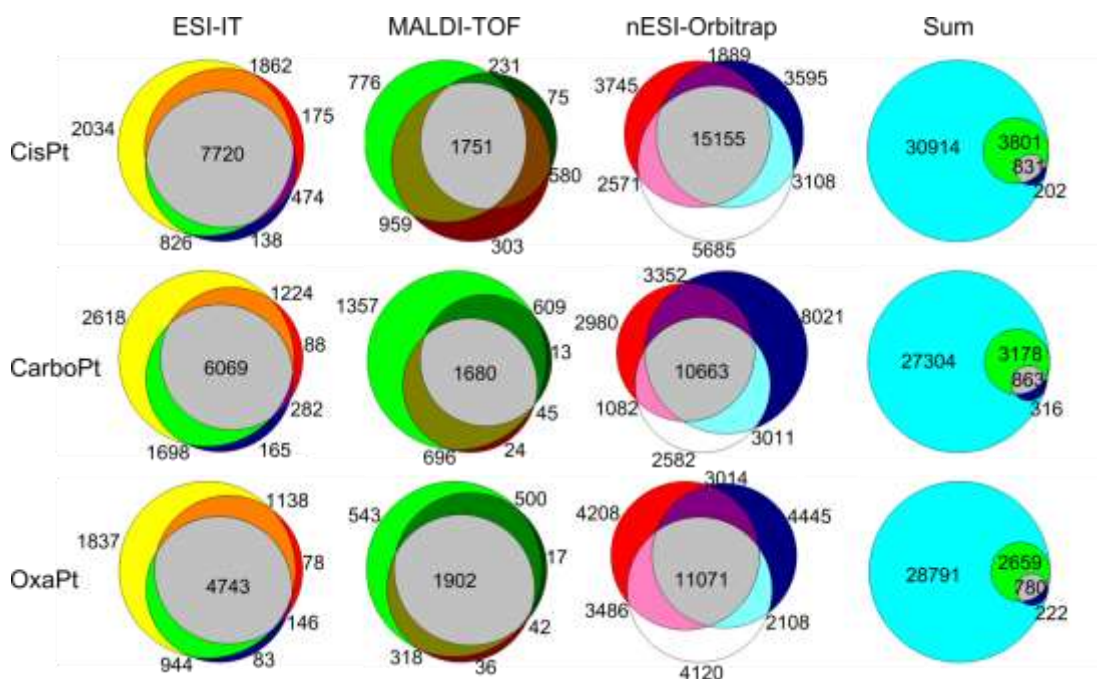
Figure 23: Numbers of identified and quantified peptides for single analyses expressed in Venn diagrams. Sum was counted from number of peptides present in at least one replicate per method. Graphics of summary diagram (Sum) is following: ▢ nESI-Orbitrap; ▢ ESI-IT a ▢ MALDI-TOF.

The count of identified proteins is important for composition of protein families or single proteins and to evaluation of their change. The more peptides were assigned to particular protein, the higher is its sequence coverage and quantification is more accurate for particular protein and method. The average count of peptide per protein was 12.0 ± 1.6 at ESI-IT, 7.4 ± 1.5 at MALDI-TOF and 5.5 ± 0.4 at nESI-Orbitrap. The respective sequence coverage was (in percent) 28.9 ± 3.0 at ESI-IT, 15.6 ± 2.8 at MALDI-TOF and 15.8 ± 1.6 at nESI-Orbitrap. The count of identified proteins was 660 ± 124 for ESI-IT with 66% coverage between all three replicates, 355 ± 68 proteins for MALDI-TOF with 41% coverage and 3430 ± 306 proteins for nESI-Orbitrap with 76% coverage (Figure 24).
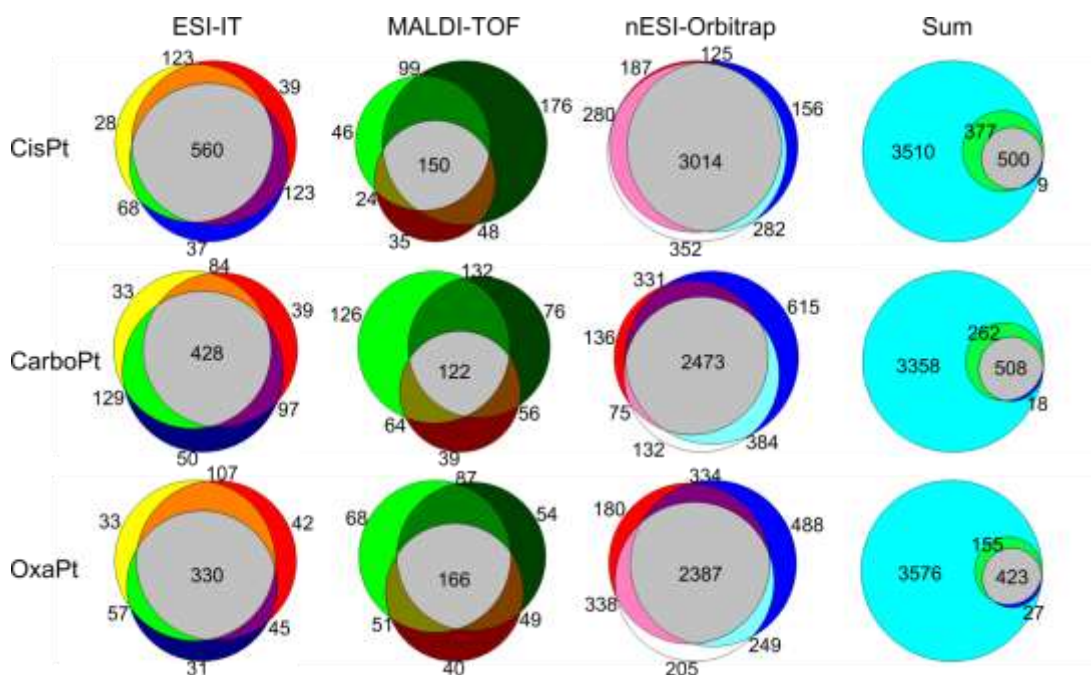
Figure 24: Numbers of identified and quantified proteins at single analyses expressed in Venn diagrams. Sum was counted from number of proteins present in at least one replicate per method. Graphics of summary diagram is following: ☐ nESI-Orbitrap; ☐ ESI-IT a ■ MALDI-TOF.

The last compared parameter is quantification accuracy expressed as R^2 among quantification values of single replicates and their average. R^2 value for ESI-IT was 0.454 ± 0.047, for MALDI-TOF 0.569 ± 0.081 and for nESI-Orbitrap 0.692 ± 0.063 (Figure 25)
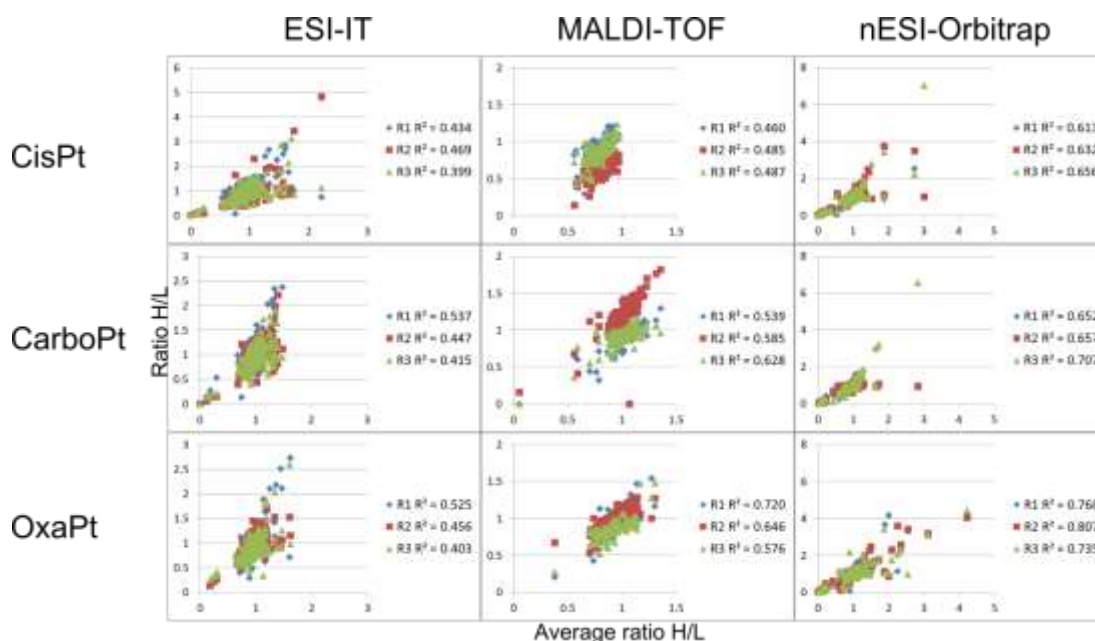
Figure 25: Quantification accuracy. For each replicate (R1, R2 and R3) was calculated coefficient of determination (R^2), to express how well regression line represents the data of individual replicate against average (all three replicates) quantification for particular drugs.

Proteins identified by ESI-IT and MALDI-TOF were evaluated together because of their reported complementarity. Total amount of identified proteins is shown in Figure 24. Lists of total identified proteins for CisPt, CarboPt and OxaPt are attached in digital form on CD. For selection of significantly regulated proteins, we have selected the most frequently used criteria [184–186], two-fold change in protein abundance. The proteins fulfilling those criteria are shown in Table 1 for CisPt, Table 2 for CarboPt and Table 3 for OxaPt.

Table 1: Proteins identified by ESI-IT and MADI-TOF and changed at least two-fold after CisPt treatment.

Accession	Protein	ESI-IT R1	ESI-IT R2	ESI-IT R3	MALDI R1	MALDI R2	MALDI R3	Average
gi|14150122	SLD5		0.05	0.09				0.07
gi|46014976	Human Mitchondrial Fission Protein Fis1		0.35	0.29				0.32
gi|33357459	Chain A, Human Pyruvate Dehydrogenase		0.63	0.34				0.49
gi|148277071	thioredoxin reductase 1 isoform 3		0.11	1.04				0.58
gi|5453549	thioredoxin peroxidase		1.12			0.15		0.64
gi|4826848	NADH dehydrogenase (ubiquinone) 1 alpha subcomplex, 5	1.01		0.27				0.64
gi|12408656	calpain 1, large subunit		0.30	1.09				0.70
gi|33468965	hypothetical protein LOC57456	0.91		0.19		1.59		0.89
gi|4503895	galactokinase 1	5.00	1.52					3.26
gi|16751921	dermcidin preproprotein	3.70	11.11	3.23				6.01
gi|223278356	glutamate receptor, ionotropic, N-methyl-D-aspartate 3A precursor				5.56		20.00	12.78
gi|3212456	Chain A, Crystal Structure Of Human Serum Albumin			6.25	100.00			53.13
gi|197927201	translocase of outer mitochondrial membrane 6 homolog	100.00	100.00					100.00
gi|7657257	translocase of outer mitochondrial membrane 20 homolog	100.00	100.00	100.00				100.00

69

Table 2: Proteins identified by ESI-IT and MADI-TOF and changed at least two-fold after CarboPt treatment.

ID	protein	ESI-IT R1	ESI-IT R2	ESI-IT R3	MALDI R1	MALDI R2	MALDI R3	Average
gi|14150122	SLD5	0.09	0.28					0.18
gi|13491174	MARCKS-like 1	0.54		0.01				0.27
gi|4505685	pyruvate dehydrogenase (lipoamide) alpha 1	0.42	0.35					0.38
A6NK07	Eukaryotic translation initiation factor 2 subunit 2-like protein	0.48	1.27			0.24		0.66
gi|36030883	RP42 homolog		2.17	0.24				1.21
gi|16751921	dermcidin preproprotein		2.33	25.00				13.66
gi|3212456	Chain A, Crystal Structure Of Human Serum Albumin	12.50	4.35		100.00	6.25	100.00	44.62
gi|39995059	caspase 2 isoform 1 preproprotein		1.14		100.00			50.57
gi|239745516	PREDICTED: hypothetical protein XP_002343510			1.49	100.00			50.75
gi|197927201	translocase of outer mitochondrial membrane 6 homolog	100.00	100.00					100.00
gi|50949594	hypothetical protein					100.00	100.00	100.00
gi|7657257	translocase of outer mitochondrial membrane 20 homolog	100.00	100.00	100.00				100.00

Table 3: Proteins identified by ESI-IT and MADI-TOF and changed at least two-fold after OxaPt treatment.

ID	Protein	ESI-IT R1	ESI-IT R2	ESI-IT R3	MALDI R1	MALDI R2	MALDI R3	Average
gi|21755646	unnamed protein product				0.01		0.01	0.01
gi|14150122	SLD5	0.15						0.15
gi|57208257	ribophorin II	0.13					1.16	0.65
gi|94536866	hypothetical protein LOC124220	2.44						2.44
gi|93279940	Human Eukaryotic Initiation Factor 4a, Eif4a				5.00	1.49	3.70	3.40
gi|47132620	keratin 2	4.17	3.70	2.44				3.44
gi|55956899	keratin 9	5.26	4.76	3.13				4.38
P02533	Keratin, type I cytoskeletal 14 (CK-14) (K14)			4.55				4.55
gi|121039	Ig gamma-1 chain C region	5.00						5.00
gi|269849769	Keratin, type I cytoskeletal 10	6.67	7.69	4.00				6.12
gi|119395750	keratin 1	3.45	3.70	2.86	8.33		12.50	6.17
gi|20150229	Chain A, Crystal Structure Of The Mrp14 Complexed With Chaps	9.09						9.09

Because ESI-IT and MALDI-TOF approaches didn't provided any meaningful results as can be seen from tables 1, 2 and 3, we have decided to re-analyze the samples using nESI-Orbitrap practically immediately after being operational in our laboratory. The numbers of identified proteins acquired from nESI-Orbitrap are shown at Figure 24. Since nESI-Orbitrap analyzes accurate mass with very high resolution, it allows more sophisticated quantification approaches. The one used in our experiment is a Significance B test, which counts protein significance not only based on its fold change, but as well based on its intensity [138]. More intense proteins are usually being quantified more precisely and thus a smaller change can be significant. Proteins significant in at least two biological replicates were selected to reduce random results. The list of B-significant proteins for CisPt is in Table 4, for CarboPt in Table 5 and for OxaPt in Table 6.

Table 4: Proteins B-significantly changed after CisPt treatment. R1, R2 and R3 are marks of biological replicates, H/L means normalized heavy to light ratio and SigB is calculated Significance B p-value. Proteins with this value below 0.05 were considered as significant.

Protein name	Gene	Protein ID	L/H R1	L/H R2	L/H R3	SigB R1	SigB R2	SigB R3
Synaptogyrin-2	SYNGR2	Q3KQZ2	0.40	0.29	0.45	2.8E-145	1.5E-296	1.5E-88
Autophagy-related protein 101	ATG101	Q9BSB4	0.75	NaN	0.18	5.3E-04	1.0E+00	5.6E-261
Mitogen-activated protein kinase 3	MAPK3	P27361	0.59	NaN	0.35	1.5E-29	1.0E+00	9.5E-89
Mothers against decapentaplegic homolog 1	SMAD1	Q15797	NaN	0.71	0.43	1.0E+00	4.1E-04	1.8E-25
SH3 domain-binding glutamic acid-rich-like protein 3	SH3BGRL3	Q9H299	0.73	NaN	0.47	2.0E-35	1.0E+00	2.5E-63
Thyroid transcription factor 1-associated protein 26	CCDC59	Q9P031	0.51	0.70	NaN	3.4E-30	1.6E-04	1.0E+00
Exportin-6	XPO6	Q96QU8	0.71	NaN	0.51	5.7E-05	1.0E+00	4.9E-15
Hydroxymethylglutaryl-CoA synthase, cytoplasmic	HMGCS1	Q01581	NaN	0.67	0.73	1.0E+00	1.1E-22	9.3E-11
Serine beta-lactamase-like protein LACTB, mitochondrial	LACTB	P83111	0.71	0.69	NaN	2.8E-07	4.8E-09	1.0E+00
Mitotic spindle assembly checkpoint protein MAD1	MAD1L1	Q9Y6D9	0.74	0.77	NaN	4.2E-09	5.5E-10	1.0E+00
Argininosuccinate synthase	ASS1	P00966	1.08	0.73	0.62	5.6E-01	7.1E-20	5.2E-43
Fatty acid desaturase 2	FADS2	O95864	0.94	0.80	0.84	1.0E-01	3.9E-11	2.1E-04
Gephyrin	GPHN	F5H039	1.04	0.85	0.85	9.5E-01	9.1E-05	6.9E-04
D-3-phosphoglycerate dehydrogenase	PHGDH	O43175	1.01	0.94	0.87	2.1E-01	2.7E-04	7.7E-09
Plasminogen receptor (KT)	PLGRKT	Q9HBL7	1.04	0.94	0.91	7.5E-01	1.6E-04	1.4E-05
Cysteine-rich protein 1	CRIP1	P50238	1.27	0.99	1.19	6.1E-08	2.0E-01	4.6E-04
Sodium/potassium-transporting ATPase subunit alpha-1	ATP1A1	F5H3A1	1.16	1.12	1.19	4.3E-05	1.1E-02	3.1E-04
CD44 antigen	CD44	P16070	0.98	1.28	1.29	3.2E-01	1.9E-04	8.4E-04
Lysophosphatidylcholine acyltransferase 1	LPCAT1	Q8NF37	2.75	1.24	1.09	1.4E-28	9.7E-09	6.5E-01
Protein S100-A9	S100A9	P06702	2.88	NaN	2.22	2.1E-09	1.0E+00	2.4E-08
Protein FAM208B	FAM208B	Q5VWN6	3.31	NaN	2.27	1.2E-18	1.0E+00	1.3E-08
WD repeat-containing protein 11	WDR11	Q9BZH6	3.13	2.98	NaN	9.8E-25	8.7E-31	1.0E+00
Lysozyme C	LYZ	P61626	3.82	3.52	4.34	4.5E-29	1.4E-12	1.5E-15
DDRGK domain-containing protein 1	DDRGK1	Q96HY6	2.64	9.07	0.96	1.3E-20	8.6E-71	2.4E-01
Cystatin-A	CSTA	P01040	7.49	3.85	3.90	1.6E-28	3.2E-15	3.7E-17
Zinc-alpha-2-glycoprotein	AZGP1	P25311	6.16	NaN	7.63	6.2E-15	1.0E+00	4.3E-17
Trypsin-3	PRSS3	P35030	14.23	NaN	4.12	6.6E-104	1.0E+00	9.3E-18
Dermcidin	DCD	P81605	16.21	7.75	8.11	0.0E+00	8.9E-68	9.8E-136
Prolactin-inducible protein	PIP	P12273	13.92	NaN	7.84	1.1E-32	1.0E+00	4.5E-50
Dystrophin	DMD	P11532	18.77	9.75	8.15	7.2E-108	1.4E-56	1.3E-23
AP-1 complex subunit mu-1	AP1M1	K7EPJ8	175.95	155.54	135.01	0.0E+00	3.0E-293	1.8E-191
Mucin-19	MUC19	Q7Z5P9	237.33	148.36	144.27	0.0E+00	4.6E-293	1.1E-191

Table 5: Proteins B-significantly changed after CarboPt treatment. R1, R2 and R3 are marks of biological replicates, H/L means normalized heavy to light ratio and SigB is calculated Significance B p-value. Proteins with this value below 0.05 were considered as significant.

Protein name	Gene	Protein ID	L/H R1	L/H R2	L/H R3	SigB R1	SigB R2	SigB R3
Histone H2A type 1-C	HIST1H2AC	Q93077	NaN	0.41	0.19	1.0E+00	0.0E+00	0.0E+00
Centrosomal protein of 41 kDa	CEP41	Q9BYV8	NaN	0.40	0.77	1.0E+00	1.4E-155	5.2E-04
tRNA (guanine(37)-N1)-methyltransferase	TRMT5	Q32P41	0.61	NaN	0.66	2.8E-24	1.0E+00	8.0E-07
F-box/LRR-repeat protein 12	FBXL12	Q9NXK8	0.68	0.79	NaN	1.6E-10	5.9E-05	1.0E+00
Ribosome biogenesis protein BOP1	BOP1	Q14137	0.66	0.88	NaN	2.6E-22	1.2E-04	1.0E+00
Galectin-1	LGALS1	P09382	0.89	0.91	0.83	1.2E-08	5.3E-06	2.9E-06
Mitotic spindle assembly checkpoint protein MAD2A	MAD2L1	Q13257	0.76	1.06	0.87	1.5E-10	7.4E-01	5.1E-04
Glucose-6-phosphate isomerase	GPI	P06744	0.85	0.95	0.93	8.1E-14	1.4E-04	5.6E-04
40S ribosomal protein S28	RPS28	P62857	NaN	0.93	0.92	1.0E+00	1.0E-04	4.5E-05
Thioredoxin reductase 1, cytoplasmic	TXNRD1	Q16881	0.95	0.96	0.93	1.5E-03	5.3E-06	2.9E-04
Stathmin-2	STMN2	Q93045-2	0.95	0.93	1.00	8.5E-04	9.2E-07	1.4E-01
Aurora kinase B	AURKB	J9JID1	1.00	1.34	1.36	7.3E-01	2.5E-05	1.9E-05
Calmodulin-like protein 5	CALML5	Q9NZT1	1.66	2.01	NaN	7.3E-23	9.3E-08	1.0E+00
Telomeric repeat-binding factor 2	TERF2	Q15554	2.74	1.09	2.70	3.6E-16	4.4E-01	1.7E-28
Leukosialin	SPN	P16150	4.23	1.01	1.73	6.7E-163	2.8E-01	4.3E-40
WD repeat-containing protein 11	WDR11	Q9BZH6	1.49	3.63	NaN	3.3E-27	1.9E-42	1.0E+00
Exportin-6	XPO6	Q96QU8	3.49	NaN	2.05	2.5E-20	1.0E+00	3.0E-06
Protein FAM208B	FAM208B	Q5VWN6	4.11	NaN	2.83	7.2E-23	1.0E+00	2.6E-09
Protein S100-A9	S100A9	P06702	NaN	4.67	3.57	1.0E+00	4.2E-50	1.5E-17
Lysozyme C	LYZ	P61626	4.72	NaN	3.75	1.2E-14	1.0E+00	1.2E-11
Protein S100-A8	S100A8	P05109	NaN	8.93	1.63	1.0E+00	1.0E-36	5.0E-04
Cystatin-A	CSTA	P01040	6.85	4.18	NaN	3.9E-73	5.6E-27	1.0E+00
Prolactin-inducible protein	PIP	P12273	9.26	NaN	3.51	7.8E-32	1.0E+00	4.0E-11
Trypsin-3	PRSS3	P35030	7.97	7.96	11.94	4.5E-121	1.3E-35	4.5E-205
Dystrophin	DMD	P11532	17.83	16.22	4.04	1.1E-35	6.7E-72	4.1E-19
Dermcidin	DCD	P81605	18.89	9.19	12.54	2.1E-255	3.0E-212	8.2E-149
Alpha-2-macroglobulin	A2M	P01023	20.74	9.43	NaN	2.2E-144	9.9E-23	1.0E+00
AP-1 complex subunit mu-1	AP1M1	K7EPJ8	199.37	191.66	NaN	5.1E-283	0.0E+00	1.0E+00
Mucin-19	MUC19	Q7Z5P9	208.11	198.08	NaN	3.8E-283	0.0E+00	1.0E+00

Table 6: Proteins B-significantly changed after OxaPt treatment. R1, R2 and R3 are marks of biological replicates, H/L means normalized heavy to light ratio and SigB is calculated Significance B p-value. Proteins with this value below 0.05 were considered as significant. Table is continuing on next page.

Protein name	Gene	Protein ID	L/H R1	L/H R2	L/H R3	SigB R1	SigB R2	SigB R3
Probable U3 small nucleolar RNA-associated protein 11	UTP11L	Q9Y3A2	0.24	0.25	0.23	0.0E+00	4.5E-134	1.5E-284
WD repeat-containing protein 46	WDR46	B4DNI0	0.28	0.28	NaN	5.2E-122	9.5E-297	1.0E+00
Ribosome biogenesis protein NSA2 homolog	CDK105	Q5J7U2	0.31	0.29	0.32	5.6E-101	8.1E-153	3.9E-60
Glioma tumor suppressor candidate region gene 2 protein	GLTSCR2	B4DVK1	NaN	0.32	0.40	1.0E+00	5.3E-112	9.4E-32
Centrosomal protein of 41 kDa	CEP41	Q9BYV8	0.23	0.47	0.45	3.0E-213	1.2E-20	5.1E-22
Protein Hook homolog 3	HOOK3	Q86VS8	0.38	NaN	0.41	3.3E-35	1.0E+00	1.5E-59
60S ribosomal protein L7-like 1	RPL7L1	Q6DKI1	0.45	0.40	0.45	7.7E-61	2.3E-301	8.6E-63
Probable ATP-dependent RNA helicase DDX56	DDX56	Q9NY93-2	0.45	0.47	0.69	1.7E-63	1.3E-34	1.9E-07
DDB1- and CUL4-associated factor 13	DCAF13	E5RHM4	0.52	NaN	0.56	3.1E-19	1.0E+00	3.7E-10
COBW domain-containing protein 1	CBWD6	F5H3X4	0.52	0.54	0.57	3.9E-18	4.2E-20	8.0E-18
Non-homologous end-joining factor 1	NHEJ1	H7C0G7	0.87	0.28	0.48	2.3E-01	2.3E-162	4.6E-17
Ribosomal L1 domain-containing protein 1	RSL1D1	Q32Q62	0.58	0.57	0.59	7.0E-53	7.3E-124	1.3E-127
Mitotic spindle assembly checkpoint protein MAD1	MAD1L1	B3KR41	0.34	0.82	NaN	3.0E-46	4.8E-04	1.0E+00
U3 small nucleolar RNA-associated protein 14 homolog A	UTP14A	Q9BVJ6-3	0.65	0.62	0.62	6.1E-15	4.7E-20	1.3E-26
Nucleolar protein 7	NOL7	Q9UMY1	0.61	0.70	0.67	1.2E-06	7.2E-11	8.8E-05
Protein FAM207A	FAM207A	C9JJU7	0.64	0.72	NaN	1.2E-05	1.1E-03	1.0E+00
U3 small nucleolar RNA-associated protein 6 homolog	UTP6	Q9NYH9	1.01	0.52	0.53	9.6E-01	1.3E-22	1.6E-12
Guanine nucleotide-binding protein-like 3	GNL3	Q9BVP2-2	0.71	0.72	0.75	8.6E-19	1.5E-22	3.0E-09
Serine/threonine-protein kinase PLK1	PLK1	P53350	0.68	0.84	0.68	4.4E-06	3.8E-03	1.2E-07
Actin-related protein 10	ACTR10	Q86TY2	0.53	0.73	0.95	1.3E-17	7.9E-14	5.6E-01
Protein CutA	CUTA	O60888-3	0.98	0.37	0.88	5.0E-01	1.5E-259	2.0E-06
ATP-dependent RNA helicase DDX54	DDX54	Q8TDD1	0.72	0.74	0.82	8.7E-05	7.7E-08	1.6E-03
Protein RRP5 homolog	PDCD11	Q14690	0.76	NaN	0.77	1.3E-12	1.0E+00	9.2E-23
Probable rRNA-processing protein EBP2	EBNA1BP2	Q99848	0.78	0.75	0.77	7.8E-06	3.0E-12	5.5E-04
Pre-mRNA-processing factor 39	PRPF39	Q86UA1	NaN	1.33	0.21	1.0E+00	4.6E-04	0.0E+00
RRP12-like protein	RRP12	Q5JTH9-2	0.78	NaN	0.78	8.6E-06	1.0E+00	1.1E-04
Probable ATP-dependent RNA helicase DDX27	DDX27	B7Z6D5	0.80	0.78	0.82	2.6E-04	2.5E-09	3.7E-05
Ribosome biogenesis protein BRX1 homolog	BXDC2	A0JLQ5	0.78	0.80	0.84	1.1E-10	5.8E-11	5.2E-03
Pumilio domain-containing protein KIAA0020	KIAA0020	Q15397	0.82	0.76	0.85	3.1E-07	8.6E-16	7.9E-09
Centromere protein F	CENPF	P49454	0.82	NaN	0.82	1.1E-03	1.0E+00	8.2E-14
Nucleolar protein 16	NOP16	Q9Y3C1	0.84	0.85	0.78	1.6E-03	2.5E-06	6.9E-05
RNA-binding protein 28	RBM28	Q9NW13	0.82	0.78	0.88	8.7E-04	1.0E-08	8.9E-03
Nucleolar and spindle-associated protein 1	NUSAP1	Q9BXS6-4	0.81	0.95	0.79	7.6E-04	1.0E-01	1.8E-06
60S ribosomal protein L35a	RPL35A	P18077	0.84	0.90	0.83	6.3E-06	6.6E-05	1.5E-02
Importin subunit alpha	KPNA2	Q6NVW7	0.79	0.93	0.86	3.1E-05	2.0E-02	2.0E-08
60S ribosomal protein L15	RPL15	P61313	0.85	0.85	0.88	4.1E-05	2.0E-08	2.8E-07
60S ribosomal protein L13	RPL13	Q6NZ55	0.88	0.82	0.88	1.5E-03	4.6E-11	4.1E-08
60S ribosomal protein L34	RPL34	P49207	0.86	0.85	0.87	1.5E-04	4.5E-08	6.6E-02
60S ribosomal protein L7a	RPL7A	P62424	0.84	0.86	0.89	1.7E-05	2.4E-07	5.2E-07
60S ribosomal protein L27a	L27a	Q9BQQ5	0.87	0.87	0.85	4.3E-04	7.6E-05	1.3E-02
60S ribosomal protein L36	RPL36	Q9Y3U8	0.85	0.83	0.92	1.6E-05	2.7E-11	2.0E-01
60S ribosomal protein L6	RPL6	Q8TBK5	0.87	0.85	0.89	3.9E-04	1.6E-08	2.4E-06
G2/mitotic-specific cyclin-B1	CCNB1	P14635	0.83	0.94	0.84	1.0E-03	4.8E-02	4.3E-04
Pescadillo homolog	PES1	B5MCF9	0.87	0.87	0.87	1.4E-02	3.1E-07	3.6E-07
Ribosomal protein L19	RPL19	Q8I WR8	0.84	0.88	0.91	6.2E-06	1.1E-05	3.2E-05
Myb-binding protein 1A	MYBBP1A	I3L1L3	0.87	0.86	0.89	6.7E-04	4.7E-02	1.5E-06
60S ribosomal protein L18a	RPL18A	M0R3D6	0.87	0.86	0.90	3.9E-04	1.4E-07	7.8E-06
40S ribosomal protein S8	RPS8	Q5JR95	0.87	0.87	0.89	3.8E-04	1.0E-06	2.2E-06
60S ribosomal protein L7	RPL7	P18124	0.86	0.87	0.91	1.2E-04	1.9E-06	2.0E-05
60S ribosomal protein L4	RPL4	P36578	0.86	0.87	0.90	1.2E-04	4.4E-06	9.5E-06
40S ribosomal protein S5	RPS5	M0R0F0	0.86	0.89	0.89	1.8E-04	4.9E-05	4.8E-07
40S ribosomal protein S16	RPS16	M0R3H0	0.88	0.87	0.89	1.1E-03	3.4E-06	1.6E-02
40S ribosomal protein S4, X isoform	RPS4X	P62701	0.88	0.87	0.90	8.8E-04	4.3E-06	6.6E-06
60S ribosomal protein L10a	RPL10A	P62906	0.87	0.88	0.91	4.8E-04	6.4E-06	5.6E-05
60S ribosomal protein L18	RPL18	G3V203	0.87	0.87	0.92	2.6E-04	3.4E-06	4.1E-04
40S ribosomal protein S23	RPS23	P62266	0.88	0.90	0.89	1.1E-03	1.6E-04	6.1E-06
60S ribosomal protein L23a	RPL23A	P62750	0.87	0.87	0.93	5.9E-04	9.9E-07	4.9E-04
60S ribosomal protein L13a	RPL13A	Q53H34	0.87	0.87	0.92	4.0E-04	4.8E-06	3.6E-04
40S ribosomal protein S24	RPS24	P62847-2	0.88	0.87	0.92	1.1E-03	3.7E-07	7.9E-02
60S ribosomal protein L14	RPL14	Q6IPH7	0.88	0.89	0.92	1.1E-03	8.4E-06	1.8E-04
60S ribosomal protein L8	RPL8	E9PKZ0	0.88	0.87	0.93	1.3E-03	3.9E-06	1.1E-03
40S ribosomal protein S3a	RPS3A	Q6NXR8	0.90	0.88	0.91	6.5E-03	7.6E-06	3.6E-05
Nucleolar GTP-binding protein 1	GTPBP4	Q9BZE4	0.87	0.90	0.92	2.2E-04	6.7E-05	1.2E-03
40S ribosomal protein S11	RPS11	M0QZC5	0.89	0.89	0.91	3.6E-03	3.3E-05	2.6E-04
60S ribosomal protein L3	RPL3	P39023	0.90	0.88	0.91	6.7E-03	1.1E-05	5.8E-05
60S ribosomal protein L21	RPL21	Q6IAX2	0.87	0.87	0.95	4.0E-04	2.6E-06	1.3E-02
Integral membrane protein 2A	ITM2A	O43736-2	0.81	1.05	0.84	2.7E-04	4.9E-01	2.7E-04

Table 6 (continued).

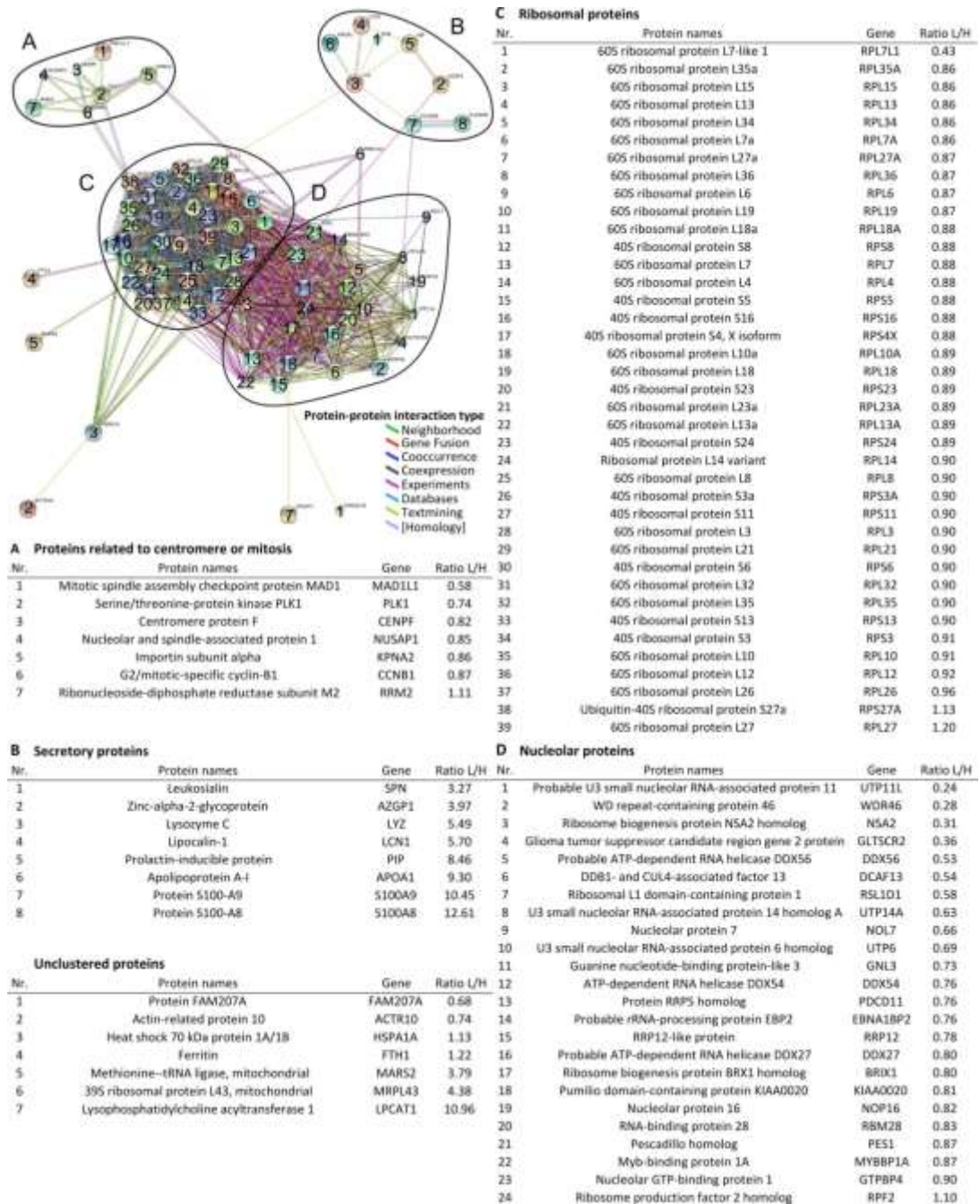| Protein name	Gene	Protein ID	L/H R1	L/H R2	L/H R3	SigB R1	SigB R2	SigB R3
40S ribosomal protein S6	RPS6	Q96DV6	0.87	0.89	0.94	4.6E-04	6.6E-05	1.9E-03
60S ribosomal protein L32	RPL32	D3YTB1	0.88	0.90	0.91	1.2E-03	4.5E-04	2.0E-01
60S ribosomal protein L35	RPL35	H0Y3A0	0.86	0.90	0.95	2.1E-04	1.4E-04	1.9E-01
40S ribosomal protein S13	RPS13	P62277	0.86	0.87	0.98	1.5E-04	1.5E-06	4.8E-01
40S ribosomal protein S3	RPS3	P23396	0.89	0.90	0.94	1.9E-03	3.4E-04	3.2E-03
60S ribosomal protein L10	RPL10	P27635	0.92	0.91	0.91	1.5E-02	5.3E-04	1.9E-04
60S ribosomal protein L12	RPL12	P30050	0.88	0.95	0.92	1.0E-03	5.7E-02	8.9E-04
Far upstream element-binding protein 3	FUBP3	Q96I24	0.96	0.91	0.92	1.8E-01	1.7E-04	8.1E-04
60S ribosomal protein L26	RPL26	Q6IBH6	0.88	0.89	1.11	7.3E-04	1.3E-04	2.6E-02
Argininosuccinate synthase	ASS	Q5T6L4	1.32	0.92	0.97	2.0E-05	1.5E-03	4.0E-01
Ubiquitin-conjugating enzyme E2 Z	UBE2Z	B4DL66	1.54	1.05	0.67	5.2E-04	7.2E-01	5.1E-05
Ribosome production factor 2 homolog	RPF2	Q5VXM9	0.94	0.87	1.50	1.7E-01	1.3E-03	2.9E-07
Ribonucleoside-diphosphate reductase subunit M2	RRM2	P31350	1.01	1.15	1.16	8.4E-01	1.2E-03	6.1E-04
Interferon regulatory factor 3	IRF3	M0R0R8	0.12	2.10	1.15	0.0E+00	9.7E-15	1.3E-01
Ubiquitin-60S ribosomal protein L40	UBB	J3QS39	1.10	1.15	1.13	5.0E-02	7.7E-07	1.2E-05
Heat shock 70 kDa protein 1A/1B	HSPA1A	P08107	1.18	1.16	1.05	3.4E-04	1.6E-07	3.0E-01
Pre-mRNA-splicing regulator WTAP	WTAP	Q15007	0.63	1.59	1.26	9.4E-09	2.2E-12	1.8E-05
60S ribosomal protein L27	RPL27	E4W6B6	1.30	0.86	1.44	2.9E-08	2.9E-07	1.8E-17
Ferritin	FTH1	G3V1D1	1.46	1.17	1.02	1.8E-03	1.3E-03	9.6E-01
Telomeric repeat-binding factor 2	TERF2	H3BR06	1.63	1.28	1.25	3.9E-08	5.0E-05	2.9E-05
DDRGK domain-containing protein 1	DDRGK1	Q96HY6	2.54	3.48	1.02	7.6E-31	2.1E-91	9.7E-01
28S ribosomal protein S26, mitochondrial	MRPS26	Q9BYN8	0.71	4.94	NaN	5.1E-05	1.7E-88	1.0E+00
Leukosialin	SPN	P16150	2.81	3.74	3.25	9.0E-35	1.0E-96	3.5E-243
Methionine--tRNA ligase, mitochondrial	MARS2	Q96GW9	2.46	5.13	NaN	2.9E-13	4.5E-54	1.0E+00
Zinc-alpha-2-glycoprotein	AZGP1	C9JEV0	5.02	4.42	2.46	8.2E-10	3.8E-14	5.4E-08
39S ribosomal protein L43, mitochondrial	MRPL43	Q8N983-4	4.81	3.95	NaN	1.5E-52	1.7E-100	1.0E+00
Lysozyme C	LYZ	P61626	13.61	2.38	0.46	1.2E-128	1.0E-17	4.2E-268
Lipocalin-1	LCN1	P31025	5.84	NaN	5.55	5.7E-18	1.0E+00	1.5E-28
Prolactin-inducible protein	PIP	P12273	7.65	9.29	8.46	1.2E-19	3.2E-66	6.4E-91
Apolipoprotein A-I	APOA1	P02647	12.48	6.13	NaN	6.5E-71	2.5E-16	1.0E+00
Dermcidin	DCD	P81605	13.67	8.46	8.19	5.4E-72	0.0E+00	1.2E-148
Protein S100-A9	S100A9	P06702	22.40	5.06	3.89	1.7E-76	6.5E-33	1.3E-23
Lysophosphatidylcholine acyltransferase 1	AYTL2	D3DTC2	9.20	NaN	12.73	5.7E-119	1.0E+00	0.0E+00
Protein S100-A8	S100A8	P05109	16.99	NaN	8.23	9.5E-133	1.0E+00	1.2E-32
Hemoglobin subunit alpha	HBA1	P69905	20.57	14.04	10.12	7.2E-76	4.9E-120	3.5E-46
Dystrophin	DMD	P11532-4	22.24	21.41	4.35	1.9E-76	1.8E-164	5.4E-114
Putative uncharacterized protein DKFZp686C11235	DKFZp686C11235	Q6MZV7	40.67	NaN	2.49	1.2E-142	1.0E+00	1.1E-15
Keratin 1	KRT1	H6VRG2	52.04	46.56	31.21	3.0E-144	2.2E-173	3.1E-181
Alternative protein CSF2RB	CSF2RB	L0R5A1	126.44	119.70	74.13	1.3E-147	0.0E+00	0.0E+00
Mucin-19	MUC19	Q7Z5P9-2	161.79	188.39	164.67	3.9E-148	0.0E+00	0.0E+00

Since proteomic profiling of OxaPt revealed the biggest amount of significantly changed proteins, I have started with biologic verification of those proteins first. The two drugs remaining, cisplatin and carboplatin are currently re-analyzed to provide better results.

2.2.2.3 Bioinformatic analysis

We have analysed data by two bioinformatic tools: DAVID for protein and process annotation and STRING for protein-protein interactions. The results from DAVID's functional annotation clustering shown that majority of significantly changed proteins are involved in ribosomes, ribosome biogenesis or those proteins are localized in nucleolus (Table 7). The output from STRING was in graphical format and is showing four main clusters (Figure 26). First cluster contains centromere or mitosis related proteins (7 proteins), second one secretory proteins involved in innate immunity response (8 proteins), third one contains ribosomal proteins (39 proteins) and fourth one contains nucleolar proteins (24 proteins). Next to those clusters, there are 7 unclustered proteins, which had connection to other proteins.

Table 7: DAVID anotation analysis of B-significant proteins. For simplicity, only annotation clusters with enrichment score higher than 10 are shown with six highest scored processes for each cluster. Annotation cluster 1 is focused to ribosomes, annotation cluster 2 to ribosome biogenesis and rRNA processing and annotation cluster 3 to nucleolus.

Annotation Cluster 1 Enrichment Score: 42.00

Category	Term	Count	%	PValue
KEGG_PATHWAY	hsa03010:Ribosome	38	36.89	6.54E-58
GOTERM_BP_FAT	GO:0006414~translational elongation	38	36.89	1.26E-57
GOTERM_MF_FAT	GO:0003735~structural constituent of ribosome	41	39.81	7.40E-56
SP_PIR_KEYWORDS	ribosomal protein	41	39.81	1.10E-54
SP_PIR_KEYWORDS	ribonucleoprotein	43	41.75	9.59E-51
SP_PIR_KEYWORDS	ribosome	31	30.10	1.14E-50

Annotation Cluster 2 Enrichment Score: 12.26

Category	Term	Count	%	PValue
GOTERM_BP_FAT	GO:0042254~ribosome biogenesis	19	18.45	7.37E-20
GOTERM_BP_FAT	GO:0022613~ribonucleoprotein complex biogenesis	19	18.45	9.22E-17
GOTERM_BP_FAT	GO:0006364~rRNA processing	15	14.56	7.94E-16
GOTERM_BP_FAT	GO:0016072~rRNA metabolic process	15	14.56	1.48E-15
GOTERM_BP_FAT	GO:0034470~ncRNA processing	15	14.56	1.76E-11
GOTERM_BP_FAT	GO:0034660~ncRNA metabolic process	16	15.53	2.27E-11

Annotation Cluster 3 Enrichment Score: 11.78

Category	Term	Count	%	PValue
GOTERM_CC_FAT	GO:0005730~nucleolus	33	32.04	1.32E-17
GOTERM_CC_FAT	GO:0031981~nuclear lumen	40	38.83	1.03E-13
GOTERM_CC_FAT	GO:0070013~intracellular organelle lumen	44	42.72	1.21E-13
GOTERM_CC_FAT	GO:0043233~organelle lumen	44	42.72	2.70E-13
GOTERM_CC_FAT	GO:0031974~membrane-enclosed lumen	44	42.72	5.36E-13
SP_PIR_KEYWORDS	nucleus	38	36.89	8.83E-04

C Ribosomal proteins

Nr.	Protein names	Gene	Ratio L/H
1	60S ribosomal protein L7-like 1	RPL7L1	0.43
2	60S ribosomal protein L35a	RPL35A	0.86
3	60S ribosomal protein L15	RPL15	0.86
4	60S ribosomal protein L13	RPL13	0.86
5	60S ribosomal protein L34	RPL34	0.86
6	60S ribosomal protein L7a	RPL7A	0.86
7	60S ribosomal protein L27a	RPL27A	0.87
8	60S ribosomal protein L36	RPL36	0.87
9	60S ribosomal protein L6	RPL6	0.87
10	60S ribosomal protein L19	RPL19	0.87
11	60S ribosomal protein L18a	RPL18A	0.88
12	40S ribosomal protein S8	RPS8	0.88
13	60S ribosomal protein L7	RPL7	0.88
14	60S ribosomal protein L4	RPL4	0.88
15	40S ribosomal protein S5	RPS5	0.88
16	40S ribosomal protein S16	RPS16	0.88
17	40S ribosomal protein S4, X isoform	RPS4X	0.88
18	60S ribosomal protein L10a	RPL10A	0.89
19	60S ribosomal protein L18	RPL18	0.89
20	40S ribosomal protein S23	RPS23	0.89
21	60S ribosomal protein L23a	RPL23A	0.89
22	60S ribosomal protein L13a	RPL13A	0.89
23	40S ribosomal protein S24	RPS24	0.89
24	Ribosomal protein L14 variant	RPL14	0.90
25	60S ribosomal protein L8	RPL8	0.90
26	40S ribosomal protein S3a	RPS3A	0.90
27	40S ribosomal protein S11	RPS11	0.90
28	60S ribosomal protein L3	RPL3	0.90
29	60S ribosomal protein L21	RPL21	0.90
30	40S ribosomal protein S6	RPS6	0.90
31	60S ribosomal protein L32	RPL32	0.90
32	60S ribosomal protein L35	RPL35	0.90
33	40S ribosomal protein S13	RPS13	0.90
34	40S ribosomal protein S3	RPS3	0.91
35	60S ribosomal protein L10	RPL10	0.91
36	60S ribosomal protein L12	RPL12	0.92
37	60S ribosomal protein L26	RPL26	0.96
38	Ubiquitin-40S ribosomal protein S27a	RPS27A	1.13
39	60S ribosomal protein L27	RPL27	1.20

A Proteins related to centromere or mitosis

Nr.	Protein names	Gene	Ratio L/H
1	Mitotic spindle assembly checkpoint protein MAD1	MAD1L1	0.58
2	Serine/threonine-protein kinase PLK1	PLK1	0.74
3	Centromere protein F	CENPF	0.82
4	Nucleolar and spindle-associated protein 1	NUSAP1	0.85
5	Importin subunit alpha	KPNA2	0.86
6	G2/mitotic-specific cyclin-B1	CCNB1	0.87
7	Ribonucleoside-diphosphate reductase subunit M2	RRM2	1.11

B Secretory proteins

Nr.	Protein names	Gene	Ratio L/H
1	Leukosialin	SPN	3.27
2	Zinc-alpha-2-glycoprotein	AZGP1	3.97
3	Lysozyme C	LYZ	5.49
4	Lipocalin-1	LCN1	5.70
5	Prolactin-inducible protein	PIP	8.46
6	Apolipoprotein A-I	APOA1	9.30
7	Protein S100-A9	S100A9	10.45
8	Protein S100-A8	S100A8	12.61

Unclustered proteins

Nr.	Protein names	Gene	Ratio L/H
1	Protein FAM207A	FAM207A	0.68
2	Actin-related protein 10	ACTR10	0.74
3	Heat shock 70 kDa protein 1A/1B	HSPA1A	1.13
4	Ferritin	FTH1	1.22
5	Methionine--tRNA ligase, mitochondrial	MARS2	3.79
6	39S ribosomal protein L43, mitochondrial	MRPL43	4.38
7	Lysophosphatidylcholine acyltransferase 1	LPCAT1	10.96

D Nucleolar proteins

Nr.	Protein names	Gene	Ratio L/H
1	Probable U3 small nucleolar RNA-associated protein 11	UTP11L	0.24
2	WD repeat-containing protein 46	WDR46	0.28
3	Ribosome biogenesis protein NSA2 homolog	NSA2	0.31
4	Glioma tumor suppressor candidate region gene 2 protein	GLTSCR2	0.36
5	Probable ATP-dependent RNA helicase DDX56	DDX56	0.53
6	DDB1- and CUL4-associated factor 13	DCAF13	0.54
7	Ribosomal L1 domain-containing protein 1	RSL1D1	0.58
8	U3 small nucleolar RNA-associated protein 14 homolog A	UTP14A	0.63
9	Nucleolar protein 7	NOL7	0.66
10	U3 small nucleolar RNA-associated protein 6 homolog	UTP6	0.69
11	Guanine nucleotide-binding protein-like 3	GNL3	0.73
12	ATP-dependent RNA helicase DDX54	DDX54	0.76
13	Protein RRP5 homolog	PDCD11	0.76
14	Probable rRNA-processing protein EBP2	EBNA1BP2	0.76
15	RRP12-like protein	RRP12	0.78
16	Probable ATP-dependent RNA helicase DDX27	DDX27	0.80
17	Ribosome biogenesis protein BRX1 homolog	BRIX1	0.80
18	Pumilio domain-containing protein KIAA0020	KIAA0020	0.81
19	Nucleolar protein 16	NOP16	0.82
20	RNA-binding protein 28	RBM28	0.83
21	Pescadillo homolog	PES1	0.87
22	Myb-binding protein 1A	MYBBP1A	0.87
23	Nucleolar GTP-binding protein 1	GTPBP4	0.90
24	Ribosome production factor 2 homolog	RPF2	1.10

Figure 26: STRING analysis of B-significant proteins. Cluster (**A**) shows centromere proteins and proteins involved in G2/M stop. Cluster (**B**) shows common secretory proteins. Cluster (**C**) shows ribosomal proteins and cluster (**D**) shows nucleolar proteins. Proteins in clusters are shown in tables with respective marking.

From STRING analysis, we have obtained four above mentioned clusters – proteins localized to centromere and involved in G2/M stop, secretory proteins involved in innate immune response, ribosomal proteins and proteins localized in nucleolus. For orthogonal verification of protein abundance, we have selected at least on protein from cluster.

A next protein cluster was created during literature survey of significantly changed protein's function. This cluster consists of proteins involved in DNA damage response. The table overview with those proteins is in Table 8.

Table 8: Significantly changed proteins involved in DNA damage response. Proteins are provided with their average fold change, intensity, Significance B p-value and reference to their involvement in DNA damage response.

Protein names	Gene names	Ratio L/H	Reference
Glioma tumor suppressor candidate region gene 2 protein	GLTSCR2	0.36	Kim et al., 2011
Non-homologous end-joining factor 1	NHEJ1	0.55	Yano et al., 2009
Pumilio domain-containing protein KIAA0020	KIAA0020	0.81	Chang et al., 2011
Nucleolar and spindle-associated protein 1	NUSAP1	0.85	Kotian et al., 2014
Ribonucleoside-diphosphate reductase subunit M2	RRM2	1.11	Lu et al., 2012
Telomeric repeat-binding factor 2	TERF2	1.39	de Lange et al., 2010

Validation of proteomic experiment

2.2.2.4 Immunoblot and ELISA verification

For antibody based verification, we have selected at least one protein from each STRING cluster (Figure 26). The most of proteins were validated by western blot, with exception of APOA1 and FTH1, for which we have used ELISA kits used in clinical biochemistry. The results from antibody verification shown at Figure 27 show nice correlation of MS results with antibody assays. There is often difference in the absolute numbers, but the trend is the same in all tested proteins.
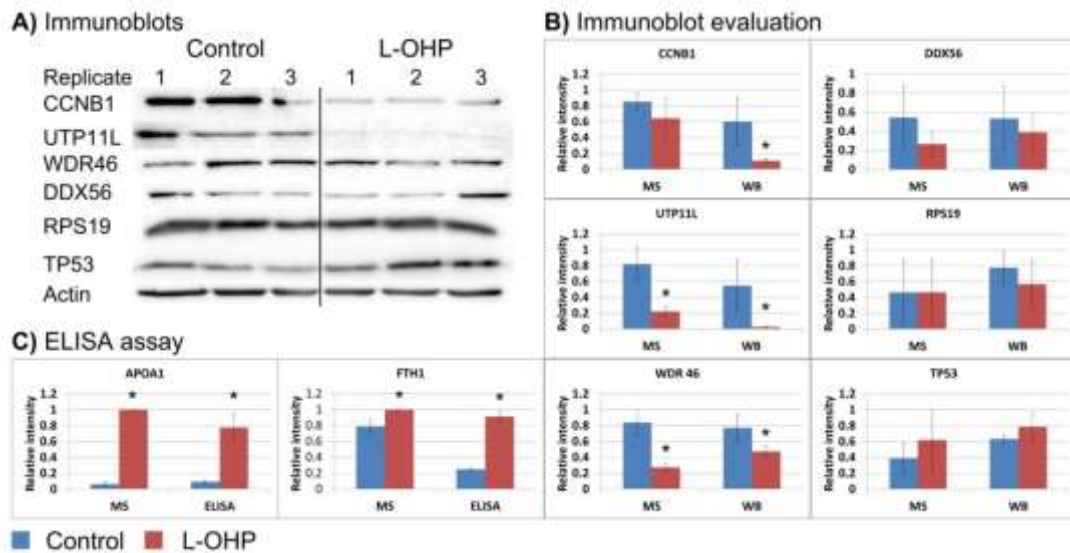
Figure 27: Immunologic verification of protein fold-change in proteins CCNB1, UTP11L, WDR 46, DDX56, RPS19 and TP53. Western blot bands in respective triplicates for OxaPt and control is shown in **A** and their normalized graphical output in **B**. Protein levels of APOA1 and FTH1 were verified by ELISA kit. Their relative intensity compared with MS intensities is in **C.** Graph columns marked with ***** have p-value below 0.05 according to t-test.

2.2.2.5 Immunofluorescence staining

In the Figure 27, we have done immunologic verification of MS results. However, MS, WB or ELISA assays don't tell anything about protein localization. To overcome this, we have performed IF staining of RPS19, DDX56 and UTP11L to elucidate their localization. As can be seen from Figure 28, all proteins localized mainly in nucleolus and followed the same expression trend after treatment.
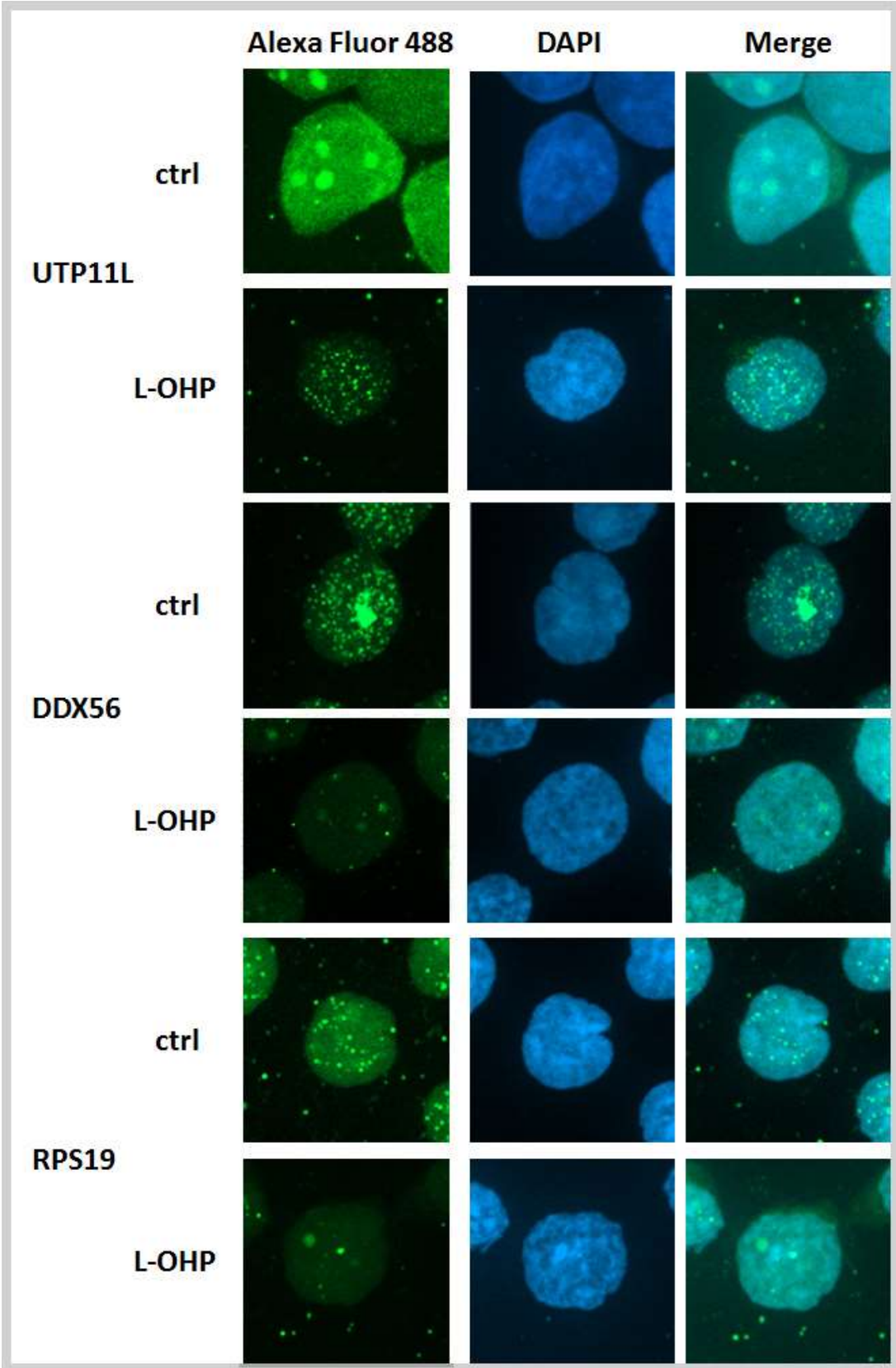
Figure 28: Immunofluorescent staining of UTP11L, DDX56 and RPS19 visualizes their nucleolar localization.

2.2.2.6 Microscopy staining

There were many of nucleolar and ribosomal proteins significantly changed after OxaPt treatment. Next step of evaluation was thus a light microscopy of whole nucleoli. For this purpose, we have used old, but proven method of toluidine blue staining with evaluation to three histologic nucleoli types: compact nucleoli, which are metabolically active; ring nucleoli, which are reversibly (e.g. during cell cycle) inactive, and micronucleoli, which are permanently inactive. In the Figure 29, we can see strong shift from compact nucleoli in control to micronucleoli after OxaPt treatment.
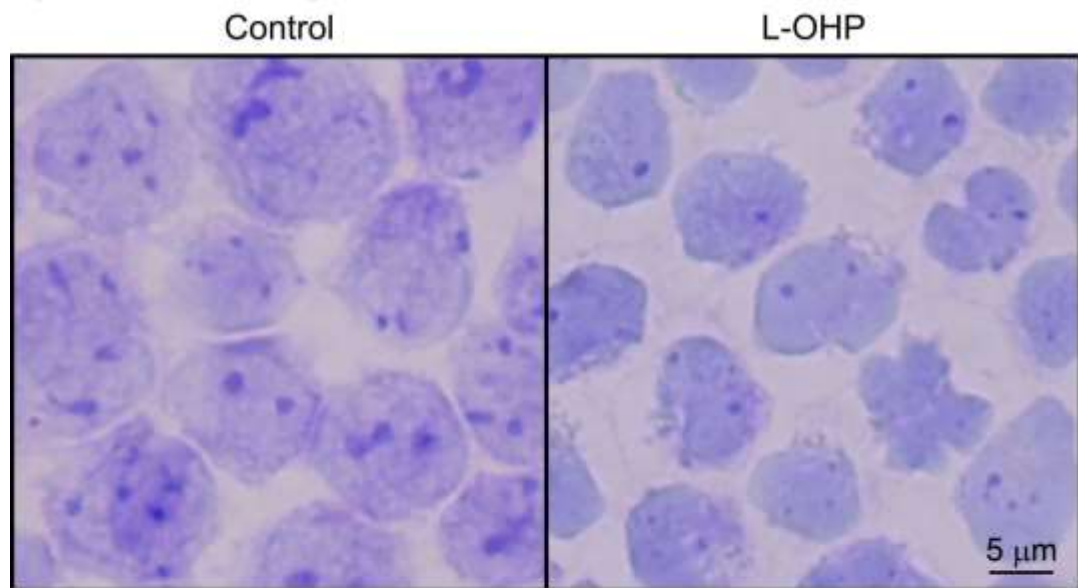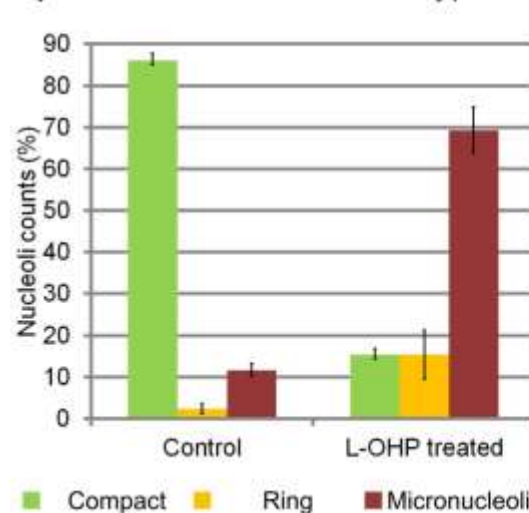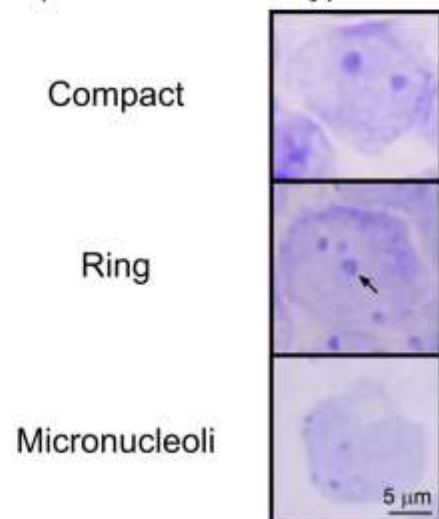
Figure 29: Toluidine blue staining of control and L-OHP treated CCRF-CEM cell line (**A**). Distribution of different nucleolar types is shown in **B** and their examples in **C.**

2.2.3 Discussion

Here I show large scale proteomic profiling experiment aimed to describe cellular response to drug treatment. In this experiment were used different mass spectrometers – ESI-IT, MALDI-TOF and nESI-Orbitrap. In this comparison, nESI-Orbitrap provided the best identification results (except of sequence coverage) from the same samples. What is interesting is a discrepancy between count of analyzed and identified spectra. This phenomenon has been already described [187] and offers a space for improvement in search engines. This discrepancy was highest in MALDI-TOF experiment, however here it is not because of search engine, but it is caused by nature of MALDI acquisition strategy. In precursor ion scan, there are all peaks with signal to noise ratio higher than 7 selected for fragmentation. The fragmentation procedure starts from most intense peaks, thus first few of spectra obtained are of high quality. However, moving to less intense peaks together with intensive laser burn of sample produces very low quality spectra for the least intense peaks.

The protein lists obtained from ESI-IT and MALDI-TOF analyses were processed together because their reported complementarity [188]. For determination of significantly changed proteins, a common condition of significance in proteomic experiments, two-time fold change ratio, was chosen [188]. Unfortunately, results are rather disappointing, since only 13-15 proteins from particular datasets were changed at least two-times. From such a few results it is impossible to elucidate cellular response to drug treatment. A re-analysis of already prepared samples on nESI-Orbitrap provided 32 proteins significantly changed in case of CisPt, 29 proteins for CarboPt and 107 proteins for OxaPt using more sophisticated Significance B algorithm [138]. The reason, why OxaPt provided such a huge amount of proteins and allowed to reconstruct several response processes (discussed below) and Cis and CarboPt didn't is still under evaluation. In the time of writing this thesis, there are new samples of CisPt and CarboPt treated cells prepared to MS analysis. Treatment conditions are based on carefully re-analyzed IC_{50} and time to apoptosis values.

We defined a protein as having been significantly up- or down-regulated if it had a significance B p-value below 0.05 in at least two replicates (Table 6). Bioinformatic analyses of these proteins were performed using the functional annotation clustering method implemented in DAVID. This method assigned high scores to ribosomal and nucleolar clusters (Table 7) A second, complementary, analysis based on protein-protein interactions was performed using STRING. This revealed four distinct clusters (Figure 26) corresponding

to proteins related to the centromere and G2/M checkpoint (Figure 26 A), secreted proteins (Figure 26 B), ribosomal (Figure 26 C) and nucleolar proteins (Figure 26 D) Proteins from clusters A, C and D are also involved in responses to cellular damage and stress [189].

Some of the proteins identified as being significantly up- or down-regulated (FAM207A, ACTR10, HSPA1A, FTH1, MARS2, MRPL43 and LPCAT1) remained unclustered in the STRING analysis (Figure 26) Literature searches on these proteins only revealed relevant information for HSPA1A, FTH1 and LPCAT1. HSPA1A functions as a protein chaperone, but it is also involved in anti-inflammatory and anti-oxidant processes [190] and DNA damage responses [191]. FTH1 is involved in innate immune reactions and is overexpressed upon lipopolysaccharide stimulation [192], anoxia or oxidative stress [193], among other things. FTH1 (ferritin) has been considered as a vehicle for platinum drugs [194]; it is possible that such drugs may form ferritin conjugates and thus induce its overexpression. LPCAT1 is reported to have anti-inflammatory properties [195]. The change in the levels of FTH1 was verified by ELISA (Figure 27).

Cluster A (Figure 26) contained proteins related to the centromere and G2/M cell cycle phase. L-OHP has been reported to arrest mitosis at the G2/M checkpoint [196]. This is consistent with our observation that L-OHP treatment caused the downregulation of cyclin B1 (CCNB1) [197] and PLK1 [198], as well as proteins interacting both with centromeres and microtubules such as NUSAP1[199], MAD1L1 [200], HOOK3 [201], CEP41[202] and CENPF [203] (Table 8). The downregulation of CCNB1 following L-OHP treatment has been observed previously [204] and was verified by immunoblotting (Figure 27).

Cluster B (Figure 26) consisted of proteins secreted extracellularly, all of which were upregulated and many of them seems to associate with histogenetic origin of CCRF-CEM cells (T lymphocytes/lymphoblasts). LYZ is involved in innate immunity and ameliorates oxidative stress [205] and has previously been reported to be elevated after cisplatin treatment [206]. SPN is a one of the major glycoproteins in T-cells; it is involved in cellular adhesion and its upregulation activates the p53 protein, which can lead to cell cycle arrest [207]. In addition, p53 plays a central role in DNA damage responses, as discussed below. Lipocalin 1 (LCN1) is a transporter of small hydrophobic molecules [208] and the increase in its expression is probably linked to L-OHP detoxification. According to a previous report [209], LCN1 is a potential biomarker of cisplatin-induced nephrotoxicity. AZGP1 is involved in lipid metabolism and has been identified as a promising biomarker of different types of carcinomas. It effects the cell cycle by downregulating the cyclin dependent kinase cdc2

and slows the G2/M transition [210]. PIP, which was also upregulated in our experiments, is a potential cancer biomarker whose abundance increases in breast and prostate cancers and which binds to AZGP1 [211] Variation of APOA1 level can be linked to poor prognosis in cancer, both without [212] and with cisplatin treatment [213]. APOA1 reportedly has anti-inflammatory, anti-apoptotic, and (of particular relevance during L-OHP treatment), antioxidant activity. As such, it is reasonable to suppose that the increase in APOA1 expression following L-OHP treatment is due to the high oxidative stress induced by the drug, so an increase in APOA1 expression can be regarded as a biomarker for effective L-OHP therapy [214]. The interaction between APOA1 and its cellular receptor, ABCA1, triggers several signaling events. These include the activation of the Cdc42 protein (which leads to cytoskeletal reorganization) and changes in the ability of other oncoproteins, including Ras and EGFR, to induce cellular transformation. In addition, binding partners of APOA1 such as APOL1 are involved in autophagy. Identifying the mechanisms that modulate APOA1 gene expression could lead to a deeper understanding of L-OHP's mechanism of action, and APOA1 expression could be monitored as a biomarker for treatment response. Two other proteins that were identified as being significantly upregulated, S100A8 and S100A9, are known to be involved in acute inflammatory responses and the induction of apoptosis via the release of nitric oxide and reactive oxygen species [215]. The differential overexpression of APOA1 following L-OHP treatment was verified using ELISA (Figure 27).

We observed statistically significant downregulation of several ribosomal proteins in cluster C (Figure 26). This is a sign of ribosomal stress, which causes the shutdown of ribosome biosynthesis [189]. However, ribosomal proteins were not regulated uniformly: some ribosomal proteins were downregulated non-significantly (e.g. RPS19, whose downregulation was verified by immunoblotting and microscopy – Figures 27 and 28) and others exhibited no detectable change in expression (e.g. RPL5, RPL11; Protein list appended electronically on CD). These differences in ribosomal protein expression may be connected to the roles of the ribosomal proteins in nucleolar stability, rRNA synthesis or p53 activation [216]. L-OHP treatment would be expected to affect ribosomes [217] and the ribosomal pathway may be involved in the development of L-OHP resistance [218].

Correspondingly to down-regulation or ribosome biosynthesis, the concentration and distribution of proteins localized in nucleolus shown in Figure 26 may indicate nucleolar stress, which reportedly causes the shrinkage or disruption of nucleoli under stress conditions [219] and has been reported to occur after L-OHP treatment [220,221]. We were

able to reproduce these results during the early stages of the L-OHP response (Figure 29) and observed a clear shift towards metabolically inactive micronucleoli (12% to 69%, Figure 29C). This phenomenon is reflected in our proteomic data, which show the downregulation of proteins involved in the small (UTP11L, UTP6, UTP14A, DCAF13, PNO1 and WDR46 [222–225]) and large (NSA2, PES1, DDX27, EBNA1BP2 (EBP2), RBM28, RPF2 and EFTUD1 [226–232]) ribosomal subunit processomes, RNA processing proteins(NOP16, DDX proteins [233,234]), nucleolar stress sensors (MYBBP1 [235]) and PES1 interactors (BRIX1, RRP12 [227]). We selected three proteins from this set – UTP11L, WDR46 and DDX56 – for immunoblot verification because of their high fold changes (Figure 26); in all three cases, immunoblotting confirmed their downregulation (Figure 27) and change of nucleolar localization was confirmed by immunofluorescence microscopy (Figure 28). Nucleolar stress was observed also using light microscopy (Figure 29).

In addition to the proteins assigned to the clusters shown in Figure 26, we found that several of the proteins identified in the proteomic analysis have previously been reported to be involved in DNA damage responses (Table 8). DNA damage is known to trigger nucleolar and ribosomal stress [189] and the formation of DNA-platinum crosslinks is well established as the main effect of platinum drugs [172]. In our experiments with L-OHP treatment, we observed the upregulation of the DNA damage response proteins TERF2 [236], RRM2 [237] as well as the downregulation of other DNA damage response proteins such as NHEJ1 [238], NUSAP1 [239], GLTSCR2 [240], and KIAA0020 [241].

Ribosomal and nucleolar stress are closely linked. Downregulation of nucleolar RNA processing and ribosomal processome proteins shuts down ribosome biosynthesis. On the other hand, nucleolar stress is triggered by DNA damage-induced inhibition of RNA polymerase I [242]. Our experiments revealed no changes in the expression of RNA polymerase I upon L-OHP treatment (Table in electronic appendix). However, nucleolar stress responses can also be activated via TOPBP1 [243] While our analysis identified TOPBP1, we were not able to quantify this protein (Table in electronic appendix).

Nonetheless, all three stress responses (ribosomal, nucleolar and DNA damage) are known to activate the p53 pathway. The activation of this pathway by DNA damage responses is well described and reviewed elsewhere, and the main pathway of DNA damage-dependent p53 activation is through nucleolar disruption [189]. Nucleolar and ribosomal stress primarily activates p53 as a consequence of the binding of MDM 2 to the free ribosomal subunits RPL5 and RPL11 [189]. We were only able to quantitate a change in the expression

of the p53 protein in one of our three replicate analyses, in which its expression increased by a factor of 1.73. The increase in its expression upon L-OHP treatment was subsequently verified by immunoblotting (Table in electronic appendix, Figure 27). Although the nucleolar stress response pathway is predominantly p53-dependent in mammals, there are other mechanisms of apoptosis activation driven by nucleolar stress. Those mechanisms occurring in cancer or organisms lacking a p53 protein homolog are reviewed by James et al. [244]. Cancer cells generally have unusually large nucleoli [245], and one of the proteins involved in the p53-independent nucleolar stress response, pescadillo (PES1), has been shown to be upregulated in p53$^{-/-}$ cells and in cancer. In keeping with our results, downregulation of PES1 reportedly leads to cell cycle arrest and apoptosis [246] (Table 6).

Introduction of nucleolar stress plays an important role in cancer treatment. This can be illustrated by the wide range of anticancer drugs that cause such stress, including L-OHP, cisplatin [247], and etoposide [242]. Indeed, some authors have suggested that nucleolar stress proteins may be viable targets for anti-cancer drug development [248] because drugs selectively targeting the nucleolus would be unlikely to cause major genotoxic stress and additional mutations.

Prior to this work, the effects of L-OHP treatment on nucleolar stress had only received passing attention, notably in Jamieson's investigation into the use of microscopy to monitor nucleolar shrinkage [220]. This shrinkage correlates with observed level of neuropathy in rats or mice, and can be modulated by sequential dosage with paclitaxel and L-OHP [220]. Study of L-OHP induced nucleolar stress was allowed using modern proteomic high resolution MS techniques based on Orbitrap mass analyzer. We have done this experiment on parallel ESI-IT and MALDI-TOF/TOF techniques previously. Hovever, using such technology, we weren't able to see significant signs of nucleolar stress except of few ribosomal proteins.

Interestingly, extensive proteomic profiling of L-OHP treated CCRF-CEM cells did not revealed activation of apoptotic machinery, thus confirming validity of our approach based on treatment of cells with cytotoxic drugs for time corresponding to half-time to induction of apoptosis.

The data from this part has been published in two articles: "The comparison of mass spectrometry approaches in proteomic profiling of drug responses" (Appendix B), accepted for publication in Chemagazín, discusses used methods and compares MS approaches. "Proteomic profiling reveals DNA damage, nucleolar and ribosomal stress are the main

responses to oxaliplatin treatment in cancer cells." shows biological interpretation of OxaPt drug response analyzed by nESI-Orbitrap. This article is submitted to Journal of Proteomics. MS data are available via ProteomeXchange with identifier PXD003543

2.3 Determination of asporin in breast cancer derived cell cultures.

The next part of this thesis is focused to 45 kDa extracellular protein asporin, which has been discovered by three independent groups in 2001 [249–251]. Asporin plays an important role in development of normal tissues, in particular of odontogenesis [252], osteoathritis [253] or in cartilage and bone development [254]. Asporin binds to transforming growth factor-β or collagen type II via aspartic acid rich N-terminal region and central part of protein [255]. Asporin inhibits collagen fibrilogenesis by competition with decorin in binding the same sites of collagen. This competition may have a role in regulating the development of extracellular matrix. The asporin N-terminal polyaspartate domain also binds calcium, and works in concert with other domains in order to initiate the mineralization of collagen [256]. The crystal structure of asporin hasn't been solved yet, however, there is a computational model of asporin (Figure 30) available from Protein model portal [257].
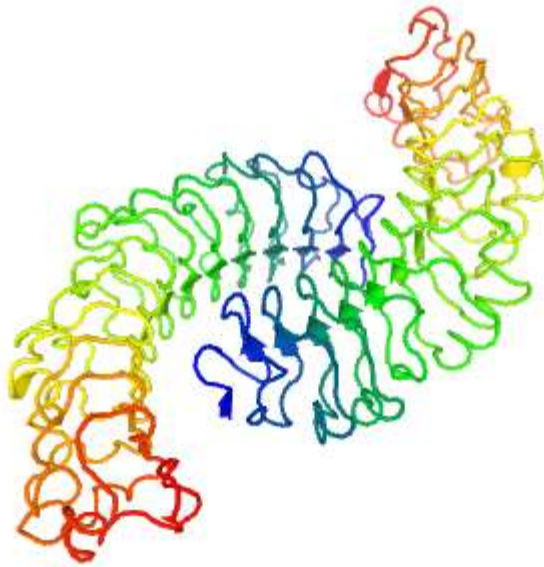
Figure 30: A computational model of asporin (accession Q9BXN1) made by Swissmodel group from University of Basel. Model is accessible via Protein model portal [257].

Asporin is involved in disease development as well. Genetic polymorphisms of asporin have been associated with various bone and joint diseases, including osteoarthritis, rheumatoid

arthritis and lumbar disc disease [258]. Asporin has been studied e.g. in Duchenne muscular dystrophy [259], in tissue samples of degenerative mitral valve disease [260], or in cancer biology, mainly in the sense of elucidating its function [261,262] or as a possible biomarker [263,264].

The current study is focused to role of asporin in cancer progression. Asporin is connected with higher metastatic potential in pancreatic cancer [263] and faster proliferation in gastric cancer [262]. On the other side it has been found as a tumor suppressor in breast cancer [265]. The role of asporin in growth, migration and invasion of cancer cells was studied by Dana Šimková (Appendix C).The proteomic task in this project was to verify antibody specificity, what is supposed to be better approach than verification of MS results by antibody based method [155], and provide orthogonal quantification information.

2.3.1 Material and methods

2.3.1.1 Production of recombinant asporin

One milliliter aliquot of frozen bacterial culture containing the plasmid carrying the gene for asporin was thawed and suspended in 100 ml of LB medium (10 g NaCl, 10 g tryptone, 5 g yeast extract per liter) with kanamycin (75 mg/L) and incubated overnight on a heated rotation shaker (37°C). The next day this starting culture was added to 1 liter of fresh LB medium containing kanamycin. The culture was again incubated on rotation shaker until reaching the sufficient density for the induction (O.D.600=0.6). Expression of the recombinant asporin was induced by Isopropyl-β-D-thiogalacto-pyranoside (IPTG, final concentration 1 mM), bacteria were incubated for another 5 hours and then harvested by the centrifugation. Bacterial pellet was resuspended in denaturating lysis buffer (8 M urea buffered by 50mM Tris, pH 8), sonicated (10x10 s, 2 min intervals) and centrifuged (10000g, 10 min, room temperature). The obtained supernatant was used for protein purification using the metaloaffinity chromatografy with Ni-NTA agarose. Washing and elution was performed using buffers with decreasing pH (8 M urea, pH 6.3; 5.9 and 4.5).

2.3.1.2 Cell lines

All experiments with human cell lines and *E. Coli* clones were done by co-authors of recently published article (Appendix C). Briefly, Hs578t cells, obtained from ECACC (Salisbury, UK), MDA-MB-231 and BT-549, obtained from American Type Culture Collection (Rockville, MD, USA) and the gingival fibroblasts obtained from healthy donors were grown in Dulbecco's Modified Eagle's Medium (DMEM; Life Technologies, Carlsbad, CA)

supplemented with 10 % foetal bovine serum (FBS, Thermo Scientific, MA, USA), 1% penicillin/streptomycin (Life Technologies) and insulin (10 µg/ml) (Life Technologies).

For asporin overexpression, *E. Coli* and two human cell lines, MDA-MB-231 and BT-549, were transfected either with ASPN full length sequence (TrueClone, pCMV6-AC, Origene) or with open reading frame (TrueORF, pCMV6, Origene) sequence using Neon Transfection System (Life Technologies). Stable cell lines were selected with 0.5 mg/ml geneticin for 2 weeks and then kept under low selection pressure at 0.1 mg/ml geneticin.

2.3.1.3 Immunoblotting

Cells were harvested into RIPA buffer (50 mM Tris HCl pH 8.0; 150 mM NaCl; 1% NP-40 0.5% sodium deoxycholate; 0.1% SDS) supplemented with protease/phosphatase inhibitor cocktail (Roche). Twenty micrograms of whole cell lysate were separated by electrophoresis in 10% Bis-Tris polyacrylamide gel followed by blotting to nitrocellulose membrane. Non-specific binding sites were blocked by incubating the blots for 2 hrs at room temperature with 5% (w/v) non-fat dry milk in PBS. Blots were incubated overnight with primary antibodies at the following concentrations: anti-ASPN (1:1000; # HPA008435; Sigma-Aldrich, Protein Atlas; plus other antibodies specified in Supplementary Table 1), anti-FAK pY397 (1:500, Life Technologies), and anti-GAPDH (1:25000; Sigma-Aldrich) was used as a loading control. Secondary antibodies were as follows anti-rabbit IgG, HRP-linked Antibody (#7074) and Anti-mouse IgG, HRP-linked Antibody (#7076), both purchased from Cell Signaling Technology, MA, USA. Signal detection was performed with Dura/ Femto ECL western blotting substrate (Thermo Scientific, MA, USA).

2.3.1.4 In-gel and on-membrane digestion

In-gel digestion was performed as described in chapter 1.2.1.4. SDS-PAGE gel was stained by Coomassie stain and photographed using Bio-Rad GS-800 calibrated densitometer. Selected bands were chopped, destained using 50 mM ammonium bicarbonate/acetonitrile washes, reduced with 50 mM Tris(2-carboxyethyl)phosphine at 90°C for 10 minutes and alkylated with 50 mM iodoacetamide for 1 hour in the dark. Samples were then washed three times with water and acetonitrile successively, with last 50% acetonitrile wash. Samples were solubilized and trypsinized in trypsin buffer (6.25 ng/µl trypsin, 50 mM 4-ethylmorpholine, 10% v/v acetonitrile, pH 8.3) overnight at 37°C. The supernatant was transferred into new eppendorf tubes and peptides were extracted from gel pieces successively in three steps by 80% acetonitrile with 0.1% TFA, 0.1% TFA in water and with 50% acetonitrile. Extracts were pooled and dried in SpeedVac (Eppendorf). Dried samples

were reconstituted in 5 µl of 80% ACN with 0.1% TFA and diluted with 145 µl of 0.1% TFA. Reconstituted samples were purified using a C-18 MacroTrap column (Michrom Bioresources, USA), dried in SpeedVac and resuspended in 20 µl of 5% acetonitrile with 0.1% formic acid for LC-MS/MS analysis.

On-membrane digestion was performed according to Luque-Garcia and Neubert [266]. Immunoblotted nitrocellulose membrane before blocking was stained with 0.2% (w/v) Ponceau S, chopped, blocked with 0.5% (w/v) polyvinylpyrrolidone PVP-40 and incubated 30 min at 37°C with agitation. After incubation with PVP-40, membrane slices were washed three times with MilliQ water and digested using 12.5 ng/µL trypsin dissolved in 50 mM ammonium bicarbonate overnight. The peptides were isolated using excess of acetone added to dry membrane. Acetone dissolves nitrocellulose and currently precipitates peptides. Samples were centrifuged 10 min at 14,000 g at room temperature, supernatant was removed and precipitate was allowed to dry. Dried precipitate was reconstituted prior MS analysis in 5% acetonitrile and 0.1% formic acid buffer.

2.3.1.5 MALDI-TOF analysis

The LC-MALDI-TOF analysis was performed similar way as in proteomic profiling experiment (Chapter 1.2.1.5). Briefly, samples were separated in Agilent Capillary 1200 (Agilent Technologies, USA) coupled with Proteineer fc (Bruker Daltonics, Germany) spotting machine and Michrom (Michrom Bioresources, USA) Magic C18AQ column with diameters 0.2 x 150 mm and particle size 5 µ and 200 Å inner material size. Fractions were collected to 384 spots, each with 8 seconds step and 1 µl of 1 µg.ml^{-1} α-cyano hydroxycinnamic acid. Mass spectra were acquired on Bruker Ultraflextreme instrument (Bruker Daltonics, Germany) with *m/z* scale 700 – 3500, reflectron positive mode. Acquired spectra were automatically evaluated by WARP-LC and FlexAnalysis software (Bruker Daltonics, Gemany) and this software also selected a parent masses for subsequent MS/MS acquisition. Obtained mass spectra were analysed by Proteinscape 3.1.0 348 software (Bruker Daltonics, Germany), where those mass spectra were searched by two algorithms – Mascot (Matrix Science, UK) and Phenyx (Genebio, Switzerland). Search results from both algorithms were collected and processed by Proteinscape's utility Protein extractor.

2.3.1.6 nESI-Orbitrap analysis

nESI-Orbitrap analysis was performed the same way as in previous project. An Orbitrap Elite (Thermo) instrument fitted with a Proxeon Easy-Spray ionization source (nESI-Orbitrap), coupled to an Ultimate 3000 RSLCnano chromatograph. One microliter of sample was

loaded on a PepMap 100 (75 µm x 2 cm, 3 µm, 100 Å pore size) desalting column (Thermo) "in-line" with a PepMap RSLC (75 µm x 15 cm, 3 µm, 100 Å pore size) analytical column (Thermo) heated at 35°C. The FTMS resolution was set to 120,000 and precursor ions were scanned across an m/z range of 300.0- 1950.0. The twenty most intense ions were selected in the linear ion-trap for fragmentation by collision (CID) in the orbitrap. Collision energy of 35eV was applied throughout. MS data search was performed using MaxQuant software package with SwissProt human database downloaded 4. 4. 2013.

2.3.1.7 MRM targeted analysis

MRM assay was performed on AB Sciex QTrap 5500 mass spectrometer coupled with Dionex Ultimate 3000 LC. Two standard peptides with sequences LYLSHNQLSEIPLNLPK and YWEMQPATFR (JPT Peptide technologies, DE) were used for optimization with five most intense transitions for the first peptide and eight for the second (Table 9).   Mass spectrometry parameters were following: Ion source gas flow (GS1) 10 l/min, curtain gas flow (CUR) 10 l/min, Ion spray voltage 2800 V. Liquid chromatography conditions were following: Column was a Thermo PepMap RSLC (75 µm x 15 cm, 3 µm, 100 Å pore size), mobile phase A was water with 0.1% formic acid, B was acetonitrile with 0.1% formic acid. Total analysis length was 60 minutes with following gradient: 0 – 10 min 5% B; 10 - 40 min increase to 35% B; 40 - 41 min increase to 95% B; 41 – 45 min hold for 95% B; 45 – 46 min decrease to 5% B and the rest of analysis was equilibration at 5% B. One microliter of analyzed sample, either standard, positive control or unknown sample was loaded. Resulting spectra were Gaussian smoothened and exported from Analyst software (AB Sciex).

Table 9: MRM transitions of LYLSHNQLSEIPLNLPK and YWEMQPATFR peptides. Q1 is mass set on first quadrupole, Q3 mass set on third one. DP is Declustering Potential, CE Collision Energy and CXP Collision Cell Exit Potential.

Peptide	Q1 (Da)	Q3 (Da)	Time (ms)	DP (V)	CE (V)	CXP (V)
LYLSHNQLSEIPLNLPK (3+)	660.43	681.50	150	121	25	18
	660.43	649.90	150	121	19	16
	660.43	593.40	150	121	23	50
	660.43	682.40	150	121	25	20
	660.43	438.40	150	121	37	36
YWEMQPATFR (2+)	664.63	608.50	50	141	27	16
	664.63	591.40	50	141	29	18
	664.63	552.40	50	141	35	40
	664.63	979.50	50	141	29	28
	664.63	551.50	50	141	35	12
	664.63	719.40	50	141	27	26
	664.63	609.40	50	141	29	16
	664.63	592.20	50	141	31	18

2.3.1.8 PRM targeted analysis

PRM targeted analysis was performed on an Orbitrap Fusion (Thermo Scientific) instrument fitted with a Proxeon Easy-Spray ionization source, coupled to an Ultimate 3000 RSLCnano chromatograph. Ten microliter of sample was loaded on a μ-Precolumn C18 PepMap 100 (300 μm x 5 mm, 5 μm, 100 Å pore size) desalting column (Thermo Scientific) "in-line" with a PepMap RSLC (75 μm x 50 cm, 3 μm, 100 Å pore size) analytical column (Thermo Scientific) heated at 35°C. The peptides were subsequently separated on the analytical column by ramping the organic phase from 5% to 35% during a total run time of 165 minutes. The aqueous and organic mobile phases were 0.1% formic acid diluted in water or acetonitrile, respectively. There were two parallel experiments running simultaneously on mass spectrometer. The first one was single FTMS scan with resolution set to 120,000 and precursor ions scanned across an m/z range of 400- 1600. The second experiment was targeted MS2 with HCD collision and Orbitrap detector. Resolution was set to 30,000 and HCD collision energy to 35%. The resulting spectra were evaluated in Skyline-daily 2.6.1.6899 software [267].

2.3.1.9 Linearity assay

HeLa cells were lysed using RIPA buffer (50 mM Tris HCl pH 8.0; 150 mM NaCl; 1% NP-40 0.5% sodium deoxycholate; 0.1% SDS) on ice for 30 min and lysate was centrifuged at cooled centrifuge with speed 20000 g for 10 min. Protein content was measured using the

BCA assay (Sigma). Ten microgram proteins from this lysate were spiked with concentration range from 100 ng to 10 zg of recombinant asporin. Spiked samples were digested in-solution using incubation with 6 µM DTT for 30 minutes at 57°C for protein reduction, 11 µM of iodacetamide for 1 hour at dark for sample alkylation and 12 ng/µl of trypsin for overnight digestion at 37°C. Samples after digestion were evaporated using vacuum centrifuge, reconstituted in water with 5% acetonitrile and 0.1% TFA, purified using Macrotrap (MiChrom) column and analyzed by PRM analysis.

2.3.2 Results

2.3.2.1 Quality control of stably transfected E. Coli

The first task of this project was to validate effectivity of two clones of stably transfected *E. Coli* expressing recombinant asporin. In the first case, His-tag affinity purified proteins from the first one has been separated on SDS-PAGE gel and chopped gel slices were subjected to MALDI-TOF analysis with search against human and *E. Coli* database. Resulting scores of *E. Coli* protein are higher and better corresponding with band intensity. The identification scores of human proteins are – with exception of Heat shock protein HSP90 in band 6 – which may be caused by a high homology of chaperones accross species (Figure 31).

Band	Database	Protein	Score
1	*E. Coli*	Glutamine--fructose-6-phosphate aminotransferase	618.6
	Human	Lymphoid-restricted membrane protein	38.9
2	*E. Coli*	Alkyl hydroperoxide reductase	2351.1
	Human	Src substrate cortactin	69.6
3	*E. Coli*	Elongation factor Tu 2	359.7
	Human	putative	34.6
4	*E. Coli*	Glycerol dehydrogenase	556.0
	Human	Lymphoid-restricted membrane protein	31.7
5	*E. Coli*	Uncharacterized protein YbiB	887.4
	Human	Lymphoid-restricted membrane protein	34.7
6	*E. Coli*	FKBP-type peptidyl-prolyl cis-trans isomerase slyD	496.2
	Human	Heat shock protein HSP 90-beta	323.8
7	*E. Coli*	GTP cyclohydrolase 1	709.6
	Human	Kinesin-like protein KIF27	74.3
8	*E. Coli*	Catabolite gene activator	1309.1
	Human	Lamin-B1	33.0

Figure 31: Coomassie blue stained SDS-PAGE gel of His-Tag purified *E. Coli* lysate. In the table are shown proteins with the highest score in particular band for *E. Coli* and human database.

The second stably transfected *E. Coli* clone was prepared as three different cellular lysates – first, "native" lysis buffer contained Tris buffered saliva with lysozyme only (N), lysis buffer

with 1% SDS (S) and lysis buffer with 8M Urea (U). Those cellular lysates has been separated on semi-preparation SDS-PAGE with 200 µg of protein loaded. Two fractions from each sample, shown on Figure 32, were chopped for further analysis.
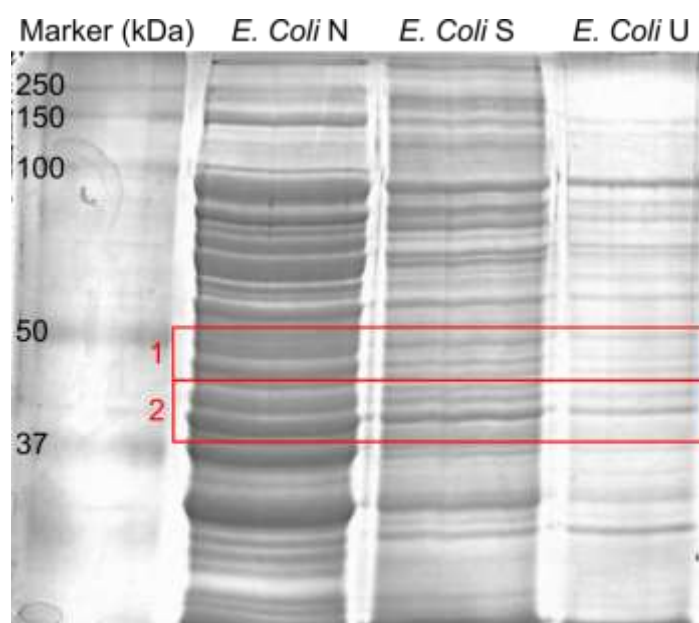
Figure 32: SDS-PAGE gel resolving three different *E. Coli* lysates. N means for cells lysed with native lysis buffer, S with SDS lysis buffer and U with urea lysis buffer. Area marked red was chopped for MS analysis.

In-gel digested bands were subjected to parallel MALDI-TOF and nESI-Orbitrap analyses. This time, asporin was identified in all samples (Table 10) with different intensity and probability score. MALDI-TOF and nESI-Orbitrap are independent methods with different scoring system. MALDI-TOF uses MASCOT probability score and MaxQuant used in evaluation of nESI-Orbitrap results summed intensity. Both values were followed by their respective sequence coverages and shown in Table 10 for comparison of lysis buffer efficiency. Lysates produced with SDS lysis buffer provided the highest yields of asporin from stably transfected *E. Coli* clone. This lysate was further used as a positive control in the targeted analyses.

Table 10: MALDI-TOF and nESI-Orbitrap results of asporin identification in *E. Coli* lysates. Score is expressed in Mascot probability score and intensity, which is sum of all peptide ion intensities per protein. Sequence coverage (SC) is the only qualitative parameter, which can compare both mass spectrometers.

Sample	MALDI-TOF		nESI-Orbitrap	
	Score	SC (%)	Intensity	SC (%)
N1	79.6	4.5	6.4E+06	12.1
N2	NA	NA	5.0E+06	21.3
S1	1191.6	33.7	8.5E+07	46.1
S2	1362.8	36.3	6.4E+07	40.8
U1	750	21.8	1.4E+07	27.1
U2	909.2	32.1	9.3E+06	21.1

2.3.2.2 Antibody verification on cell lines

The next task was verifying of antibody specificity by MS techniques according to Aebersold [155]. For this reasons, two nitrocellulose membranes were run in parallel. One was processed in regular way with anti-asporin antibody and chemiluminiscence detection (Figure 33). Based on this information, positively stained bands from non-blocked second membrane was chopped and processed by on-membrane digestion protocol.
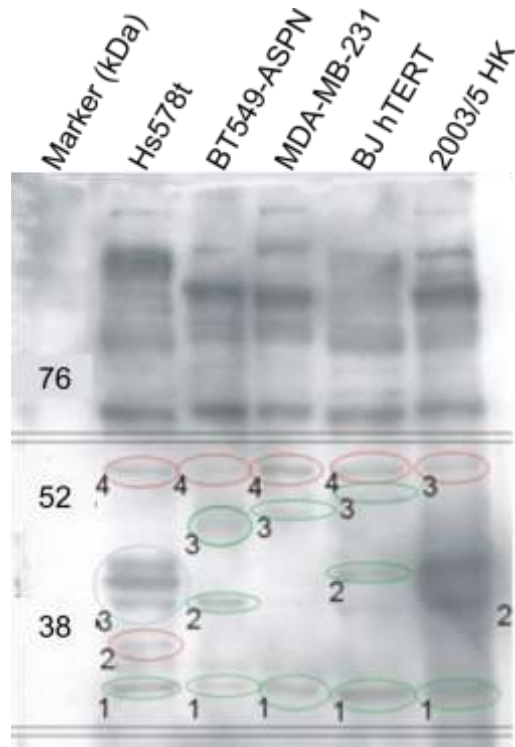
Figure 33: WB membrane of cell lines incubated with Lund anti-asporin antibody (Appendix C). Rounded bands were chopped and prepared for MS analysis.

MS analysis was performed in parallel by MALDI-TOF and nESI-Orbitrap. In average, we have identified 9 ± 4 proteins per spot by MALDI-TOF and 546 ± 123 proteins by nESI-Orbitrap. However, asporin haven't been identified in any of samples. The next step was to analyze a preparative electrophoretic gel with 200 µg of protein loaded (Figure 34). This gel was run in parallel with immunoblot.
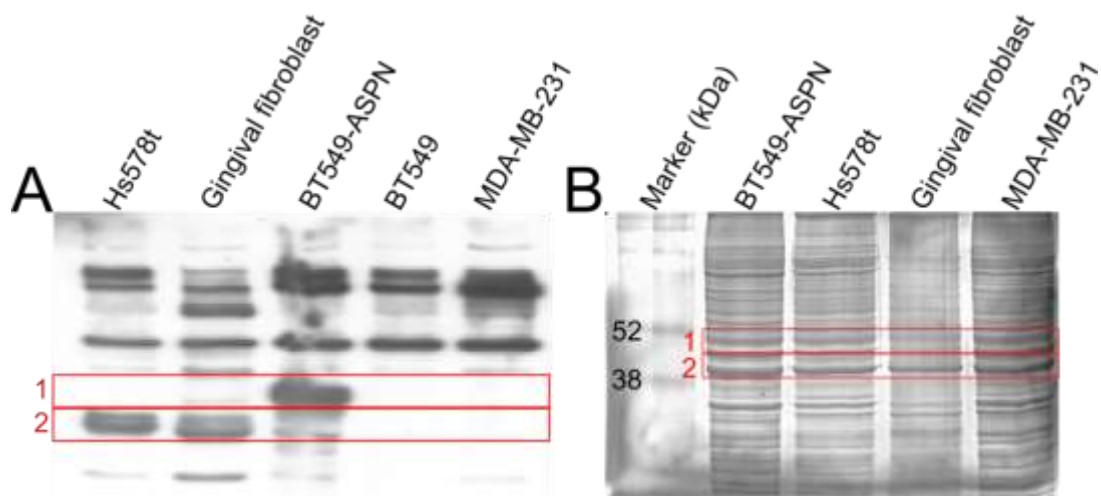
Figure 34: Parallel WB analysis (A) and preparative SDS-PAGE gel (B). Red squares indicate the range of molecular weights chopped for successive MS analysis.

The digests from preparative electrophoresis were analyzed on nESI-Orbitrap. The results were very similar to above mentioned. Unfortunately, asporin wasn't detected even in those samples. The fail of identification approach led us to development of targeted MS assays. Digested extracts frompreparative electrophoresis were used for optimization of targeted approach.

2.3.2.3 Targeted MS analysis

Targeted MS analyses were performed on two instruments. The first one was quadrupole-ion trap hybrid mass spectrometer. We have optimized transitions for peptides LYLSHNQLSEIPLNLPK and YWEMQPATFR according to Table I. The chromatogram of synthetic peptides with 10 fmol amount is provided in the Figure 35.
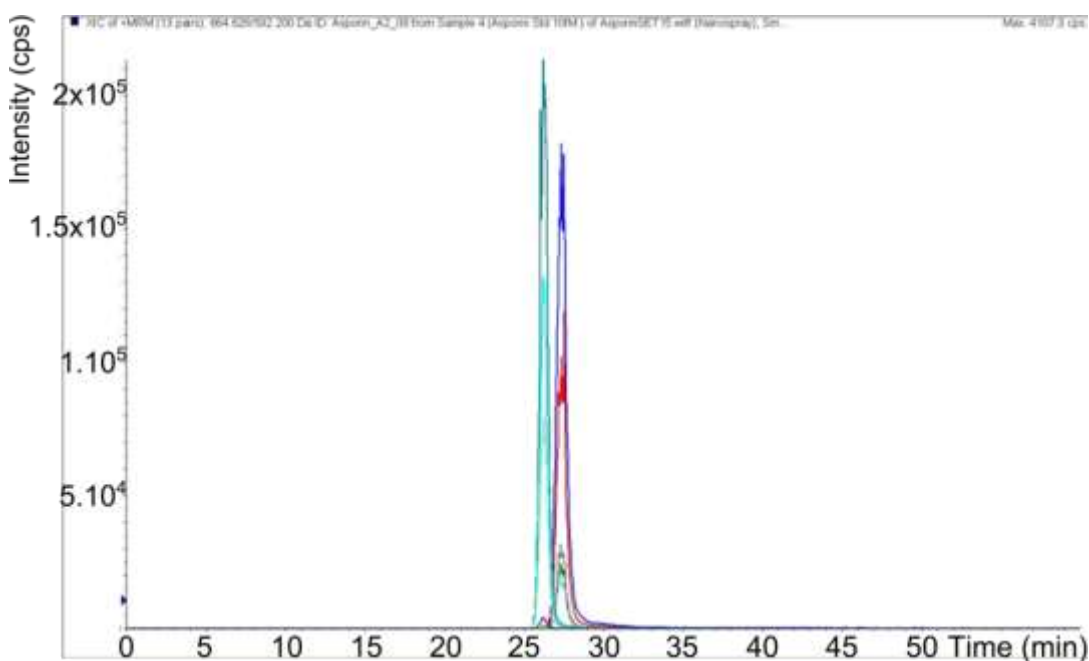
Figure 35: Chromatogram of MRM transitions of 10 fmol peptide standards with sequence LYLSHNQLSEIPLNLPK (blue-red) and YWEMQPATFR (green).

For further validation of sample processing efficiency, 11.9 pmol of recombinant asporin was loaded into SDS-PAGE gel, chopped, digested and analysed using MRM (Figure 36). A weak, but still detectable signal has been observed.
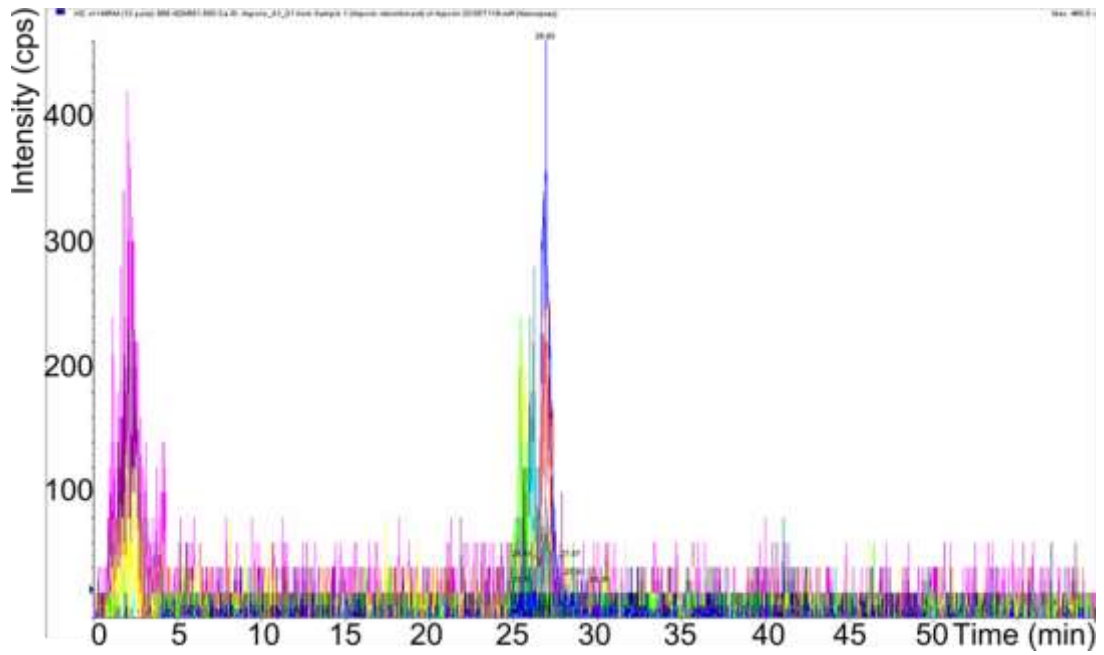
Figure 36: MRM chromatogram for 11.9 pmol of recombinant asporin. Monitored peptides were LYLSHNQLSEIPLNLPK (blue-red) and YWEMQPATFR (green) as above.

The *E.Coli* S1 digest was selected as a positive control for determination of matrix effect of complex biological sample. Processing of this lysate has been described before. The resulting peak resembles signal of standard peptides (Figure 37).
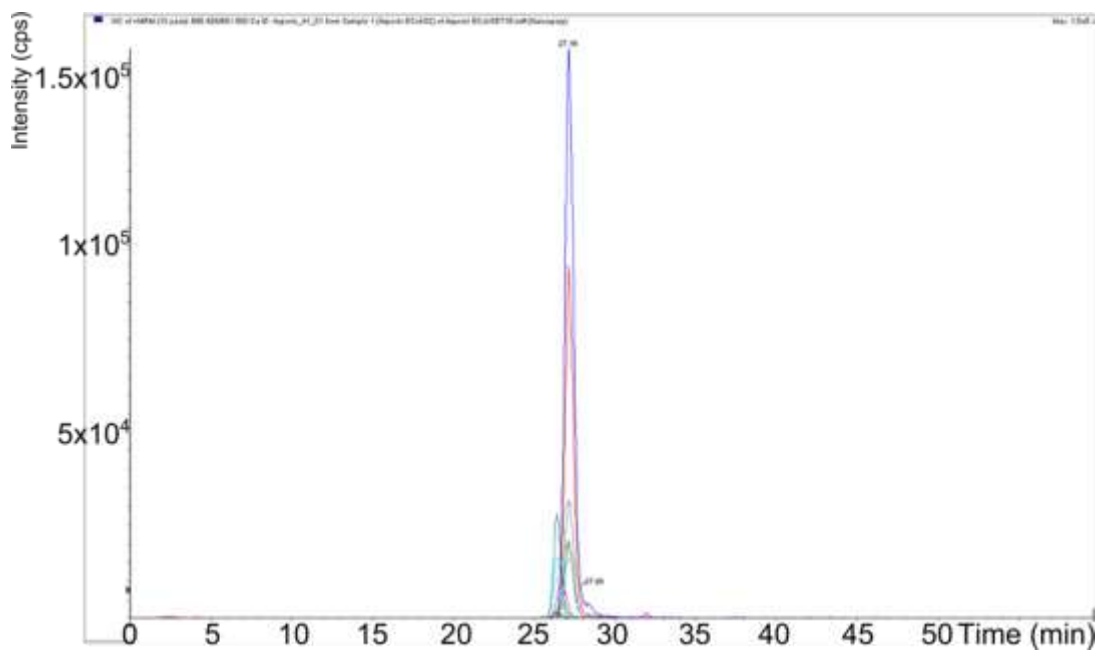
Figure 37: MRM transitions spectra for *E. Coli* S1 cellular lysate. Monitored peptides were LYLSHNQLSEIPLNLPK (blue-red) and YWEMQPATFR (green).

Those optimization results looked promising. Cellular lysate of BT549-ASPN, which has the strongest asporin expression (Figure 34), were thus analyzed by MRM. On the figure 38, there is a zoom-in to spectrum of BT549-ASPN showing a weak peak masked by noise. Asporin content in other samples was, unfortunately, below limit of detection.
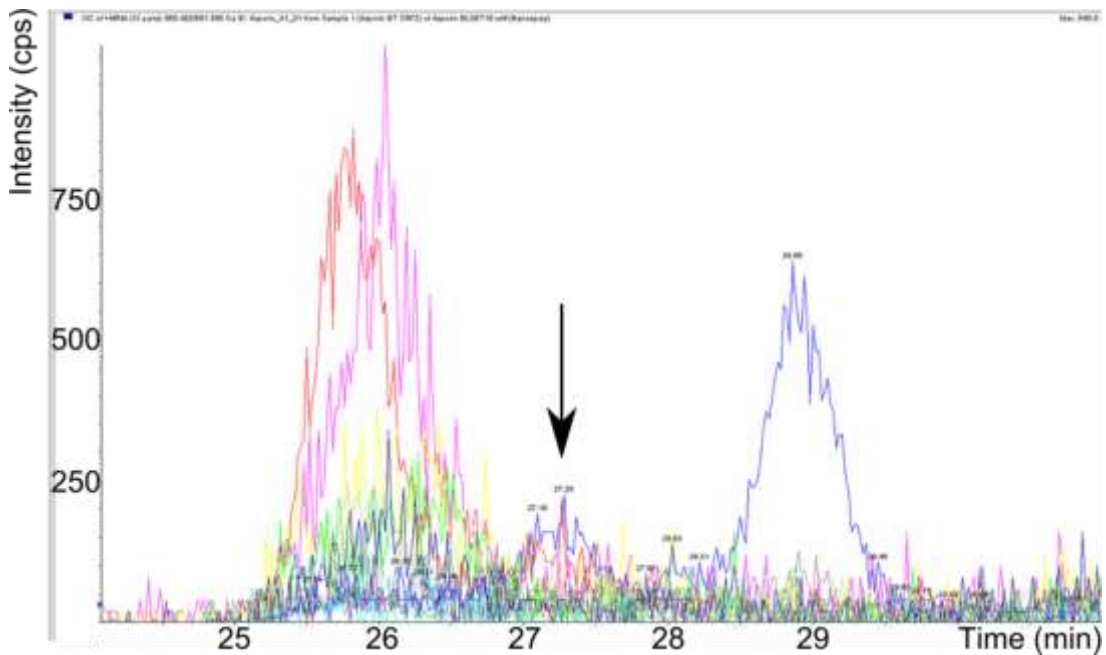
Figure 38: Zoom-in of MRM chromatogram of BT549-ASPN cell lysate. Arrow points peak of LYLSHNQLSEIPLNLPK (blue-red) peptide. This peak is unfortunately under limit of detection.

As with MS identification approach, we have failed to verify antibody specificity with sensitive MRM approach. Therefore, we have tried PRM targeted analysis performed on orbitrap with high resolution and accuracy. PRM is thus very clean transition spectra. Reduced noise thus increases sensitivity. Another advantage of PRM analysis is a good tutorial in Skyline program, which allowed setting up transitions for more peptides than in MRM analysis. Those peptides were MLDLQNNK 2+ (Peptide 1), LYLSHNQLSEIPLNLPK 3+ (Peptide 2), ISTVELEDFKR (2+ Peptide 3 and 3+ Peptide 4) and ITDIENGSLANIPR 2+ (Peptide 5). The intensities of these peptides for samples from Figure 34 are shown at Table 11. Although weak, those MS intensities correspond to intensity WB signals reported above. Signal of this strength laid under limit of quantification (see below) and asporin concentration in different cell lines couldn't been thus determined.

Table 11: PRM transition intensities (in counts per second) of asporin measured in selected cellular samples.

sample	Pep1	Pep2	Pep3	Pep4	Pep5
Hs578T 1	0	0	600	0	0
Hs578T 2	0	0	0	0	0
Gingival fibro. 1	0	500	1500	0	600
Gingival fibro. 2	0	0	0	0	0
BT549-ASPN 1	500	1500	1500	2500	2000
BT549-ASPN 2	0	1000	0	0	0
MDA-MB-231 1	0	0	400	0	0
MDA-MB-231 2	0	0	0	0	0

Successful validation of antibody specificity led to a need of orthogonal verification of results from odontogenic differentiation (Figure 39). MS and WB intensities were in good correlation in this case as well.

Sample	Pep2	Pep3	Pep4
Odonto.diff#1 1	3000	1500	3000
Odonto.diff#1 2	0	0	0
Odonto.diff#2 1	2000	1500	1500
Odonto.diff#2 2	0	0	0
Odonto.diff#3 1	3000	1000	2000
Odonto.diff#3 2	0	0	0
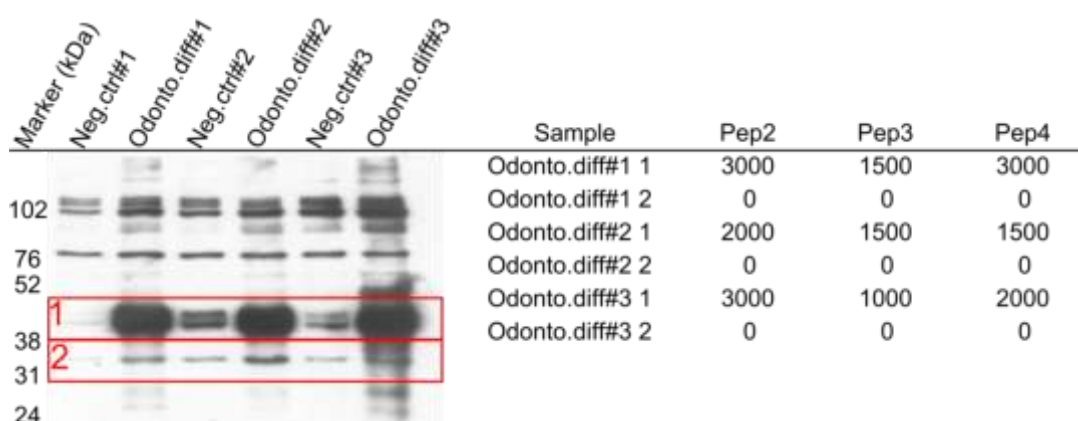
Figure 39: WB and MS analysis of samples after odontogenic differentiation. Peptides 1 and 5 weren't detected in those samples.

2.3.2.4 Linearity and sensitivity

In the figures shown previously, there is a clear evidence that WB is in case of asporin more sensitive then MS. To verify this, we have spiked 10 µg of HeLa cell lysate with asporin in concentration range 3 µmol – 3 amol and analyzed in parallel by WB and PRM MS. Limit of detection was 3 pmol for both MS and for WB (Figure 40). Limit of quantification for MS was 30 nmol contrary to WB, which is linear since limit of detection.
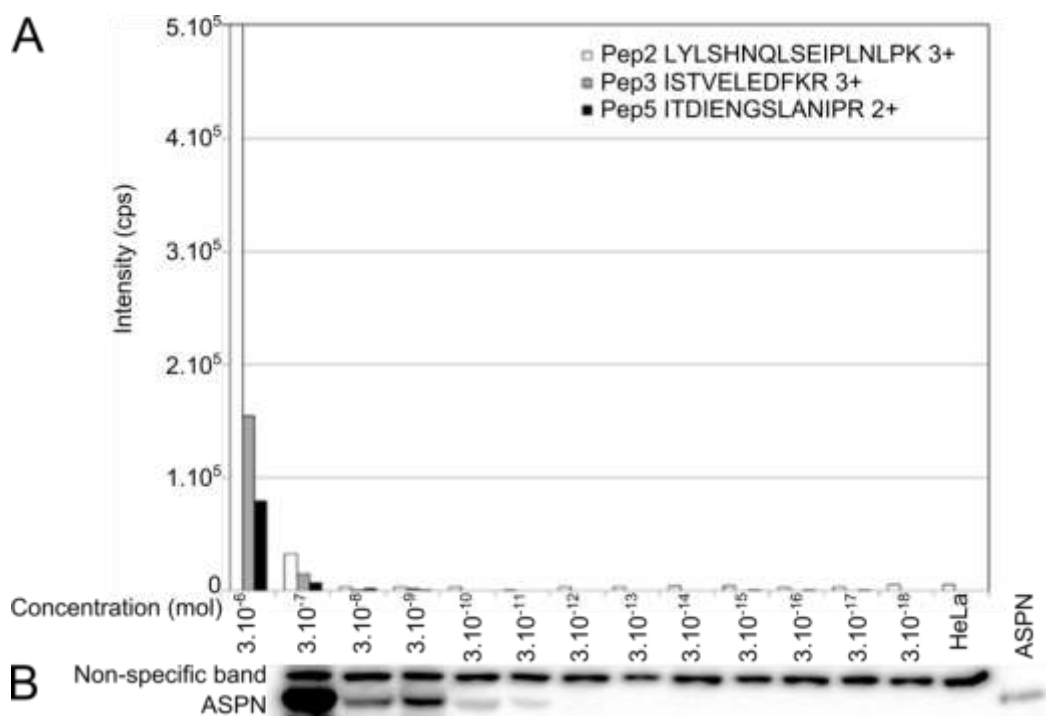
Figure 40: Comparison of sensitivity and linear range between MS (A) and WB (B). Above asporin-specific band (ASPN) is one band of non-specific binding, which serves as a loading control. In this case, WB is three orders of magnitude more sensitive.

2.3.3 Discussion

Initial tasks in this project were thus verifying its secretion in two clones of stably transfected *E. Coli* and verifying antibody specificity of Sigma-Aldrich anti-ASPN antibody. The approach of choice for those experiments was LC/MS identification strategy on MALDI-TOF and later on nESI-Orbitrap. Both analyses are routinely done in our laboratory and proven to be sensitive enough for such task (Chapter 2.2). Both strategies were successful to identify asporin only in one *E. Coli* clone, which exprimed asporin in large amounts. This was further supported by semi-preparative gel electrophoresis with 200 µg of loaded protein and favorable lysis buffer.

The fail of LC/MS approach led us to switch to more sensitive methods, which is MRM and later PRM. We were successful to verify antibody specificity, however only by highly sensitive and specific PRM approach and with border intensities. The linearity assay (Figure 40) revealed the reason. Limit of quantification for recombinant asporin spiked into HeLa cell lysate was 30 nmol. Asporin could be detected by MS only in high concentrations. This is in contrary with declared sensitivity of instrument (factory specification is 167 amol of

reserpine) or with sensitivity of different peptides analyzed in the same instrument (limit of detection for K-ras peptide LVVVGACGVGK is 15 amol [268]), or to senstivity of synthetic asporin peptides as well (Figure 35). In our concentration range, we are calculating with amount of protein, not peptide. The difference between protein and peptides is shown in MRM section, where 11.9 pmol of recombinant asporin provides poor signal detectable only thanks to low complexity of sample (Figure 36) and on the other side 10 fmol of synthetic asporin peptides LYLSHNQLSEIPLNLPK and YWEMQPATFR were analyzed with very high intensities.

LC-MS proteomic analysis is considered as very sensitive, although there are several factors influencing sensitivity of protein analysis. First one can be different ionization of peptides. This was really observed specially in PRM analysis (Figure 40) by different intensities of Skyline generated transitions. However, this doesn't explain fall of intensity of all peptides in relatively high protein concentration. Another issue can be matrix effects, which can cause ion suppression [152]. This is not the cause of low sensitivity, because we have seen low intensity signal in MRM transition of pure recombinant asporin with 11.9 pmol concentration (Figure 36) and on other side very high signal was observed in *E. Coli* lysates (Figure 37) with rich matrix background.

The most probable explanation of the poor asporin sensitivity phenomenon is a poor processing efficiency, since protocols for sample preparation are well established and proven to be effective (see Chapter 2.2). My hypothesis is poor digestion of asporin by trypsin. Trypsin is used as digestion enzyme in proteomics almost exclusively (see Chapter 1.3.1.2). Using trypsin in proteomics is advantageous, but have several limitations [113], mainly with poor coverage of portion of proteome lacking arginine and lysine in protein sequence [113]. In the literature was as well described differential efficiency in trypsin digestion of native vs. denatured or fibrillar vs. globular proteins [269] or proteins with different levels of disulfide bonds [270]. The kinetics of trypsin digestion is studied for long time, the oldest accessible study is from 1924 [271], but it is studied now as well. Differential speed of trypsin digestion in proteomics is described by Ye [272]. This literature evidence makes poor trypsin efficiency on asporin viable hypothesis. On the other side, asporin was successfully detected by LC-MS in dental cementum [252], in decorin deficient Duchenne muscular dystrophy [259], in tissue samples of degenerative mitral valve disease [260] or as possible biomarker of pancreatic cancer [264]. However, those studies don't contain the absolute asporin quantification. The issue of asporin sensitivity in proteomics thus needs a further study.

This research has been done as a part of article published in Oncotarget journal at July 7, 2016 (Appendix C).

3 SUMMARY

The thesis is divided into three relatively independent parts. Theoretical introduction is outlined as review of proteomic workflow divided into most crucial steps of current proteomic analysis – protein separation, protein identification and both top-down and bottom-up proteomics. Protein separation chapter discussed a broad scale of separation methods including the most important methods such as SDS-PAGE, affinity purification or liquid chromatography. Protein identification methods chapter offer a brief review of protein identification strategies including antibody based methods (for example immunoassays, western blot or immunomicroscopic methods), mass spectrometry and protein sequencing methods. The most popular discovery approach in current proteomics is a combination of liquid chromatography as a separation step with mass spectrometry for protein identification. This approach is divided on top-down or bottom-up proteomics in dependence if whole proteins or digested peptides are analyzed. In this thesis, I have used almost exclusively bottom-up approach and Chapter 1.3 is thus more focused on this method. The theoretical part is closed by overview of less often mass spectrometry approaches in proteomics.

The workflow reviewed in theoretical part can be used in a broad scale of current science's challenges as is shown in experimental part, where two different projects are solved. In the first one, proteomics is used to discover cellular response to anticancer drugs. In the second project, advanced proteomic methods were used to validate antibody specificity and confirm presence and quantity of potential cancer biomarker.

It is beneficial to use large-scale analysis in determination of cellular response of the drug. Among other large-scale experiments, such as genomic or transcriptomic, analysis of proteome is closest to actual place of drug effect. Unfortunately, compared to DNA or mRNA, it is not possible to amplify proteins yet. Proteomic analysis has to be very sensitive to be effective. This has been shown on proteomic profile of cellular response to three common, platinum-based, anti-cancer drugs. If MS approach based on "traditional" low resolution instruments was used, number of identified proteins didn't exceed 1000 and list of proteins significantly changed after treatment was very brief. Unfortunately, in such brief lists were repeatedly present few proteins uniformly changed after treatment with broad scale of different drugs (data not shown). The introduction of state-of-art high resolution instrument led to significant increase of identified proteins followed by similar increase in count of significantly changed proteins. This increment showed the best results for OxaPt,

where it allowed reconstructing main response pathways. The changes in case of CisPt and CarboPt were less distinctive.

There were five main areas of CCRF-CEM cellular response to OxaPt treatment. Two of them are very closely related and it is nucleolar and ribosomal stress. Nucleolar and ribosomal stress were for OxaPt confirmed by WB and IF microscopy. Those effects were also observed by other research groups as nucleolar shrinkage after OxaPt treatment [220,273]. In those studies, nucleolar shrinkage was given in correlation with neuropathy. Relation between OxaPt and ribosomes has been reported as two-sided. On one side OxaPt affects ribosomes [217], on the other side ribosomes are involved in OxaPt resistance [218].

The other two interesting groups of proteins are proteins related to centromere/involved in G2/M stop or proteins involved in DNA damage response. Contrary to most commonly mentioned effect of platinum drugs, covalent bond of platinum to DNA, those protein groups were less abundant and DNA damage response group wasn't even discovered by common bioinformatic tools. On the other side, DNA damage response is known to trigger nucleolar stress [244].

The last group of regulated proteins was secretory proteins which are often involved in innate immunity or in response to oxidative stress. Oxidative stress is one of often reported responses to platinum drugs in general. The role of innate immunity in OxaPt treatment is described in vivo [274], but it is still unclear in case of cell line. However, one from those proteins, Apolipoprotein A1, has shown itself as potential biomarker of effective OxaPt therapy [214].

Whereas in the proteomic profiling experiment antibodies were used to verify MS result, in the second project of this thesis, MS was used to verify antibody specificity. This is useful particularly in case, when more bands are observed and is necessary to decide which band is specific. In this project, the level of potential cancer biomarker, asporin, and its role in cell biology was determined.

Proteomic analysis of asporin was done in two branches. In the first branch, I have tested transfected *E. Coli* lysate for proper asporin expression. This task has been done relatively easily, because in the second analyzed sample bacteria expressed asporin in a big amount. This sample served as a positive control for the rest of experiments.

The second branch of experiments consisted of qualitative and quantitative analysis of asporin in cell lines – the particular antibody verification. This task was particularly

challenging, because standard proteomic method as MS protein identification or MRM failed to detect asporin due to poor sensitivity which was specific only to asporin. The task of antibody verification was fulfilled using orbitrap based PRM analysis, however again with unsatisfactory sensitivity to asporin. The main issue of proteomic determination of asporin has been found in sample processing. The 10 fmoles of synthetic asporin peptides provided very strong signal on MRM, even compared to digest from 12 pmol recombinant protein. The reason of such a discrepancy is still unknown and is a object for future research.

In those two projects, almost all main proteomic methods discussed in theoretical part were used to fill the research objectives. Proteomic profiling of cancer cell lines treated by platinum drugs used SILAC metabolic labelling, SDS-PAGE separation, LC-MS detection using bottom-up approach and result verification using western blot and immunofluorescence microscopy. Asporin validation project was solved using regular bottom-up LC-MS approach with successive movement to more sensitive methods, like MRM or PRM targeted approaches.

4 SOUHRN

Tato dizertační práce je rozdělena na tři relativně samostatné části. Teoretická část je koncipována jako přehled současných proteomických metod rozdělených na základní kroky současné proteomické analýzy – separaci a identifikaci proteinů a top-down a bottom-up proteomiku. V kapitole o proteinové separaci je diskutována široká škála separačních metod včetně těch nejpoužívanějších, jako je SDS-PAGE elektroforéza, afinitní purifikace nebo kapalinová chromatografie. Stručný přehled metod identifikace proteinů obsahuje metody využívající protilárky (například imunostanovení (ELISA), western blot nebo imunomikroskopické metody) hmotnostní spektrometrii a metody sekvenace proteinů. Nejoblíbenější přístup v současné proteomice je kombinace kapalinové chromatografie a hmotnostní spektrometrie. Tento přístup se dále dělí na top-down a bottom-up proteomiku v závislosti na tom, jestli se měří celé proteiny, nebo jejich tryptické digesty. V této práci byl využíván zejména přístup bottom-up, proto je kapitola 1.3 zaměřena zejména na popis tohoto přístupu. Teoretická část je zakončena přehledem méně častých hmotnostně spektrometrických přístupů.

Postupy popsané v teoretické části jsou využívány při řešení celé řady úkolů v současné vědě. To je mimo jiné ukázáno v experimentální části, kde jsou řešeny dva projekty. V prvním z nich je proteomika využita k určení buněčné odpovědi na protinádorová léčiva. V druhém projektu byly použity pokročilé proteomické metody k validaci specifity protilátky a potvrzení přítomnosti a množství potenciálního nádorového biomarkeru.

Při určování buněčné odpovědi na léčivo je výhodné používat tzv. large-scale analýzu. Typickými příklady tohoto druhu analýzy jsou genomické nebo transkriptomické experimenty, ale také proteomika, která mapuje expresní změny na úrovni proteinů. Na rozdíl od DNA nebo RNA není možné v současné době proteiny amplifikovat. Z toho důvodu musí být proteomická analýza velmi citlivá. To je mimo jiné ukázáno na příkladu proteomického profilování buněčné odpovědi na tři běžná platinová cytostatika. Pokud byla použita identifikace proteinů pomocí „tradičních" hmotnostních spektrometrů, počet identifikovaných proteinů nepřesáhl tisíc a seznam signifikantně změněných proteinů byl velmi strohý. Bohužel se v tomto seznamu často opakovaly stejné proteiny nezávislé na použitém léčivu, což může souviset s jejich abundancí a relativní jednoduchostí identifikace. Zavedení nejnovějšího vysokorozlišovacího hmotnostního spektrometru na principu orbitrapu vedlo k výraznému zvýšení počtu identifikovaných proteinů a obdobnému zvýšení signifikantně regulovaných proteinů. Toto zvýšení se nejlépe projevilo u oxaliplatiny, kde

umožnilo analyzovat hlavní dráhy buněčné odpovědi. Změny pozorované u cisplatiny a karboplatiny byly méně výrazné.

Buněčnou odpověď linie CCRF-CEM na ošetření oxaliplatinou lze rozdělit na pět hlavních oblastí a to DNA poškození, nukleolární a ribozomální stres, alterace buněčného cyklu a změny na úrovni sekretomu. Nukleolární a ribozomální stres jsou příbuzné a související procesy, přičemž tyto výsledky byly ověřeny western blotem a imunofluorescenční mikroskopií. Nukleolární a ribozomální stres byl pozorován mikroskopicky jako zmenšení jadérka po ošetření oxaliplatinou také jinými vědeckými týmy [220,273]. V těchto studiích bylo zmenšení jadérka dáváno do souvislosti s nežádoucí neuropatií, která je indukována oxaliplatinou. Vztah oxaliplatiny a ribozomů je oboustranný. Na jedné straně oxaliplatina ovlivňuje produkci ribozomů [217], na straně druhé jsou ribozomy zapojeny do rezistence na oxaliplatinu [218].

Další dvě zajímavé skupiny jsou skupiny proteinů zapojených v zastavení mitózy v G2/M fázi a proteiny účastnící se v opravě poškození DNA. Poškození DNA je přitom nejčastěji uváděným efektem platinových léčiv. Je proto zajímavé, že dráha poškození DNA byla nejméně zastoupená a nebyla odhalena běžnými bioinformatickými nástroji. Na druhou stranu, odpověď na poškození DNA je jedním ze spouštěčů nukleolárního stresu [244].

Poslední skupinou regulovaných proteinů jsou sekretované proteiny často se podílející na vrozené imunitě, mezibuněčných interakcích nebo při odpovědi na oxidativní stres. Oxidativní stres je jedním z často popisovaných efektů platinových léčiv. Role vrozené imunity je při léčbě nádorů oxaliplatinou popsána in vivo [274], v buněčné linii je tato role nejasná. Jeden z nadregulovaných sekretovaných proteinů, apolipoprotein A1, je potenciálním biomarkerem efektivní terapie oxaliplatinou [214].

Zatímco v profilovacím projektu byly protilátkové metody využity k ověření MS výsledků, v druhé části experimentální části byla využita MS k ověření specificity protilátky. To je důležité zejména v případě, kdy se protilátka váže na více proteinů, jak můžeme často na westernblottu pozorovat pozitivitou více bandů a je potřeba rozlišit, který band, protein, je specifický. V tomto případě byla ověřována koncentrace a role potencionálního nádorového biomarkeru – asporinu.

Proteomická analýza asporinu byla prováděna ve dvou experimentálních schématech. První z nich byla validace exprese asporinu v lyzátu transfekované baterie *E. Coli.* Tento projekt byl úspěšně splněn, jelikož již druhý testovaný klon linie exprimoval asporin v množství

bezpečně dostačujícím pro identifikaci pomocí hmotnostní spektrometrie. Tyto vzorky byly v dalších analýzách využity jako pozitivní kontrola.

Druhá větev experimentů spočívala v kvalitativní a kvantitativní analýze asporinu v buněčných liniích – tedy vlastní validaci protilátky. Tento úkol se ukázal být velkou výzvou, jelikož běžné proteomické postupy jako MS identifikace, nebo MRM analýza měly vůči asporinu velmi malou citlivost. Úkol validovat specificitu protilátky byl splněn až použitím metody PRM na hmotnostním spektrometru s s vysokým rozlišením. Citlivost však nebyla zcela uspokojující ani v tomto případě. Hlavní příčina nízké citlivosti byla nalezena v přípravě vzorků, protože 10 fmol roztok syntetických peptidů asporinu poskytoval při MRM velmi silný signál, zatímco vzorek připravený z 12 pmol rekombinantního proteinu signál spíše slabý. Důvod tohoto rozdílu bude předmětem dalšího výzkumu.

V těchto dvou projektech byla využita většina metod popisovaných v teoretické části. V proteomickém profilování byly využity metody metabolického značení SILAC, SDS-PAGE, LC-MS detekce s využitím bottom-up proteomiky a výsledky byly validovány metodami western blot a imunofluorescenční mikroskopie. Projekt validace asporinu byl řešen zpočátku s využitím běžného bottom-up LC-MS proteomického přístupu s postupným přesunem k citlivějším metodám jako jsou MRM nebo PRM cílené metody.

5 REFERENCES

[1]   L. Dayon, M. Kussmann, Proteomics of human plasma: A critical comparison of analytical workflows in terms of effort, throughput and outcome, EuPA Open Proteomics. 1 (2013) 8–16. doi:10.1016/j.euprot.2013.08.001.

[2]   A. Michalski, E. Damoc, O. Lange, E. Denisov, D. Nolting, M. Müller, R. Viner, J. Schwartz, P. Remes, M. Belford, J.-J. Dunyach, J. Cox, S. Horning, M. Mann, A. Makarov, Ultra High Resolution Linear Ion Trap Orbitrap Mass Spectrometer (Orbitrap Elite) Facilitates Top Down LC MS/MS and Versatile Peptide Fragmentation Modes, Mol. Cell. Proteomics MCP. 11 (2012). doi:10.1074/mcp.O111.013698.

[3]   C. Arndt, S. Koristka, H. Bartsch, M. Bachmann, Native polyacrylamide gels, Methods Mol. Biol. Clifton NJ. 869 (2012) 49–53. doi:10.1007/978-1-61779-821-4_5.

[4]   B. Bjellqvist, K. Ek, P. Giorgio Righetti, E. Gianazza, A. Görg, R. Westermeier, W. Postel, Isoelectric focusing in immobilized pH gradients: Principle, methodology and some applications, J. Biochem. Biophys. Methods. 6 (1982) 317–339. doi:10.1016/0165-022X(82)90013-6.

[5]   Z. Zhu, J.J. Lu, S. Liu, Protein Separation by Capillary Gel Electrophoresis: A Review, Anal. Chim. Acta. 709 (2012) 21–31. doi:10.1016/j.aca.2011.10.022.

[6]   U.K. Laemmli, Cleavage of structural proteins during the assembly of the head of bacteriophage T4, Nature. 227 (1970) 680–685.

[7]   J.A. Reynolds, C. Tanford, The Gross Conformation of Protein-Sodium Dodecyl Sulfate Complexes, J. Biol. Chem. 245 (1970) 5161–5165.

[8]   J.Y. Kim, S.H. Ahn, S.T. Kang, B.J. Yoon, Electrophoretic mobility equation for protein with molecular shape and charge multipole effects, J. Colloid Interface Sci. 299 (2006) 486–492. doi:10.1016/j.jcis.2006.02.003.

[9]   P.G. Righetti, C. Simó, R. Sebastiano, A. Citterio, Carrier ampholytes for IEF, on their fortieth anniversary (1967-2007), brought to trial in court: the verdict, Electrophoresis. 28 (2007) 3799–3810. doi:10.1002/elps.200700232.

[10]   P.H. O'Farrell, High resolution two-dimensional electrophoresis of proteins, J. Biol. Chem. 250 (1975) 4007–4021.

[11]   M. Unlu, M.E. Morgan, J.S. Minden, Difference gel electrophoresis: A single gel method for detecting changes in protein extracts, Electrophoresis. 18 (1997) 2071–2077. doi:10.1002/elps.1150181133.

[12]   J. Tyleckova, R. Hrabakova, K. Mairychova, P. Halada, L. Radova, P. Dzubak, M. Hajduch, S.J. Gadher, H. Kovarova, Cancer Cell Response to Anthracyclines Effects: Mysteries of the Hidden Proteins Associated with These Drugs, Int. J. Mol. Sci. 13 (2012) 15536–15564. doi:10.3390/ijms131215536.

[13]   A. Champagne, H. Rischer, K.-M. Oksman-Caldentey, M. Boutry, In-depth proteome mining of cultured Catharanthus roseus cells identifies candidate proteins involved in the synthesis and transport of secondary metabolites, Proteomics. 12 (2012) 3536–3547. doi:10.1002/pmic.201200218.

[14]   J. Petrak, R. Ivanek, O. Toman, R. Cmejla, J. Cmejlova, D. Vyoral, J. Zivny, C.D. Vulpe, Déjà vu in proteomics. A hit parade of repeatedly identified differentially expressed proteins, Proteomics. 8 (2008) 1744–1749. doi:10.1002/pmic.200700919.

[15]   D.C. Simpson, R.D. Smith, Combining capillary electrophoresis with mass spectrometry for applications in proteomics, ELECTROPHORESIS. 26 (2005) 1291–1305. doi:10.1002/elps.200410132.

[16]   G. Mitulović, New HPLC Techniques for Proteomics Analysis: A Short Overview of Latest Developments, J. Liq. Chromatogr. Relat. Technol. 38 (2015) 390–403. doi:10.1080/10826076.2014.941266.

[17]   J.F. Banks Jr., High-sensitivity peptide mapping using packed-capillary liquid chromatography and electrospray ionization mass spectrometry, J. Chromatogr. A. 743 (1996) 99–104. doi:10.1016/0021-9673(96)00319-6.

[18]   M. Vollmer, T. van de Goor, HPLC-Chip/MS Technology in Proteomic Profiling, in: R.S. Foote, J.W. Lee (Eds.), Micro Nano Technol. Bioanal., Humana Press, 2009: pp. 3–15. doi:10.1007/978-1-59745-483-4_1.

[19]   D.L. Swaney, C.D. Wenger, J.J. Coon, The value of using multiple proteases for large-scale mass spectrometry-based proteomics, J. Proteome Res. 9 (2010) 1323–1329. doi:10.1021/pr900863u.

[20]   H. Wang, Y. Yang, Y. Li, B. Bai, X. Wang, H. Tan, T. Liu, T.G. Beach, J. Peng, Z. Wu, Systematic Optimization of Long Gradient Chromatography Mass Spectrometry for Deep Analysis of Brain Proteome, J. Proteome Res. 14 (2015) 829–838. doi:10.1021/pr500882h.

[21]   Differences Between Using Acetonitrile and Methanol for Reverse Phase Chromatography : SHIMADZU (Shimadzu Corporation), Shimadzu Corp. (n.d.). http://www.shimadzu.com/an/hplc/support/lib/lctalk/35/35lab.html (accessed May 27, 2016).

[22]   S. Fekete, J.-L. Veuthey, D. Guillarme, New trends in reversed-phase liquid chromatographic separations of therapeutic peptides and proteins: Theory and applications, J. Pharm. Biomed. Anal. 69 (2012) 9–27. doi:10.1016/j.jpba.2012.03.024.

[23]   J. Rozenbrand, W.P. van Bennekom, Silica-based and organic monolithic capillary columns for LC: Recent trends in proteomics, J. Sep. Sci. 34 (2011) 1934–1944. doi:10.1002/jssc.201100294.

[24]   R. Hayes, A. Ahmed, T. Edge, H. Zhang, Core–shell particles: Preparation, fundamentals and applications in high performance liquid chromatography, J. Chromatogr. A. 1357 (2014) 36–52. doi:10.1016/j.chroma.2014.05.010.

[25]   D.A. Wolters, M.P. Washburn, J.R. Yates, An automated multidimensional protein identification technology for shotgun proteomics, Anal. Chem. 73 (2001) 5683–5690.

[26]   M.J. Edelmann, Strong cation exchange chromatography in analysis of posttranslational modifications: innovations and perspectives, J. Biomed. Biotechnol. 2011 (2011) 936508. doi:10.1155/2011/936508.

[27]   E. Mostovenko, C. Hassan, J. Rattke, A.M. Deelder, P.A. van Veelen, M. Palmblad, Comparison of peptide and protein fractionation methods in proteomics, EuPA Open Proteomics. 1 (2013) 30–37. doi:10.1016/j.euprot.2013.09.001.

[28]   D. Bensaddek, A. Nicolas, A.I. Lamond, Evaluating the use of HILIC in large-scale, multi dimensional proteomics: Horses for courses?, Int. J. Mass Spectrom. 391 (2015) 105–114. doi:10.1016/j.ijms.2015.07.029.

[29]   O. Hernandez-Hernandez, J.E. Quintanilla-Lopez, R. Lebron-Aguilar, M.L. Sanz, F.J. Moreno, Characterization of post-translationally modified peptides by hydrophilic interaction and reverse phase liquid chromatography coupled to quadrupole-time-of-flight mass spectrometry, J. Chromatogr. A. 1428 (2016) 202–211. doi:10.1016/j.chroma.2015.07.096.

[30]   G. Pohlentz, K. Marx, M. Mormann, Characterization of Protein N-Glycosylation by Analysis of ZIC-HILIC-Enriched Intact Proteolytic Glycopeptides, Methods Mol. Biol. Clifton NJ. 1394 (2016) 163–179. doi:10.1007/978-1-4939-3341-9_12.

[31]   Q. Quan, S.S.W. Szeto, H.C.H. Law, Z. Zhang, Y. Wang, I.K. Chu, Fully automated multidimensional reversed-phase liquid chromatography with tandem anion/cation exchange columns for simultaneous global endogenous tyrosine nitration detection, integral membrane protein characterization, and quantitative proteomics mapping in cerebral infarcts, Anal. Chem. 87 (2015) 10015–10024. doi:10.1021/acs.analchem.5b02619.

[32] J. Fíla, D. Honys, Enrichment techniques employed in phosphoproteomics, Amino Acids. 43 (2012) 1025–1047. doi:10.1007/s00726-011-1111-z.

[33] L. Beltran, P.R. Cutillas, Advances in phosphopeptide enrichment techniques for phosphoproteomics, Amino Acids. 43 (2012) 1009–1024. doi:10.1007/s00726-012-1288-9.

[34] X.-S. Yue, A.B. Hummon, Mass spectrometry-based phosphoproteomics in cancer research, Front. Biol. 7 (2012) 566–586. doi:10.1007/s11515-012-2022-4.

[35] T. Sivaraman, T.K. Kumar, G. Jayaraman, C. Yu, The mechanism of 2,2,2-trichloroacetic acid-induced protein precipitation, J. Protein Chem. 16 (1997) 291–297.

[36] M. Zellner, W. Winkler, H. Hayden, M. Diestinger, M. Eliasen, B. Gesslbauer, I. Miller, M. Chang, A. Kungl, E. Roth, R. Oehler, Quantitative validation of different protein precipitation methods in proteome analysis of blood platelets, ELECTROPHORESIS. 26 (2005) 2481–2489. doi:10.1002/elps.200410262.

[37] T.P. King, Separation of proteins by ammonium sulfate gradient solubilization, Biochemistry (Mosc.). 11 (1972) 367–371. doi:10.1021/bi00753a010.

[38] G. Köhler, C. Milstein, Continuous cultures of fused cells secreting antibody of predefined specificity, Nature. 256 (1975) 495–497.

[39] G.L. Perkins, E.D. Slater, G.K. Sanders, J.G. Prichard, Serum tumor markers, Am. Fam. Physician. 68 (2003) 1075–1082.

[40] H. Koprowski, M. Herlyn, Z. Steplewski, H.F. Sears, Specific antigen in serum of patients with colon carcinoma, Science. 212 (1981) 53–55.

[41] H.-W. Sun, R.-F. Wang, P. Yan, C.-L. Zhang, P. Hao, H. Ma, X.-Q. Chen, Radioactive iodine labeling of monoclonal antibody against Hsp90α and its use in diagnostic imaging in prostate cancer xenograft model, Eur. Rev. Med. Pharmacol. Sci. 19 (2015) 835–843.

[42] L.T. Goldstein, N.R. Klinman, L.A. Manson, A Microtest Radioimmunoassay for Noncytotoxic Tumor-Specific Antibody to Cell-Surface Antigens, J. Natl. Cancer Inst. 51 (1973) 1713–1715. doi:10.1093/jnci/51.5.1713.

[43] T.-B. Xin, H. Chen, Z. Lin, S.-X. Liang, J.-M. Lin, A secondary antibody format chemiluminescence immunoassay for the determination of estradiol in human serum, Talanta. 82 (2010) 1472–1477. doi:10.1016/j.talanta.2010.07.023.

[44] R.S. Yalow, S.A. Berson, IMMUNOASSAY OF ENDOGENOUS PLASMA INSULIN IN MAN, J. Clin. Invest. 39 (1960) 1157–1175.

[45] L. Cox, B. Williams, S. Sicherer, J. Oppenheimer, L. Sher, R. Hamilton, D. Golden, American College of Allergy, Asthma and Immunology Test Task Force, American Academy of Allergy, Asthma and Immunology Specific IgE Test Task Force, Pearls and pitfalls of allergy diagnostic testing: report from the American College of Allergy, Asthma and Immunology/American Academy of Allergy, Asthma and Immunology Specific IgE Test Task Force, Ann. Allergy Asthma Immunol. Off. Publ. Am. Coll. Allergy Asthma Immunol. 101 (2008) 580–592.

[46] S. Ta, K. Jn, M. Je, J. Im, F. F, R. Ea, Y. N, Prostate specific antigen in the diagnosis and treatment of adenocarcinoma of the prostate. II. Radical prostatectomy treated patients., J. Urol. 141 (1989) 1076–1083.

[47] S.D. Gan, K.R. Patel, Enzyme Immunoassay and Enzyme-Linked Immunosorbent Assay, J. Invest. Dermatol. 133 (2013) 1–3. doi:10.1038/jid.2013.287.

[48] MitoSciences: Sandwich ELISA Kits - Microplates, (n.d.). http://www.mitosciences.com/microplate-sandwich-elisa-kits.html (accessed August 23, 2016).

[49] H. Towbin, T. Staehelin, J. Gordon, Electrophoretic transfer of proteins from polyacrylamide gels to nitrocellulose sheets: procedure and some applications, Proc. Natl. Acad. Sci. U. S. A. 76 (1979) 4350–4354.

[50] S.C. Taylor, T. Berkelman, G. Yadav, M. Hammond, A Defined Methodology for Reliable Quantification of Western Blot Data, Mol. Biotechnol. 55 (2013) 217–226. doi:10.1007/s12033-013-9672-6.

[51] Western blot, Wikipedia Free Encycl. (2016). https://en.wikipedia.org/w/index.php?title=Western_blot&oldid=735048447 (accessed August 23, 2016).

[52] Transferring separated components in gel electrophoresis via nylon membrane, n.d. http://www.google.com/patents/US4455370 (accessed April 3, 2016).

[53] N. LeGendre, Immobilon-P transfer membrane: applications and utility in protein biochemical analysis, BioTechniques. 9 (1990) 788–805.

[54] K.C. Reed, D.A. Mann, Rapid transfer of DNA from agarose gels to nylon membranes, Nucleic Acids Res. 13 (1985) 7207–7221. doi:10.1093/nar/13.20.7207.

[55] A. Elbaggari, K. McDonald, A. Alburo, C. Martin, Imaging of Chemiluminescent Western Blots: Comparison of Digital Imaging and X-ray Film, (n.d.).

[56] J.C. Gingrich, D.R. Davis, Q. Nguyen, Multiplex detection and quantitation of proteins on western blots using fluorescent probes, BioTechniques. 29 (2000) 636–642.

[57] J. Park, M. Mabuchi, A. Sharma, Multiplexed Fluorescent Immunodetection Using Low Autofluorescence Immobilon®-FL Membrane, Methods Mol. Biol. Clifton NJ. 1314 (2015) 195–205. doi:10.1007/978-1-4939-2718-0_22.

[58] M. Zellner, R. Babeluk, M. Diestinger, P. Pirchegger, S. Skeledzic, R. Oehler, Fluorescence-based Western blotting for quantitation of protein biomarkers in clinical samples, ELECTROPHORESIS. 29 (2008) 3621–3627. doi:10.1002/elps.200700935.

[59] A. Vigelsø, R. Dybboe, C.N. Hansen, F. Dela, J.W. Helge, A.G. Grau, GAPDH and β-actin protein decreases with aging, making Stain-Free technology a superior loading control in Western blotting of human skeletal muscle, J. Appl. Physiol. 118 (2015) 386–394. doi:10.1152/japplphysiol.00840.2014.

[60] I. Romero-Calvo, B. Ocón, P. Martínez-Moya, M.D. Suárez, A. Zarzuelo, O. Martínez-Augustin, F.S. de Medina, Reversible Ponceau staining as a loading control alternative to actin in Western blots, Anal. Biochem. 401 (2010) 318–320. doi:10.1016/j.ab.2010.02.036.

[61] M. Hagiwara, K.-I. Kobayashi, T. Tadokoro, Y. Yamamoto, Application of SYPRO Ruby- and Flamingo-stained polyacrylamide gels to Western blot analysis, Anal. Biochem. 397 (2010) 262–264.

[62] C. Welinder, L. Ekblad, Coomassie staining as loading control in Western blot analysis, J. Proteome Res. 10 (2011) 1416–1419. doi:10.1021/pr1011476.

[63] A. Gürtler, N. Kunz, M. Gomolka, S. Hornhardt, A.A. Friedl, K. McDonald, J.E. Kohn, A. Posch, Stain-Free technology as a normalization tool in Western blot analysis, Anal. Biochem. 433 (2013) 105–111. doi:10.1016/j.ab.2012.10.010.

[64] A.H. Coons, H.J. Creech, R.N. Jones, Immunological Properties of an Antibody Containing a Fluorescent Group., Exp. Biol. Med. 47 (1941) 200–202. doi:10.3181/00379727-47-13084P.

[65] P. Grell, P. Fabian, M. Khoylou, L. Radova, O. Slaby, R. Hrstka, R. Vyzula, M. Hajduch, M. Svoboda, Akt expression and compartmentalization in prediction of clinical outcome in HER2-positive metastatic breast cancer patients treated with trastuzumab, Int. J. Oncol. 41 (2012) 1204–1212. doi:10.3892/ijo.2012.1576.

[66] M. Mistrik, E. Vesela, T. Furst, H. Hanzlikova, I. Frydrych, J. Gursky, D. Majera, J. Bartek, Cells and Stripes: A novel quantitative photo-manipulation technique, Sci. Rep. 6 (2016) 19567. doi:10.1038/srep19567.

[67] B.F. Cravatt, G.M. Simon, J.R. Yates, The biological impact of mass-spectrometry-based proteomics, Nature. 450 (2007) 991–1000. doi:10.1038/nature06525.

[68] M. Ramström, J. Bergquist, Miniaturized proteomics and peptidomics using capillary liquid separation and high resolution mass spectrometry, FEBS Lett. 567 (2004) 92–95. doi:10.1016/j.febslet.2004.04.074.

[69] D.J. Pappin, P. Hojrup, A.J. Bleasby, Rapid identification of proteins by peptide-mass fingerprinting, Curr. Biol. CB. 3 (1993) 327–332.

[70] W. Lotz, Electron-impact ionization cross-sections and ionization rate coefficients for atoms and ions from hydrogen to calcium, Z. Für Phys. 216 (1968) 241–247. doi:10.1007/BF01392963.

[71] M.S.B. Munson, F.H. Field, Chemical Ionization Mass Spectrometry. I. General Introduction, J. Am. Chem. Soc. 88 (1966) 2621–2630. doi:10.1021/ja00964a001.

[72] M. Karas, D. Bachmann, F. Hillenkamp, Influence of the wavelength in high-irradiance ultraviolet laser desorption mass spectrometry of organic molecules, Anal. Chem. 57 (1985) 2935–2939. doi:10.1021/ac00291a042.

[73] J.B. Fenn, M. Mann, C.K. Meng, S.F. Wong, C.M. Whitehouse, Electrospray ionization for mass spectrometry of large biomolecules, Science. 246 (1989) 64–71.

[74] K. Tanaka, H. Waki, Y. Ido, S. Akita, Y. Yoshida, T. Yoshida, T. Matsuo, Protein and polymer analyses up to m/z 100 000 by laser ionization time-of-flight mass spectrometry, Rapid Commun. Mass Spectrom. 2 (1988) 151–153. doi:10.1002/rcm.1290020802.

[75] R. Zenobi, R. Knochenmuss, Ion formation in MALDI mass spectrometry, Mass Spectrom. Rev. 17 (1998) 337–366. doi:10.1002/(SICI)1098-2787(1998)17:5<337::AID-MAS2>3.0.CO;2-S.

[76] P. Kebarle, U.H. Verkerk, Electrospray: from ions in solution to ions in the gas phase, what we know now, Mass Spectrom. Rev. 28 (2009) 898–917. doi:10.1002/mas.20247.

[77] V.M. Bautista-de Lucio, M. Ortiz-Casas, L. Antonio Bautista-Hernndez, N. Luz, C. Gaona- Jurez, ngel Gustavo, D.A. Frausto-del Ro, H. Meja-Lpez, Diagnostics Methods in Ocular Infections–From Microorganism Culture to Molecular Methods, in: I. Chaudhry (Ed.), Common Eye Infect., InTech, 2013. http://www.intechopen.com/books/common-eye-infections/diagnostics-methods-in-ocular-infections-from-microorganism-culture-to-molecular-methods (accessed August 23, 2016).

[78] S.M. Rivera, P. Christou, R. Canela-Garayoa, Identification of carotenoids using mass spectrometry, Mass Spectrom. Rev. 33 (2014) 353–372. doi:10.1002/mas.21390.

[79] D. Jain, P.K. Basniwal, Forced degradation and impurity profiling: Recent trends in analytical perspectives, J. Pharm. Biomed. Anal. 86 (2013) 11–35. doi:10.1016/j.jpba.2013.07.013.

[80] S.H. Lee, M.V. Williams, R.N. DuBois, I.A. Blair, Targeted lipidomics using electron capture atmospheric pressure chemical ionization mass spectrometry, Rapid Commun. Mass Spectrom. RCM. 17 (2003) 2168–2176. doi:10.1002/rcm.1170.

[81] A. Raffaelli, A. Saba, Atmospheric pressure photoionization mass spectrometry, Mass Spectrom. Rev. 22 (2003) 318–331. doi:10.1002/mas.10060.

[82] J. Tillner, J.S. McKenzie, E.A. Jones, A.V. Speller, J.L. Walsh, K.A. Veselkov, J. Bunch, Z. Takats, I.S. Gilmore, Investigation of the impact of DESI sprayer geometry on its performance in imaging of biological tissue, Anal. Chem. (2016). doi:10.1021/acs.analchem.6b00345.

[83] S. Okutan, H.S. Hansen, C. Janfelt, A simplified approach to cryo-sectioning of mice for whole-body imaging of drugs and metabolites with desorption electrospray ionization mass spectrometry imaging, Proteomics. (2016). doi:10.1002/pmic.201500422.

[84] C.-C. Hsu, P.-T. Chou, R.N. Zare, Imaging of Proteins in Tissue Samples Using Nanospray Desorption Electrospray Ionization Mass Spectrometry, Anal. Chem. 87 (2015) 11171–11175. doi:10.1021/acs.analchem.5b03389.

[85] W. Stephens, A Pulsed Mass Spectrometer with Time Dispersion, Phys. Rev. 69 (1946) 691–691.

[86] B. Mamyrin, V. Karataev, D. Shmikk, V. Zagulin, Mass-Reflectron a New Nonmagnetic Time-of-Flight High-Resolution Mass-Spectrometer, Zhurnal Eksp. Teor. Fiz. 64 (1973) 82–89.

[87] W. Paul, H. Steinwedel, Notizen: Ein neues Massenspektrometer ohne Magnetfeld, Z. Für Naturforschung A. 8 (1953) 448–450. doi:10.1515/zna-1953-0710.

[88] M.B. Comisarow, A.G. Marshall, Fourier transform ion cyclotron resonance spectroscopy, Chem. Phys. Lett. 25 (1974) 282–283. doi:10.1016/0009-2614(74)89137-2.

[89] A. Makarov, Electrostatic Axially Harmonic Orbital Trapping: A High-Performance Technique of Mass Analysis, Anal. Chem. 72 (2000) 1156–1162. doi:10.1021/ac991131p.

[90] H. Steen, B. Küster, M. Mann, Quadrupole time-of-flight versus triple-quadrupole mass spectrometry for the determination of phosphopeptides by precursor ion scanning, J. Mass Spectrom. JMS. 36 (2001) 782–790. doi:10.1002/jms.174.

[91] L. Tong, X.Y. Zhou, A. Jylha, U. Aapola, D.N. Liu, S.K. Koh, D. Tian, J. Quah, H. Uusitalo, R.W. Beuerman, L. Zhou, Quantitation of 47 human tear proteins using high resolution multiple reaction monitoring (HR-MRM) based-mass spectrometry, J. Proteomics. 115 (2015) 36–48. doi:10.1016/j.jprot.2014.12.002.

[92] L.C. Gillet, P. Navarro, S. Tate, H. Röst, N. Selevsek, L. Reiter, R. Bonner, R. Aebersold, Targeted data extraction of the MS/MS spectra generated by data-independent acquisition: a new concept for consistent and accurate proteome analysis, Mol. Cell. Proteomics MCP. 11 (2012) O111.016717. doi:10.1074/mcp.O111.016717.

[93] J.R. Yates, D. Cociorva, L. Liao, V. Zabrouskov, Performance of a Linear Ion Trap-Orbitrap Hybrid for Peptide Analysis, Anal. Chem. 78 (2006) 493–500. doi:10.1021/ac0514624.

[94] W. Fleischmann, S. Möller, A. Gateau, R. Apweiler, A novel method for automatic functional annotation of proteins, Bioinforma. Oxf. Engl. 15 (1999) 228–233.

[95] J. Schmutz, J. Wheeler, J. Grimwood, M. Dickson, J. Yang, C. Caoile, E. Bajorek, S. Black, Y.M. Chan, M. Denys, J. Escobar, D. Flowers, D. Fotopulos, C. Garcia, M. Gomez, E. Gonzales, L. Haydu, F. Lopez, L. Ramirez, J. Retterer, A. Rodriguez, S. Rogers, A. Salazar, M. Tsai, R.M. Myers, Quality assessment of the human genome sequence, Nature. 429 (2004) 365–368. doi:10.1038/nature02390.

[96] P. Edman, Method for Determination of the Amino Acid Sequence in Peptides, Acta Chem. Scand. 4 (1950) 283–293. doi:10.3891/acta.chem.scand.04-0283.

[97] R.M. Hewick, M.W. Hunkapiller, L.E. Hood, W.J. Dreyer, A gas-liquid solid phase peptide and protein sequenator., J. Biol. Chem. 256 (1981) 7990–7997.

[98] T. Rabilloud, Paleoproteomics explained to youngsters: how did the wedding of two-dimensional electrophoresis and protein sequencing spark proteomics on: Let there be light, J. Proteomics. 107 (2014) 5–12. doi:10.1016/j.jprot.2014.03.011.

[99] A. Devabhaktuni, J.E. Elias, Application of de Novo Sequencing to Large-Scale Complex Proteomics Data Sets, J. Proteome Res. 15 (2016) 732–742. doi:10.1021/acs.jproteome.5b00861.

[100] R. Turetschek, D. Lyon, G. Desalegn, H.-P. Kaul, S. Wienkoop, A Proteomic Workflow Using High-Throughput De Novo Sequencing Towards Complementation of Genome Information for Improved Comparative Crop Science, Methods Mol. Biol. Clifton NJ. 1394 (2016) 233–243. doi:10.1007/978-1-4939-3341-9_17.

[101] K.W. Rickert, L. Grinberg, R.M. Woods, S. Wilson, M.A. Bowen, M. Baca, Combining phage display with de novo protein sequencing for reverse engineering of monoclonal antibodies, mAbs. 8 (2016) 501–512. doi:10.1080/19420862.2016.1145865.

[102] V.C. Carregari, J. Dai, T. Verano-Braga, T. Rocha, L.A. Ponce-Soto, S. Marangoni, P. Roepstorff, Revealing the functional structure of a new PLA2 K49 from Bothriopsis taeniata snake venom employing automatic "de novo" sequencing using CID/HCD/ETD MS/MS analyses, J. Proteomics. 131 (2016) 131–139. doi:10.1016/j.jprot.2015.10.020.

[103] Z.R. Gregorich, Y. Ge, Top-down Proteomics in Health and Disease: Challenges and Opportunities, Proteomics. 14 (2014) 1195–1210. doi:10.1002/pmic.201300432.

[104] L.C. Gillet, A. Leitner, R. Aebersold, Mass Spectrometry Applied to Bottom-Up Proteomics: Entering the High-Throughput Era for Hypothesis Testing, Annu. Rev. Anal. Chem. Palo Alto Calif. (2016). doi:10.1146/annurev-anchem-071015-041535.

[105] J.F. Kellie, J.C. Tran, J.E. Lee, D.R. Ahlf, H.M. Thomas, I. Ntai, A.D. Catherman, K.R. Durbin, L. Zamdborg, A. Vellaichamy, P.M. Thomas, N.L. Kelleher, The emerging process of Top Down mass spectrometry for protein analysis: biomarkers, protein-therapeutics, and achieving high throughput, Mol. Biosyst. 6 (2010) 1532. doi:10.1039/c000896f.

[106] J.R. Wiśniewski, A. Zougman, N. Nagaraj, M. Mann, Universal sample preparation method for proteome analysis, Nat. Methods. 6 (2009) 359–362. doi:10.1038/nmeth.1322.

[107] A. Zougman, P.J. Selby, R.E. Banks, Suspension trapping (STrap) sample preparation method for bottom-up proteomics analysis, Proteomics. 14 (2014) 1006–1000. doi:10.1002/pmic.201300553.

[108] C.S. Hughes, S. Foehr, D.A. Garfield, E.E. Furlong, L.M. Steinmetz, J. Krijgsveld, Ultrasensitive proteome analysis using paramagnetic bead technology, Mol. Syst. Biol. 10 (2014) 757.

[109] M. Ethier, W. Hou, H.S. Duewel, D. Figeys, The proteomic reactor: a microfluidic device for processing minute amounts of protein prior to mass spectrometry analysis, J. Proteome Res. 5 (2006) 2754–2759. doi:10.1021/pr060312m.

[110] J.R. Wiśniewski, Quantitative Evaluation of Filter Aided Sample Preparation (FASP) and Multienzyme Digestion FASP Protocols, Anal. Chem. 88 (2016) 5438–5443. doi:10.1021/acs.analchem.6b00859.

[111] J. Erde, R.R.O. Loo, J.A. Loo, Enhanced FASP (eFASP) to increase proteome coverage and sample recovery for quantitative proteomic experiments, J. Proteome Res. 13 (2014) 1885–1895. doi:10.1021/pr4010019.

[112] P. Giansanti, L. Tsiatsiani, T.Y. Low, A.J.R. Heck, Six alternative proteases for mass spectrometry-based proteomics beyond trypsin, Nat. Protoc. 11 (2016) 993–1006. doi:10.1038/nprot.2016.057.

[113] L. Tsiatsiani, A.J.R. Heck, Proteomics beyond trypsin, FEBS J. 282 (2015) 2612–2626. doi:10.1111/febs.13287.

[114] J.D. Egertson, B. MacLean, R. Johnson, Y. Xuan, M.J. MacCoss, Multiplexed peptide analysis using data-independent acquisition and Skyline, Nat. Protoc. 10 (2015) 887–903. doi:10.1038/nprot.2015.055.

[115] M. Carrascal, K. Schneider, R. Elena Calaf, S. van Leeuwen, D. Canosa, E. Gelpí, J. Abian, Quantitative electrospray LC–MS and LC–MS/MS in biomedicine1, J. Pharm. Biomed. Anal. 17 (1998) 1129–1138. doi:10.1016/S0731-7085(98)00078-8.

[116] J.-H. Kang, D. Asai, R. Toita, H. Kitazaki, Y. Katayama, Plasma protein kinase C (PKC)alpha as a biomarker for the diagnosis of cancers, Carcinogenesis. 30 (2009) 1927–1931. doi:10.1093/carcin/bgp210.

[117] D.M. Mallawaaratchy, S. Mactier, K.L. Kaufman, K. Blomfield, R.I. Christopherson, The phosphoinositide 3-kinase inhibitor LY294002, decreases aminoacyl-tRNA synthetases, chaperones and glycolytic enzymes in human HT-29 colorectal cancer cells, J. Proteomics. 75 (2012) 1590–1599. doi:10.1016/j.jprot.2011.11.032.

[118] A. Thompson, J. Schäfer, K. Kuhn, S. Kienle, J. Schwarz, G. Schmidt, T. Neumann, R. Johnstone, A.K.A. Mohammed, C. Hamon, Tandem mass tags: a novel quantification strategy for comparative analysis of complex protein mixtures by MS/MS, Anal. Chem. 75 (2003) 1895–1904.

[119] P.L. Ross, Y.N. Huang, J.N. Marchese, B. Williamson, K. Parker, S. Hattan, N. Khainovski, S. Pillai, S. Dey, S. Daniels, S. Purkayastha, P. Juhasz, S. Martin, M. Bartlet-Jones, F. He, A. Jacobson, D.J. Pappin, Multiplexed protein quantitation in Saccharomyces cerevisiae using amine-reactive isobaric tagging reagents, Mol. Cell. Proteomics MCP. 3 (2004) 1154–1169. doi:10.1074/mcp.M400129-MCP200.

[120] L. Choe, M. D'Ascenzo, N.R. Relkin, D. Pappin, P. Ross, B. Williamson, S. Guertin, P. Pribil, K.H. Lee, 8-plex quantitation of changes in cerebrospinal fluid protein expression in subjects undergoing intravenous immunoglobulin treatment for Alzheimer's disease, Proteomics. 7 (2007) 3651–3660. doi:10.1002/pmic.200700316.

[121] T. Werner, G. Sweetman, M.F. Savitski, T. Mathieson, M. Bantscheff, M.M. Savitski, Ion coalescence of neutron encoded TMT 10-plex reporter ions, Anal. Chem. 86 (2014) 3594–3601. doi:10.1021/ac500140s.

[122] D.C. Frost, T. Greer, F. Xiang, Z. Liang, L. Li, Development and characterization of novel 8-plex DiLeu isobaric labels for quantitative proteomics and peptidomics, Rapid Commun. Mass Spectrom. RCM. 29 (2015) 1115–1124. doi:10.1002/rcm.7201.

[123] S.-E. Ong, B. Blagoev, I. Kratchmarova, D.B. Kristensen, H. Steen, A. Pandey, M. Mann, Stable isotope labeling by amino acids in cell culture, SILAC, as a simple and accurate approach to expression proteomics, Mol. Cell. Proteomics MCP. 1 (2002) 376–386.

[124] S. Zanivan, M. Krueger, M. Mann, In vivo quantitative proteomics: the SILAC mouse, Methods Mol. Biol. Clifton NJ. 757 (2012) 435–450. doi:10.1007/978-1-61779-166-6_25.

[125] M. Bantscheff, B. Dümpelfeld, B. Kuster, Femtomol sensitivity post-digest (18)O labeling for relative quantification of differential protein complex composition, Rapid Commun. Mass Spectrom. RCM. 18 (2004) 869–876. doi:10.1002/rcm.1418.

[126] J.A. Vogt, K. Schroer, K. Hölzer, C. Hunzinger, M. Klemm, K. Biefang-Arndt, S. Schillo, M.A. Cahill, A. Schrattenholz, H. Matthies, W. Stegmann, Protein abundance quantification in embryonic stem cells using incomplete metabolic labelling with 15N amino acids, matrix-assisted laser desorption/ionisation time-of-flight mass spectrometry, and analysis of relative isotopologue abundances of peptides, Rapid Commun. Mass Spectrom. RCM. 17 (2003) 1273–1282. doi:10.1002/rcm.1045.

[127] S.P. Gygi, B. Rist, S.A. Gerber, F. Turecek, M.H. Gelb, R. Aebersold, Quantitative analysis of complex protein mixtures using isotope-coded affinity tags, Nat. Biotechnol. 17 (1999) 994–999. doi:10.1038/13690.

[128] A. Schmidt, J. Kellermann, F. Lottspeich, A novel strategy for quantitative proteomics using isotope-coded protein labels, Proteomics. 5 (2005) 4–15. doi:10.1002/pmic.200400873.

[129] J. Ji, A. Chakraborty, M. Geng, X. Zhang, A. Amini, M. Bina, F. Regnier, Strategy for qualitative and quantitative analysis in proteomics based on signature peptides, J. Chromatogr. B. Biomed. Sci. App. 745 (2000) 197–210.

[130] X. Ye, B. Luke, T. Andresson, J. Blonder, 18O stable isotope labeling in MS-based proteomics, Brief. Funct. Genomic. Proteomic. 8 (2009) 136–144. doi:10.1093/bfgp/eln055.

[131] J.W. Gouw, B.B.J. Tops, J. Krijgsveld, Metabolic labeling of model organisms using heavy nitrogen (15N), Methods Mol. Biol. Clifton NJ. 753 (2011) 29–42. doi:10.1007/978-1-61779-148-2_2.

[132] P.V. Bondarenko, D. Chelius, T.A. Shaler, Identification and relative quantitation of protein mixtures by enzymatic digestion followed by capillary reversed-phase liquid chromatography-tandem mass spectrometry, Anal. Chem. 74 (2002) 4741–4749.

[133] H. Liu, R.G. Sadygov, J.R. Yates, A model for random sampling and estimation of relative protein abundance in shotgun proteomics, Anal. Chem. 76 (2004) 4193–4201. doi:10.1021/ac0498563.

[134] K.A. Neilson, N.A. Ali, S. Muralidharan, M. Mirzaei, M. Mariani, G. Assadourian, A. Lee, S.C. van Sluyter, P.A. Haynes, Less label, more free: approaches in label-free quantitative mass spectrometry, Proteomics. 11 (2011) 535–553. doi:10.1002/pmic.201000553.

[135] SWATH-MS, (n.d.). http://www.imsb.ethz.ch/research/aebersold/research/swath-ms.html (accessed August 23, 2016).

[136] D.N. Perkins, D.J.C. Pappin, D.M. Creasy, J.S. Cottrell, Probability-based protein identification by searching sequence databases using mass spectrometry data, ELECTROPHORESIS. 20 (1999) 3551–3567. doi:10.1002/(SICI)1522-2683(19991201)20:18<3551::AID-ELPS3551>3.0.CO;2-2.

[137] J.K. Eng, A.L. McCormack, J.R. Yates, An approach to correlate tandem mass spectral data of peptides with amino acid sequences in a protein database, J. Am. Soc. Mass Spectrom. 5 (1994) 976–989. doi:10.1016/1044-0305(94)80016-2.

[138] J. Cox, M. Mann, MaxQuant enables high peptide identification rates, individualized p.p.b.-range mass accuracies and proteome-wide protein quantification, Nat. Biotechnol. 26 (2008) 1367–1372. doi:10.1038/nbt.1511.

[139] J. Griss, Spectral library searching in proteomics, PROTEOMICS. 16 (2016) 729–740. doi:10.1002/pmic.201500296.

[140] The Gene Ontology project in 2008, Nucleic Acids Res. 36 (2008) D440–D444. doi:10.1093/nar/gkm883.

[141] M. Kanehisa, S. Goto, KEGG: kyoto encyclopedia of genes and genomes, Nucleic Acids Res. 28 (2000) 27–30.

[142] V.N. Bhatia, D.H. Perlman, C.E. Costello, M.E. McComb, Software tool for researching annotations of proteins: open-source protein annotation software with data visualization, Anal. Chem. 81 (2009) 9819–9823. doi:10.1021/ac901335x.

[143] K.D. Dahlquist, N. Salomonis, K. Vranizan, S.C. Lawlor, B.R. Conklin, GenMAPP, a new tool for viewing and analyzing microarray data on biological pathways, Nat. Genet. 31 (2002) 19–20. doi:10.1038/ng0502-19.

[144] P.D. Karp, S. Paley, P. Romero, The Pathway Tools software, Bioinforma. Oxf. Engl. 18 Suppl 1 (2002) S225-232.

[145] M. Ashburner, C.A. Ball, J.A. Blake, D. Botstein, H. Butler, J.M. Cherry, A.P. Davis, K. Dolinski, S.S. Dwight, J.T. Eppig, M.A. Harris, D.P. Hill, L. Issel-Tarver, A. Kasarskis, S. Lewis, J.C. Matese, J.E. Richardson, M. Ringwald, G.M. Rubin, G. Sherlock, Gene ontology: tool for the unification of biology. The Gene Ontology Consortium, Nat. Genet. 25 (2000) 25–29. doi:10.1038/75556.

[146] J.A. Falkner, J.W. Falkner, P.C. Andrews, ProteomeCommons.org IO Framework: reading and writing multiple proteomics data formats, Bioinforma. Oxf. Engl. 23 (2007) 262–263. doi:10.1093/bioinformatics/btl573.

[147] D.W. Huang, B.T. Sherman, R.A. Lempicki, Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources, Nat. Protoc. 4 (2008) 44–57. doi:10.1038/nprot.2008.211.

[148] C. von Mering, M. Huynen, D. Jaeggi, S. Schmidt, P. Bork, B. Snel, STRING: a database of predicted functional associations between proteins, Nucleic Acids Res. 31 (2003) 258–261.

[149] C. Stark, B.-J. Breitkreutz, T. Reguly, L. Boucher, A. Breitkreutz, M. Tyers, BioGRID: a general repository for interaction datasets, Nucleic Acids Res. 34 (2006) D535-539. doi:10.1093/nar/gkj109.

[150] C. Prieto, J. De Las Rivas, APID: Agile Protein Interaction DataAnalyzer, Nucleic Acids Res. 34 (2006) W298-302. doi:10.1093/nar/gkl128.

[151] M.D. McDowall, M.S. Scott, G.J. Barton, PIPs: human protein–protein interaction prediction database, Nucleic Acids Res. 37 (2009) D651–D656. doi:10.1093/nar/gkn870.

[152] D.C. Liebler, L.J. Zimmerman, Targeted Quantitation of Proteins by Mass Spectrometry, Biochemistry (Mosc.). 52 (2013) 3797–3806. doi:10.1021/bi400110b.

[153] W.E. Heywood, D. Galimberti, E. Bliss, E. Sirka, R.W. Paterson, N.K. Magdalinou, M. Carecchio, E. Reid, A. Heslegrave, C. Fenoglio, E. Scarpini, J.M. Schott, N.C. Fox, J. Hardy, K. Bahtia, S. Heales, N.J. Sebire, H. Zetterburg, K. Mills, Identification of novel CSF biomarkers for neurodegeneration and their validation by a high-throughput multiplexed targeted proteomic assay, Mol. Neurodegener. 10 (2015) 64. doi:10.1186/s13024-015-0059-y.

[154] J. Adachi, R. Narumi, T. Tomonaga, Targeted Phosphoproteome Analysis Using Selected/Multiple Reaction Monitoring (SRM/MRM), Methods Mol. Biol. Clifton NJ. 1394 (2016) 87–100. doi:10.1007/978-1-4939-3341-9_7.

[155] R. Aebersold, A.L. Burlingame, R.A. Bradshaw, Western blots versus selected reaction monitoring assays: time to turn the tables?, Mol. Cell. Proteomics MCP. 12 (2013) 2381–2382. doi:10.1074/mcp.E113.031658.

[156] G.R. Hilton, J.L.P. Benesch, Two decades of studying non-covalent biomolecular assemblies by means of electrospray ionization mass spectrometry, J. R. Soc. Interface R. Soc. 9 (2012) 801–816. doi:10.1098/rsif.2011.0823.

[157] A. Abuhammad, E. Fullam, E.D. Lowe, D. Staunton, A. Kawamura, I.M. Westwood, S. Bhakta, A.C. Garner, D.L. Wilson, P.T. Seden, S.G. Davies, A.J. Russell, E.F. Garman, E. Sim, Piperidinols that show anti-tubercular activity as inhibitors of arylamine N-acetyltransferase: an essential enzyme for mycobacterial survival inside macrophages, PloS One. 7 (2012) e52790. doi:10.1371/journal.pone.0052790.

[158] H. Li, X. Bu, J. Lu, C. Xu, X. Wang, X. Yang, Interaction study of ciprofloxacin with human telomeric DNA by spectroscopy and molecular docking, Spectrochim. Acta. A. Mol. Biomol. Spectrosc. 107 (2013) 227–234. doi:10.1016/j.saa.2013.01.069.

[159] T. Hutchens, T. Yip, New Desorption Strategies for the Mass-Spectrometric Analysis of Macromolecules, Rapid Commun. Mass Spectrom. 7 (1993) 576–580. doi:10.1002/rcm.1290070703.

[160] S.E. Warder, L.A. Tucker, S.M. McLoughlin, T.J. Strelitzer, J.L. Meuth, Q. Zhang, G.S. Sheppard, P.L. Richardson, R. Lesniewski, S.K. Davidsen, R.L. Bell, J.C. Rogers, J. Wang, Discovery, identification, and characterization of candidate pharmacodynamic markers of methionine aminopeptidase-2 inhibition, J. Proteome Res. 7 (2008) 4807–4820. doi:10.1021/pr800388p.

[161] R. Mojzikova, P. Koralkova, D. Holub, Z. Zidova, D. Pospisilova, J. Cermak, Z. Striezencova Laluhova, K. Indrak, M. Sukova, M. Partschova, J. Kucerova, M. Horvathova, V. Divoky, Iron status in patients with pyruvate kinase deficiency: neonatal hyperferritinaemia associated with a novel frameshift deletion in the PKLR gene (p.Arg518fs), and low hepcidin to ferritin ratios, Br. J. Haematol. 165 (2014) 556–563. doi:10.1111/bjh.12779.

[162] T. Yoneyama, S. Ohtsuki, K. Honda, M. Kobayashi, M. Iwasaki, Y. Uchida, T. Okusaka, S. Nakamori, M. Shimahara, T. Ueno, A. Tsuchida, N. Sata, T. Ioka, Y. Yasunami, T.

Kosuge, T. Kaneda, T. Kato, K. Yagihara, S. Fujita, W. Huang, T. Yamada, M. Tachikawa, T. Terasaki, Identification of IGFBP2 and IGFBP3 As Compensatory Biomarkers for CA19-9 in Early-Stage Pancreatic Cancer Using a Combination of Antibody-Based and LC-MS/MS-Based Proteomics, PloS One. 11 (2016) e0161009. doi:10.1371/journal.pone.0161009.

[163] T. De Marchi, E. Kuhn, L.J. Dekker, C. Stingl, R.B.H. Braakman, M. Opdam, S.C. Linn, F.C.G.J. Sweep, P.N. Span, T.M. Luider, J.A. Foekens, J.W.M. Martens, S.A. Carr, A. Umar, Targeted MS Assay Predicting Tamoxifen Resistance in Estrogen-Receptor-Positive Breast Cancer Tissues and Sera, J. Proteome Res. 15 (2016) 1230–1242. doi:10.1021/acs.jproteome.5b01119.

[164] A. Dogan, Advances in clinical applications of tissue proteomics: opportunities and challenges, Expert Rev. Proteomics. 11 (2014) 531–533. doi:10.1586/14789450.2014.953062.

[165] D. Holub, PROTEOMICS IN TRANSLATIONAL AND CLINICAL RESEARCH, Univerzita Palackého v Olomouci; Ph.D. thesis, Olomouc, n.d.

[166] L. Najdekr, A. Gardlo, L. Mádrová, D. Friedecký, H. Janečková, E.S. Correa, R. Goodacre, T. Adam, Oxidized phosphatidylcholines suggest oxidative stress in patients with medium-chain acyl-CoA dehydrogenase deficiency, Talanta. 139 (2015) 62–66. doi:10.1016/j.talanta.2015.02.041.

[167] A.C.M. Veloo, E.D. de Vries, H. Jean-Pierre, U.S. Justesen, T. Morris, E. Urban, I. Wybo, A.J. van Winkelhoff, ENRIA workgroup, The optimization and validation of the Biotyper MALDI-TOF MS database for the identification of Gram-positive anaerobic cocci, Clin. Microbiol. Infect. Off. Publ. Eur. Soc. Clin. Microbiol. Infect. Dis. (2016). doi:10.1016/j.cmi.2016.06.016.

[168] T. Maier, S. Klepel, U. Renner, M. Kostrzewa, Fast and reliable MALDI-TOF MS–based microorganism identification, Nat. Methods. 3 (2006). doi:10.1038/nmeth870.

[169] A. Scherl, Clinical protein mass spectrometry, Methods San Diego Calif. 81 (2015) 3–14. doi:10.1016/j.ymeth.2015.02.015.

[170] O.S. Skinner, N.L. Kelleher, Illuminating the dark matter of shotgun proteomics, Nat. Biotechnol. 33 (2015) 717–718. doi:10.1038/nbt.3287.

[171] A. Lisitsa, S. Moshkovskii, A. Chernobrovkin, E. Ponomarenko, A. Archakov, Profiling proteoforms: promising follow-up of proteomics for biomarker discovery, Expert Rev. Proteomics. 11 (2014) 121–129. doi:10.1586/14789450.2014.878652.

[172] S. Ahmad, Platinum-DNA interactions and subsequent cellular processes controlling sensitivity to anticancer platinum complexes, Chem. Biodivers. 7 (2010) 543–566. doi:10.1002/cbdv.200800340.

[173] N.J. Wheate, S. Walker, G.E. Craig, R. Oun, The status of platinum anticancer drugs in the clinic and in clinical trials, Dalton Trans. 39 (2010) 8113–8127. doi:10.1039/C0DT00292E.

[174] L.P. Martin, T.C. Hamilton, R.J. Schilder, Platinum Resistance: The Role of DNA Repair Pathways, Clin. Cancer Res. 14 (2008) 1291–1295. doi:10.1158/1078-0432.CCR-07-2238.

[175] B. Stordal, N. Pavlakis, R. Davey, Oxaliplatin for the treatment of cisplatin-resistant cancer: a systematic review, Cancer Treat. Rev. 33 (2007) 347–357. doi:10.1016/j.ctrv.2007.01.009.

[176] B. Gustavsson, G. Carlsson, D. Machover, N. Petrelli, A. Roth, H.-J. Schmoll, K.-M. Tveit, F. Gibson, A review of the evolution of systemic chemotherapy in the management of colorectal cancer, Clin. Colorectal Cancer. 14 (2015) 1–10. doi:10.1016/j.clcc.2014.11.002.

[177] P.J. Loehrer, L.H. Einhorn, Drugs five years later. Cisplatin, Ann. Intern. Med. 100 (1984) 704–713.

[178] A. Schmitt, L. Gladieff, C.M. Laffont, A. Evrard, J.-C. Boyer, A. Lansiaux, C. Bobin-Dubigeon, M.-C. Etienne-Grimaldi, M. Boisdron-Celle, M. Mousseau, F. Pinguet, A. Floquet, E.M. Billaud, C. Durdux, C.L. Guellec, J. Mazières, T. Lafont, F. Ollivier, D. Concordet, E. Chatelut, Factors for Hematopoietic Toxicity of Carboplatin: Refining the Targeting of Carboplatin Systemic Exposure, J. Clin. Oncol. 28 (2010) 4568–4574. doi:10.1200/JCO.2010.29.3597.

[179] M.W. Saif, J. Reardon, Management of oxaliplatin-induced peripheral neuropathy, Ther. Clin. Risk Manag. 1 (2005) 249–258.

[180] M. Hajdúch, V. Mihál, J. Minarík, E. Faber, M. Safárovă, E. Weigl, P. Antálek, Decreased in vitro chemosensitivity of tumour cells in patients suffering from malignant diseases with poor prognosis, Cytotechnology. 19 (1996) 243–245.

[181] D. Biedermann, B. Eignerova, M. Hajduch, J. Sarek, Synthesis and Evaluation of Biological Activity of the Quaternary Ammonium Salts of Lupane-, Oleanane-, and Ursane-Type Acids, Synthesis. 2010 (2010) 3839–3848. doi:10.1055/s-0030-1258293.

[182] Karel Smetana, Nukleolární test, in: Vybr. Diagn. Metody Lékařské Imunol., Avicenum - zdravotnické nakladatelství, 1986: pp. 273–279.

[183] H.C.J. Hoefsloot, S. Smit, A.K. Smilde, A classification model for the Leiden proteomics competition, Stat. Appl. Genet. Mol. Biol. 7 (2008) Article8. doi:10.2202/1544-6115.1351.

[184] J. Han, P. Gao, S. Zhao, X. Bie, Z. Lu, C. Zhang, F. Lv, iTRAQ-based proteomic analysis of LI-F type peptides produced by Paenibacillus polymyxa JSa-9 mode of action against Bacillus cereus, J. Proteomics. (2016). doi:10.1016/j.jprot.2016.08.019.

[185] C. Treitz, B. Enjalbert, J.-C. Portais, F. Letisse, A. Tholey, Differential quantitative proteome analysis of Escherichia coli grown on acetate vs. glucose, Proteomics. (2016). doi:10.1002/pmic.201600303.

[186] R. Li, F. Wang, H. Liu, R. 'an Wu, H. Zou, Nano LC-MS Based Proteomic Analysis as a Predicting Approach to Study Cellular Responses of Carbon Nanotubes, J. Nanosci. Nanotechnol. 16 (2016) 2350–2359.

[187] Y. Liu, W. Sun, J. John, G. Lajoie, B. Ma, K. Zhang, De Novo Sequencing Assisted Approach for Characterizing Mixture MS/MS Spectra, IEEE Trans. Nanobioscience. (2016). doi:10.1109/TNB.2016.2519841.

[188] E. Hammer, S. Bien, M.G. Salazar, L. Steil, C. Scharf, P. Hildebrandt, H.W.S. Schroeder, H.K. Kroemer, U. Völker, C.A. Ritter, Proteomic analysis of doxorubicin-induced changes in the proteome of HepG2cells combining 2-D DIGE and LC-MS/MS approaches, Proteomics. 10 (2010) 99–114. doi:10.1002/pmic.200800626.

[189] I. Grummt, The nucleolus—guardian of cellular homeostasis and genome integrity, Chromosoma. 122 (2013) 487–497. doi:10.1007/s00412-013-0430-0.

[190] D.C. Henstridge, M. Whitham, M.A. Febbraio, Chaperoning to the metabolic party: The emerging therapeutic role of heat-shock proteins in obesity and type 2 diabetes, Mol. Metab. 3 (2014) 781–793. doi:10.1016/j.molmet.2014.08.003.

[191] Y. Duan, S. Huang, J. Yang, P. Niu, Z. Gong, X. Liu, L. Xin, R.W. Currie, T. Wu, HspA1A facilitates DNA repair in human bronchial epithelial cells exposed to Benzo[a]pyrene and interacts with casein kinase 2, Cell Stress Chaperones. 19 (2014) 271–279. doi:10.1007/s12192-013-0454-7.

[192] D.S.T. Ong, L. Wang, Y. Zhu, B. Ho, J.L. Ding, The response of ferritin to LPS and acute phase of Pseudomonas infection, J. Endotoxin Res. 11 (2005) 267–280. doi:10.1179/096805105X58698.

[193] D. Finazzi, P. Arosio, Biology of ferritin in mammals: an update on iron storage, oxidative damage and neurodegeneration, Arch. Toxicol. 88 (2014) 1787–1802. doi:10.1007/s00204-014-1329-0.

[194] R. Xing, X. Wang, C. Zhang, Y. Zhang, Q. Wang, Z. Yang, Z. Guo, Characterization and cellular uptake of platinum anticancer drugs encapsulated in apoferritin, J. Inorg. Biochem. 103 (2009) 1039–1044. doi:10.1016/j.jinorgbio.2009.05.001.

[195] E. Stanca, G. Serviddio, F. Bellanti, G. Vendemiale, L. Siculella, A.M. Giudetti, Down-regulation of LPCAT expression increases platelet-activating factor level in cirrhotic rat liver: potential antiinflammatory effect of silybin, Biochim. Biophys. Acta. 1832 (2013) 2019–2026. doi:10.1016/j.bbadis.2013.07.005.

[196] S. William-Faltaos, D. Rouillard, P. Lechat, G. Bastian, Cell cycle arrest by oxaliplatin on cancer cells, Fundam. Clin. Pharmacol. 21 (2007) 165–172. doi:10.1111/j.1472-8206.2007.00462.x.

[197] V.R. Menon, E.J. Peterson, K. Valerie, N.P. Farrell, L.F. Povirk, Ligand modulation of a dinuclear platinum compound leads to mechanistic differences in cell cycle progression and arrest, Biochem. Pharmacol. 86 (2013) 1708–1720. doi:10.1016/j.bcp.2013.10.012.

[198] K.S. Lee, D.-Y. Oh, Y.H. Kang, J.-E. Park, Self-regulated mechanism of Plk1 localization to kinetochores: lessons from the Plk1-PBIP1 interaction, Cell Div. 3 (2008) 4. doi:10.1186/1747-1028-3-4.

[199] T. Raemaekers, K. Ribbeck, J. Beaudouin, W. Annaert, M. Van Camp, I. Stockmans, N. Smets, R. Bouillon, J. Ellenberg, G. Carmeliet, NuSAP, a novel microtubule-associated protein involved in mitotic spindle organization, J. Cell Biol. 162 (2003) 1017–1029. doi:10.1083/jcb.200302129.

[200] T. Kruse, M.S.Y. Larsen, G.G. Sedgwick, J.O. Sigurdsson, W. Streicher, J.V. Olsen, J. Nilsson, A direct role of Mad1 in the spindle assembly checkpoint beyond Mad2 kinetochore recruitment, EMBO Rep. 15 (2014) 282–290. doi:10.1002/embr.201338101.

[201] N. Melling, L. Harutyunyan, C. Hube-Magg, M. Kluth, R. Simon, P. Lebok, S. Minner, M.C. Tsourlakis, C. Koop, M. Graefen, M. Adam, A. Haese, C. Wittmer, S. Steurer, J. Izbicki, G. Sauter, W. Wilczak, T. Schlomm, T. Krech, High-Level HOOK3 Expression Is an Independent Predictor of Poor Prognosis Associated with Genomic Instability in Prostate Cancer, PloS One. 10 (2015) e0134614. doi:10.1371/journal.pone.0134614.

[202] J.E. Lee, J.L. Silhavy, M.S. Zaki, J. Schroth, S.L. Bielas, S.E. Marsh, J. Olvera, F. Brancati, M. Iannicelli, K. Ikegami, A.M. Schlossman, B. Merriman, T. Attié-Bitach, C.V. Logan, I.A. Glass, A. Cluckey, C.M. Louie, J.H. Lee, H.R. Raynes, I. Rapin, I.P. Castroviejo, M. Setou, C. Barbot, E. Boltshauser, S.F. Nelson, F. Hildebrandt, C.A. Johnson, D.A. Doherty, E.M. Valente, J.G. Gleeson, CEP41 is mutated in Joubert syndrome and is required for tubulin glutamylation at the cilium, Nat. Genet. 44 (2012) 193–199. doi:10.1038/ng.1078.

[203] A.M. Waters, R. Asfahani, P. Carroll, L. Bicknell, F. Lescai, A. Bright, E. Chanudet, A. Brooks, S. Christou-Savina, G. Osman, P. Walsh, C. Bacchelli, A. Chapgier, B. Vernay, D.M. Bader, C. Deshpande, M. O' Sullivan, L. Ocaka, H. Stanescu, H.S. Stewart, F. Hildebrandt, E. Otto, C.A. Johnson, K. Szymanska, N. Katsanis, E. Davis, R. Kleta, M. Hubank, S. Doxsey, A. Jackson, E. Stupka, M. Winey, P.L. Beales, The kinetochore protein, CENPF, is mutated in human ciliopathy and microcephaly phenotypes, J. Med. Genet. 52 (2015) 147–156. doi:10.1136/jmedgenet-2014-102691.

[204] P. Noordhuis, A.C. Laan, K. van de Born, N. Losekoot, I. Kathmann, G.J. Peters, Oxaliplatin activity in selected and unselected human ovarian and colorectal cancer cell lines, Biochem. Pharmacol. 76 (2008) 53–61. doi:10.1016/j.bcp.2008.04.007.

[205] H. Liu, F. Zheng, Q. Cao, B. Ren, L. Zhu, G. Striker, H. Vlassara, Amelioration of oxidant stress by the defensin lysozyme, Am. J. Physiol. Endocrinol. Metab. 290 (2006) E824-832. doi:10.1152/ajpendo.00349.2005.

[206] K. Pai, A. Sodhi, Effect of cisplatin, rIFN-Y, LPS and MDP on release of H2O2, O2- and lysozyme from human monocytes in vitro, Indian J. Exp. Biol. 29 (1991) 910–915.

[207] L. Kadaja, S. Laos, T. Maimets, Overexpression of leukocyte marker CD43 causes activation of the tumor suppressor proteins p53 and ARF, Oncogene. 23 (2004) 2523–2530. doi:10.1038/sj.onc.1207359.

[208] D.R. Flower, A.C. North, T.K. Attwood, Structure and sequence relationships in the lipocalins and related proteins, Protein Sci. Publ. Protein Soc. 2 (1993) 753–761. doi:10.1002/pro.5560020507.

[209] J. Mishra, K. Mori, Q. Ma, C. Kelly, J. Barasch, P. Devarajan, Neutrophil gelatinase-associated lipocalin: a novel early urinary biomarker for cisplatin nephrotoxicity, Am. J. Nephrol. 24 (2004) 307–315. doi:10.1159/000078452.

[210] M.I. Hassan, A. Waheed, S. Yadav, T.P. Singh, F. Ahmad, Zinc α2-Glycoprotein: A Multidisciplinary Protein, Mol. Cancer Res. 6 (2008) 892–906. doi:10.1158/1541-7786.MCR-07-2195.

[211] M.I. Hassan, A. Waheed, S. Yadav, T.P. Singh, F. Ahmad, Prolactin inducible protein in cancer, fertility and immunoregulation: structure, function and its clinical implications, Cell. Mol. Life Sci. CMLS. 66 (2009) 447–459. doi:10.1007/s00018-008-8463-x.

[212] Q. Quan, Q. Chen, P. Chen, L. Jiang, T. Li, H. Qiu, B. Zhang, Decreased apolipoprotein A-I level indicates poor prognosis in extranodal natural killer/T-cell lymphoma, nasal type, OncoTargets Ther. 9 (2016) 1281–1290. doi:10.2147/OTT.S96549.

[213] T. Cheng, X. Dai, D.-L. Zhou, Y. Lv, L.-Y. Miao, Correlation of apolipoprotein A-I kinetics with survival and response to first-line platinum-based chemotherapy in advanced non-small cell lung cancer, Med. Oncol. Northwood Lond. Engl. 32 (2015) 407. doi:10.1007/s12032-014-0407-8.

[214] M. Zamanian-Daryoush, J.A. DiDonato, Apolipoprotein A-I and Cancer, Front. Pharmacol. 6 (2015). doi:10.3389/fphar.2015.00265.

[215] J.M. Ehrchen, C. Sunderkötter, D. Foell, T. Vogl, J. Roth, The endogenous Toll-like receptor 4 agonist S100A8/S100A9 (calprotectin) as innate amplifier of infection, autoimmunity, and cancer, J. Leukoc. Biol. 86 (2009) 557–566. doi:10.1189/jlb.1008647.

[216] E. Nicolas, P. Parisot, C. Pinto-Monteiro, R. de Walque, C. De Vleeschouwer, D.L.J. Lafontaine, Involvement of human ribosomal proteins in nucleolar structure and p53-dependent nucleolar stress, Nat. Commun. 7 (2016) 11390. doi:10.1038/ncomms11390.

[217] K. Burger, B. Mühl, T. Harasim, M. Rohrmoser, A. Malamoussi, M. Orban, M. Kellner, A. Gruber-Eber, E. Kremmer, M. Hölzel, D. Eick, Chemotherapeutic drugs inhibit ribosome biogenesis at various levels, J. Biol. Chem. 285 (2010) 12416–12425. doi:10.1074/jbc.M109.074211.

[218] G. Samimi, G. Manorek, R. Castel, J.K. Breaux, T.C. Cheng, C.C. Berry, G. Los, S.B. Howell, cDNA microarray-based identification of genes and pathways associated with oxaliplatin resistance, Cancer Chemother. Pharmacol. 55 (2005) 1–11. doi:10.1007/s00280-004-0819-9.

[219] M. Horky, G. Wurzer, V. Kotala, M. Anton, B. Vojtesek, J. Vacha, J. Wesierska-Gadek, Segregation of nucleolar components coincides with caspase-3 activation in cisplatin-treated HeLa cells, J. Cell Sci. 114 (2001) 663–670.

[220] S.M.F. Jamieson, J. Liu, T. Hsu, B.C. Baguley, M.J. McKeage, Paclitaxel induces nucleolar enlargement in dorsal root ganglion neurons in vivo reducing oxaliplatin toxicity, Br. J. Cancer. 88 (2003) 1942–1947. doi:10.1038/sj.bjc.6601012.

[221] C.L. Renn, V.A. Carozzi, P. Rhee, D. Gallop, S.G. Dorsey, G. Cavaletti, Multimodal assessment of painful peripheral neuropathy induced by chronic oxaliplatin-based chemotherapy in mice, Mol. Pain. 7 (2011) 29. doi:10.1186/1744-8069-7-29.

[222] L. Hu, J. Wang, Y. Liu, Y. Zhang, L. Zhang, R. Kong, Z. Zheng, X. Du, Y. Ke, A small ribosomal subunit (SSU) processome component, the human U3 protein 14A

(hUTP14A) binds p53 and promotes p53 degradation, J. Biol. Chem. 286 (2011) 3119–3128. doi:10.1074/jbc.M110.157842.

[223] Y.H. Lim, J.M. Charette, S.J. Baserga, Assembling a Protein-Protein Interaction Map of the SSU Processome from Existing Datasets, PLoS ONE. 6 (2011). doi:10.1371/journal.pone.0017701.

[224] M.G. Campbell, K. Karbstein, Protein-protein interactions within late pre-40S ribosomes, PloS One. 6 (2011) e16194. doi:10.1371/journal.pone.0016194.

[225] Y. Hirai, E. Louvet, T. Oda, M. Kumeta, Y. Watanabe, T. Horigome, K. Takeyasu, Nucleolar scaffold protein, WDR46, determines the granular compartmental localization of nucleolin and DDX21, Genes Cells Devoted Mol. Cell. Mech. 18 (2013) 780–797. doi:10.1111/gtc.12077.

[226] J. Baßler, H. Paternoga, I. Holdermann, M. Thoms, S. Granneman, C. Barrio-Garcia, A. Nyarko, W. Lee, G. Stier, S.A. Clark, D. Schraivogel, M. Kallas, R. Beckmann, D. Tollervey, E. Barbar, I. Sinning, E. Hurt, A network of assembly factors is involved in remodeling rRNA elements during preribosome maturation, J. Cell Biol. 207 (2014) 481–498. doi:10.1083/jcb.201408111.

[227] M. Kellner, M. Rohrmoser, I. Forné, K. Voss, K. Burger, B. Mühl, A. Gruber-Eber, E. Kremmer, A. Imhof, D. Eick, DEAD-box helicase DDX27 regulates 3' end formation of ribosomal 47S RNA and stably associates with the PeBoW-complex, Exp. Cell Res. 334 (2015) 146–159. doi:10.1016/j.yexcr.2015.03.017.

[228] K. Shimoji, J. Jakovljevic, K. Tsuchihashi, Y. Umeki, K. Wan, S. Kawasaki, J. Talkish, J.L. Woolford, K. Mizuta, Ebp2 and Brx1 function cooperatively in 60S ribosomal subunit assembly in Saccharomyces cerevisiae, Nucleic Acids Res. 40 (2012) 4574–4588. doi:10.1093/nar/gks057.

[229] J. Nousbeck, R. Spiegel, A. Ishida-Yamamoto, M. Indelman, A. Shani-Adir, N. Adir, E. Lipkin, S. Bercovici, D. Geiger, M.A. van Steensel, P.M. Steijlen, R. Bergman, A. Bindereif, M. Choder, S. Shalev, E. Sprecher, Alopecia, Neurological Defects, and Endocrinopathy Syndrome Caused by Decreased Expression of RBM28, a Nucleolar Protein Associated with Ribosome Biogenesis, Am. J. Hum. Genet. 82 (2008) 1114–1121. doi:10.1016/j.ajhg.2008.03.014.

[230] C. Sun, J.L. Woolford, The yeast nucleolar protein Nop4p contains four RNA recognition motifs necessary for ribosome biogenesis, J. Biol. Chem. 272 (1997) 25345–25352.

[231] K. Saito, Y. Iizuka, S. Ohta, S. Takahashi, K. Nakamura, H. Saya, K. Yoshida, Y. Kawakami, M. Toda, Functional analysis of a novel glioma antigen, EFTUD1, Neuro-Oncol. 16 (2014) 1618–1629. doi:10.1093/neuonc/nou132.

[232] N. Asano, K. Kato, A. Nakamura, K. Komoda, I. Tanaka, M. Yao, Structural and functional analysis of the Rpf2-Rrs1 complex in ribosome biogenesis, Nucleic Acids Res. 43 (2015) 4746–4757. doi:10.1093/nar/gkv305.

[233] C. Zhang, C. Yin, L. Wang, S. Zhang, Y. Qian, J. Ma, Z. Zhang, Y. Xu, S. Liu, HSPC111 Governs Breast Cancer Growth by Regulating Ribosomal Biogenesis, Mol. Cancer Res. 12 (2014) 583–594. doi:10.1158/1541-7786.MCR-13-0168.

[234] Z. Xu, T.C. Hobman, The helicase activity of DDX56 is required for its role in assembly of infectious West Nile virus particles, Virology. 433 (2012) 226–235. doi:10.1016/j.virol.2012.08.011.

[235] T. Kumazawa, K. Nishimura, N. Katagiri, S. Hashimoto, Y. Hayashi, K. Kimura, Gradual reduction in rRNA transcription triggers p53 acetylation and apoptosis via MYBBP1A, Sci. Rep. 5 (2015). doi:10.1038/srep10854.

[236] T. de Lange, How Shelterin Solves the Telomere End-Protection Problem, Cold Spring Harb. Symp. Quant. Biol. 75 (2010) 167–177. doi:10.1101/sqb.2010.75.017.

[237] A.-G. Lu, H. Feng, P.-X.-Z. Wang, D.-P. Han, X.-H. Chen, M.-H. Zheng, Emerging roles of the ribonucleotide reductase M2 in colorectal cancer and ultraviolet-induced DNA

damage repair, World J. Gastroenterol. WJG. 18 (2012) 4704–4713. doi:10.3748/wjg.v18.i34.4704.

[238] K. Yano, K. Morotomi-Yano, N. Adachi, H. Akiyama, Molecular mechanism of protein assembly on DNA double-strand breaks in the non-homologous end-joining pathway, J. Radiat. Res. (Tokyo). 50 (2009) 97–108.

[239] S. Kotian, T. Banerjee, A. Lockhart, K. Huang, U.V. Catalyurek, J.D. Parvin, NUSAP1 influences the DNA damage response by controlling BRCA1 protein levels, Cancer Biol. Ther. 15 (2014) 533–543. doi:10.4161/cbt.28019.

[240] J.-Y. Kim, K.-O. Seok, Y.-J. Kim, W.K. Bae, S. Lee, J.-H. Park, Involvement of GLTSCR2 in the DNA Damage Response, Am. J. Pathol. 179 (2011) 1257–1264. doi:10.1016/j.ajpath.2011.05.041.

[241] H.-Y. Chang, C.-C. Fan, P.-C. Chu, B.-E. Hong, H.J. Lee, M.-S. Chang, hPuf-A/KIAA0020 modulates PARP-1 cleavage upon genotoxic stress, Cancer Res. 71 (2011) 1126–1134. doi:10.1158/0008-5472.CAN-10-1831.

[242] M. Pietrzak, S.C. Smith, J.T. Geralds, T. Hagg, C. Gomes, M. Hetman, Nucleolar disruption and apoptosis are distinct neuronal responses to etoposide-induced DNA damage, J. Neurochem. 117 (2011) 1033–1046. doi:10.1111/j.1471-4159.2011.07279.x.

[243] M. Sokka, K. Rilla, I. Miinalainen, H. Pospiech, J.E. Syväoja, High levels of TopBP1 induce ATR-dependent shut-down of rRNA transcription and nucleolar segregation, Nucleic Acids Res. 43 (2015) 4975–4989. doi:10.1093/nar/gkv371.

[244] A. James, Y. Wang, H. Raje, R. Rosby, P. DiMario, Nucleolar stress with and without p53, Nucleus. 5 (2014) 402–426. doi:10.4161/nucl.32235.

[245] L. Montanaro, D. Treré, M. Derenzini, Nucleolus, Ribosomes, and Cancer, Am. J. Pathol. 173 (2008) 301–310. doi:10.2353/ajpath.2008.070752.

[246] J. Li, L. Yu, H. Zhang, J. Wu, J. Yuan, X. Li, M. Li, Down-regulation of pescadillo inhibits proliferation and tumorigenicity of breast cancer cells, Cancer Sci. 100 (2009) 2255–2260. doi:10.1111/j.1349-7006.2009.01325.x.

[247] P. Jordan, M. Carmo-Fonseca, Cisplatin inhibits synthesis of ribosomal RNA in vivo, Nucleic Acids Res. 26 (1998) 2831–2836. doi:10.1093/nar/26.12.2831.

[248] R.D. Hannan, D. Drygin, R.B. Pearson, Targeting RNA polymerase I transcription and the nucleolus for cancer therapy, Expert Opin. Ther. Targets. 17 (2013) 873–878. doi:10.1517/14728222.2013.818658.

[249] S.P. Henry, M. Takanosu, T.C. Boyd, P.M. Mayne, H. Eberspaecher, W. Zhou, B. de Crombrugghe, M. Hook, R. Mayne, Expression pattern and gene characterization of asporin. a newly discovered member of the leucine-rich repeat protein family, J. Biol. Chem. 276 (2001) 12212–12221. doi:10.1074/jbc.M011290200.

[250] P. Lorenzo, A. Aspberg, P. Onnerfjord, M.T. Bayliss, P.J. Neame, D. Heinegard, Identification and characterization of asporin. a novel member of the leucine-rich repeat protein family closely related to decorin and biglycan, J. Biol. Chem. 276 (2001) 12201–12211. doi:10.1074/jbc.M010932200.

[251] S. Yamada, S. Murakami, R. Matoba, Y. Ozawa, T. Yokokoji, Y. Nakahira, K. Ikezawa, S. Takayama, K. Matsubara, H. Okada, Expression profile of active genes in human periodontal ligament and isolation of PLAP-1, a novel SLRP family gene, Gene. 275 (2001) 279–286.

[252] C.R. Salmon, A.P.O. Giorgetti, A.F. Paes Leme, R.R. Domingues, E.A. Sallum, M.C. Alves, T.N. Kolli, B.L. Foster, F.H. Nociti Jr., Global proteome profiling of dental cementum under experimentally-induced apposition, J. Proteomics. 141 (2016) 12–23. doi:10.1016/j.jprot.2016.03.036.

[253] L. Xu, Z. Li, S.-Y. Liu, S.-Y. Xu, G.-X. Ni, Asporin and osteoarthritis, Osteoarthr. Cartil. OARS Osteoarthr. Res. Soc. 23 (2015) 933–939. doi:10.1016/j.joca.2015.02.011.

[254] J. Sun, T. Zhang, P. Zhang, L. Lv, Y. Wang, J. Zhang, S. Li, Overexpression of the PLAP-1 gene inhibits the differentiation of BMSCs into osteoblast-like cells, J. Mol. Histol. 45 (2014) 599–608. doi:10.1007/s10735-014-9585-0.

[255] I. Kou, M. Nakajima, S. Ikegawa, Binding characteristics of the osteoarthritis-associated protein asporin, J. Bone Miner. Metab. 28 (2010) 395–402. doi:10.1007/s00774-009-0145-8.

[256] S. Kalamajski, A. Aspberg, K. Lindblom, D. Heinegård, A. Oldberg, Asporin competes with decorin for collagen binding, binds calcium and promotes osteoblast collagen mineralization, Biochem. J. 423 (2009) 53–59. doi:10.1042/BJ20090542.

[257] Q9BXN1: 3D Model From SWISSMODEL Based On 2ft3F -- Protein Model Portal - PSI SBKB, (n.d.). http://www.proteinmodelportal.org/?pid=modelDetail&provider=SWISSMODEL&template=2ft3F&pmpuid=1000866460871&range_from=1&range_to=380&ref_ac=Q9BXN1&mapped_ac=Q9BXN1&zid=async (accessed May 19, 2016).

[258] S. Ikegawa, The genetics of common degenerative skeletal disorders: osteoarthritis and degenerative disc disease, Annu. Rev. Genomics Hum. Genet. 14 (2013) 245–256. doi:10.1146/annurev-genom-091212-153427.

[259] S. Murphy, P. Dowling, M. Zweyer, R.R. Mundegar, M. Henry, P. Meleady, D. Swandulla, K. Ohlendieck, Proteomic analysis of dystrophin deficiency and associated changes in the aged mdx-4cv heart model of dystrophinopathy-related cardiomyopathy, J. Proteomics. (n.d.). doi:10.1016/j.jprot.2016.03.011.

[260] H.T. Tan, T.K. Lim, A.M. Richards, T. Kofidis, K.L.-K. Teoh, L.H. Ling, M.C.M. Chung, Unravelling the proteome of degenerative human mitral valves, PROTEOMICS. 15 (2015) 2934–2944. doi:10.1002/pmic.201500040.

[261] R. Satoyoshi, S. Kuriyama, N. Aiba, M. Yashiro, M. Tanaka, Asporin activates coordinated invasion of scirrhous gastric cancer and cancer-associated fibroblasts, Oncogene. 34 (2015) 650–660. doi:10.1038/onc.2013.584.

[262] Q. Ding, M. Zhang, C. Liu, Asporin participates in gastric cancer cell growth and migration by influencing EGF receptor signaling, Oncol. Rep. 33 (2015) 1783–1790. doi:10.3892/or.2015.3791.

[263] P.J. Hurley, D. Sundi, B. Shinder, B.W. Simons, R.M. Hughes, R.M. Miller, B. Benzon, S.F. Faraj, G.J. Netto, I.A. Vergara, N. Erho, E. Davicioni, R.J. Karnes, G. Yan, C. Ewing, S.D. Isaacs, D.M. Berman, J.R. Rider, K.M. Jordahl, L.A. Mucci, J. Huang, S.S. An, B.H. Park, W.B. Isaacs, L. Marchionni, A.E. Ross, E.M. Schaeffer, Germline Variants in Asporin Vary by Race, Modulate the Tumor Microenvironment, and Are Differentially Associated with Metastatic Prostate Cancer, Clin. Cancer Res. Off. J. Am. Assoc. Cancer Res. 22 (2016) 448–458. doi:10.1158/1078-0432.CCR-15-0256.

[264] D. Ansari, L. Aronsson, A. Sasor, C. Welinder, M. Rezeli, G. Marko-Varga, R. Andersson, The role of quantitative mass spectrometry in the discovery of pancreatic cancer biomarkers for translational science, J. Transl. Med. 12 (2014) 87. doi:10.1186/1479-5876-12-87.

[265] P. Maris, A. Blomme, A.P. Palacios, B. Costanza, A. Bellahcène, E. Bianchi, S. Gofflot, P. Drion, G.E. Trombino, E. Di Valentin, P.G. Cusumano, S. Maweja, G. Jerusalem, P. Delvenne, E. Lifrange, V. Castronovo, A. Turtoi, Asporin Is a Fibroblast-Derived TGF-β1 Inhibitor and a Tumor Suppressor Associated with Good Prognosis in Breast Cancer, PLoS Med. 12 (2015) e1001871. doi:10.1371/journal.pmed.1001871.

[266] J.L. Luque-Garcia, T.A. Neubert, On-membrane tryptic digestion of proteins for mass spectrometry analysis, Methods Mol. Biol. Clifton NJ. 536 (2009) 331–341. doi:10.1007/978-1-59745-542-8_35.

[267] B. MacLean, D.M. Tomazela, N. Shulman, M. Chambers, G.L. Finney, B. Frewen, R. Kern, D.L. Tabb, D.C. Liebler, M.J. MacCoss, Skyline: an open source document editor

for creating and analyzing targeted proteomics experiments, Bioinformatics. 26 (2010) 966–968. doi:10.1093/bioinformatics/btq054.

[268] J. Přichystal, Identifikace tumor-specifických mutovaných forem K-ras proteinu pomocí hmotnostní spektrometrie, Univerzita Palackého v Olomouci; diplomová práce, n.d.

[269] F. Haurowitz, M. Tunca, P. Schwerin, V. Goksu, The Action of Trypsin on Native and Denatured Proteins, J. Biol. Chem. 157 (1945) 621–625.

[270] M. Miyakawa, G. Zukeran, Y. Wada, K. Akama, Conserved C272/278 in b domain regulate the function of PDI-P5 to make lysozymes trypsin-resistant forms via significant intermolecular disulfide cross-linking, Biochim. Biophys. Acta. 1854 (2015) 485–491. doi:10.1016/j.bbapap.2015.02.014.

[271] J.H. Northrop, THE KINETICS OF TRYPSIN DIGESTION, J. Gen. Physiol. 6 (1924) 439–452.

[272] M. Ye, Y. Pan, K. Cheng, H. Zou, Protein digestion priority is independent of protein abundances, Nat. Methods. 11 (2014) 220–222. doi:10.1038/nmeth.2850.

[273] M.J. McKeage, T. Hsu, D. Screnci, G. Haddad, B.C. Baguley, Nucleolar damage correlates with neurotoxicity induced by different platinum drugs, Br. J. Cancer. 85 (2001) 1219–1225. doi:10.1054/bjoc.2001.2024.

[274] C. Pfirschke, C. Engblom, S. Rickelt, V. Cortez-Retamozo, C. Garris, F. Pucci, T. Yamazaki, V. Poirier-Colame, A. Newton, Y. Redouane, Y.-J. Lin, G. Wojtkiewicz, Y. Iwamoto, M. Mino-Kenudson, T.G. Huynh, R.O. Hynes, G.J. Freeman, G. Kroemer, L. Zitvogel, R. Weissleder, M.J. Pittet, Immunogenic Chemotherapy Sensitizes Tumors to Checkpoint Blockade Therapy, Immunity. 44 (2016) 343–354. doi:10.1016/j.immuni.2015.11.024.

6 ABBREVIATIONS

2DE	2D electrophoresis
2DIGE	2D difference gel electrophoresis
APCI	Atmospheric pressure chemical ionization
APPI	Atmospheric pressure photoionization
AUC	Area under curve
CarboPt	Carboplatin
CCD	Charge-Coupled Device
CI	Chemical ionization
CID	Collision induced dissociation
CisPt	Cisplatin
DAVID	Database for annotation, visualization and integrated discovery
DDA	Data dependent analysis
DESI	Desorption electrospray ionization
DIA	Data independent analysis
DTT	Dithiothreitol
ECD	Electron capture dissociation
EI	Electron ionization
ELISA	Enzyme-linked immunosorbent assay
ERLIC	Electrostatic repulsion-hydrophilic interaction chromatography
ESI	Electrospray ionization
ETD	Electron transfer dissociation
FASP	Filter aided sample preparation
FT	Fourier transformation
GIST	Global internal standard technology
GO	Gene Ontology
HILIC	Hydrophilic interaction liquid chromatography
HPLC	High performance liquid chromatography
HR-MRM	High-resolution multiple reaction monitoring
IC_{50}	50% inhibition constant
ICAT	Isotope coded affinity tag
ICPL	Isotope coded protein labelling
IEF	Isoelectric focusing
IF	Immunofluorescence microscopy

IHC	Immunohistochemistry
IMAC	Immobilized metal affinity chromatography
IPTG	Isopropyl-β-D-thiogalacto-pyranoside
IT	Ion trap
iTRAQ	Isobaric tags for relative and absolute quantitation
KEGG	Kyoto encyclopedia of genes and genomes
LC	Liquid chromatography
Lys-C	Lysyl endopeptidase from Lysobacter enzymogenesis
MALDI	Matrix assisted laser desorption/ionization
MOAC	Metal oxide affinity chromatography
MRM	Multiple reactions monitoring
MS	Mass spectrometry
MTT	3-(4,5-dimethylthiazol-2-yl)-2,5-diphenyltetrazolium bromide
MudPIT	Multidimensional Protein Identification Technology
OxaPt	Oxaliplatin
PBS	Phosphate buffered saliva
PMSF	Phenylmethanesulfonyl fluoride
PRM	Parallel reaction monitoring
PTM	Post-translational modification
PVDF	Polyvinylidene fluoride
PVP	Polyvinylpyrollidon
Q	Quadrupole
Q-TOF	Quadrupole with time-of-flight
RP	Reverse phase
SAX	Strong anion exchange
SCX	Strong cation exchange
SDS	Sodium dodecyl sulphate
SDS-PAGE	Sodium dodecyl sulphate polyacrylamide gel electrophoresis
SELDI	Surface enhanced laser desorption/ionisation
SILAC	Stable isotope labelled amino acid in cell culture
SRM	Selected reaction monitoring
SWATH	Sequential window acquisition of all theoretical fragment
TCA	Trichloroacetic acid
TFA	Trifluoroacetic acid
TMT	Tandem mass tags

TOF   Time-of-flight

UV   Ultraviolet light

WB   Western blot

7 BIBLIOGRAPHY

Articles related to PhD thesis are labelled*

7.1 Original articles and reviews

*RYLOVÁ, G.[+], **T. OŽDIAN**[+], L. VARANASI, M. SOURAL, J. HLAVÁČ, D. HOLUB, P. DŽUBÁK a M. HAJDÚCH. Affinity-Based Methods in Drug-Target Discovery. Current Drug Targets. 2015, 16(1), 60-76. ISSN 1389-4501. IF: 3.597. PMID: 25410410 ([+]both authors contributed equally to the study)

***OŽDIAN, T.**, D. HOLUB, G. RYLOVÁ, J. VÁCLAVKOVÁ, M. HAJDÚCH a P. DŽUBÁK. Porovnání hmotnostně spektrometrických přístupů v proteomickém profilování léčiv. Chemagazín. 2016, XXVI(5), 8-11. ISSN 1210-7409.

*ŠIMKOVÁ, D., G. KHARAISHVILI, G. KOŘÍNKOVÁ, **T. OŽDIAN**, T. SUCHANKOVA-KLEPLOVA, T. SOUKUP, M. KRUPKA, A. GALANDÁKOVÁ, P. DŽUBÁK, M. JANÍKOVÁ, J. NAVRATIL, Z. KAHOUNOVA, K. SOUCEK a J. BOUCHAL. CHYBI DEDI a PDF_The dual role of asporin in breast cancer progression. Oncotarget. 2016, -, -. ISSN 1949-2553 . IF: 5.008. PMID: 27409832

LIGASOVÁ, A., R. LIBOSKA, D. FRIEDECKÝ, K. MIČOVÁ, T. ADAM, **T. OŽDIAN**, I. ROSENBERG a K. KOBERNA. Dr Jekyll and Mr Hyde: A Strange Case of 5-Ethynyl-2'-deoxyuridine and 5-ethynyl-2'-deoxycytidine. Open Biology. 2016, 6: 150172, 1-11. ISSN 2046-2441. IF: 5.784. PMID: 26740587

LEMROVÁ, B., P. SMYSLOVÁ, I. POPA, **T. OŽDIAN**, P. ZAJDEL a M. SOURAL. Directed Solid-Phase Synthesis of Trisubstituted Imidazo[4,5-c]pyridines and Imidazo[4,5-b]pyridines. ACS Combinatorial Science. 2014, 16(10), 558-565. ISSN 2156-8952. IF: 3.401. PMID: 25046560

7.2 Oral and poster presentations

OŽDIAN, T., D. HOLUB, G. RYLOVÁ, P. DŽUBÁK a M. HAJDÚCH. Oxaliplatin treatment causes nucleolar and ribosomal protein decrease. In : XI. Diagnostic, Predictive and Experimetnal Oncology Days: Abstract Book. 2.12.-3.12.2015, hotel NH Congress, Olomouc, 2015. DEDIKACE: LF_2015_031, LO1304.

OŽDIAN, T., D. HOLUB, G. RYLOVÁ, M. KOLLAREDDY, P. DŽUBÁK a M. HAJDÚCH. Oxaliplatin treatment causes nucleolar and ribosomal protein decrease. In: 9th Central and Eastern

European Proteomics Conference: Book of Abstracts and Program. 14.6. – 18.6.2015, Poznań, 2015. DEDIKACE: LF_2015_031, LO1304, TE02000058.

OŽDIAN, T., D. HOLUB, G. RYLOVÁ, M. KOLLAREDDY, P. DŽUBÁK a M. HAJDÚCH. Proteomic profilling of oxaliplatin treated cell line revealed nucleoli and ribosomal protein downregulation. In : X. Diagnostic, Predictive and Experimetnal Oncology Days: Abstract Book. 2.12.-3.12.2014, hotel NH Congress, Olomouc, 2015. DEDIKACE: LF_2014_010, TE02000058, CZ.1.05/3.1.00/14.0307, CZ.1.05/2.1.00/01.0030.

OŽDIAN, T., D. HOLUB, P. DŽUBÁK, G. RYLOVÁ, M. KOLLAREDDY, J. ŘEHULKA, L. RADOVÁ a M. HAJDÚCH. In vitro proteomické profilování buněčné linie CEM platinovými protinádorovými léčivy. In: IX. Diagnostic, Predictive and Experimetnal Oncology Days: Abstract Book. 21.11.-22.11.2013, hotel NH Congress, Olomouc, 2013, A37. DEDIKACE: LF_2013_016, LF_2013_015, BIOMEDREG CZ.1.05/25.1.00/01.0030.

OŽDIAN, T., J. VÁCLAVKOVÁ, D. HOLUB, P. DŽUBÁK, G. RYLOVÁ, J. ŘEHULKA, M. HRUŠKA a M. HAJDÚCH. Proteomic profiling of vinca and taxus alkaloid based anticancer drugs. In: Sborník abstrakt: Konference Chemické biologie a genetiky. 12.5.-14.5.2013, Karlov pod Pradědem, 2013, 19. DEDIKACE: LF_2013_016, BIOMEDREG CZ.1.05/25.1.00/01.0030.

OŽDIAN, T., D. HOLUB, P. DŽUBÁK, G. RYLOVÁ, M. KOLLAREDDY, J. ŘEHULKA, D. CAHOVÁ, V. HAVLÍČEK a M. HAJDÚCH. In vitro proteomické profilování buněčné linie CEM cisplatinou, karboplatinou, oxaliplatinou, daunorubicinem a vinkristinem. In: II. Neformální proteomické setkání. Kniha výtahů. 2012, 15. BEZ IF DEDIKACE: LF_2012_018, BIOMEDREG CZ.1.05/25.1.00/01.0030.

OŽDIAN, T., D. HOLUB, P. DŽUBÁK, G. RYLOVÁ, M. R. KOLLAREDDY, J. ŘEHULKA, L. RADOVÁ, D. CAHOVÁ, V. HAVLÍČEK a M. HAJDÚCH. In vitro protetické profilování buněčné linie CEM platinovými protinádorovými léčivy. In: VIII. dny diagnostické, prediktivní a experimentální onkologie. Onkologie. 2012, 5, Suppl. B, B24. ISSN 1802-4475. BEZ IF. DEDIKACE: UP (LF_2012_018), BIOMEDREG CZ.1.05/25.1.00/01.0030.

OŽDIAN, T., D. HOLUB, P. DŽUBÁK, G. RYLOVÁ, M. KOLLAREDY, J. ŘEHULKA, D. CAHOVÁ, V. HAVLÍČEK a M. HAJDÚCH. In Vitro Proteomic Profiling of Platinum Drugs Used in Cancer Therapy. In: 6th Central and Eastern European Proteomics Conference: Book of Abstracts and Program. 2012, 84. ISBN 978-963-9970-28-1. BEZ IF. DEDIKACE: LF_2012_018, CZ.1.05/25.1.00/01.0030.

OŽDIAN, T., D. HOLUB, P. DŽUBÁK, G. RYLOVÁ, M. KOLLAREDDY, J. ŘEHULKA, R. ČÁBELKA, D. CAHOVÁ, V. HAVLÍČEK a M. HAJDÚCH. Comparison of proteomic profiles of cisplatin, carboplatin, oxaliplatin, daunorubicin and vincristine-treated CEM cell line. In: CHEMICA 49S: COLLECTED REPORTS OF THE NATURAL SCIENCE FACULTY, PALACKÝ UNIVERSITY OLOMOUC, CZECH REPUBLIC. Book of Abstracts and Program. 30th Informal Meeting on Mass Spectrometry. Olomouc: Přírodovědecká fakulta UP, 2012, 102. ISBN 978-80-244-3047-8 / ISSN 0232-0061. BEZ IF. DEDIKACE: IGA UPOL 9011100251, BIOMEDREG CZ.1.05/2.1.00/01.0030.

OŽDIAN, T., D. HOLUB, P. DŽUBÁK, G. RYLOVÁ, M. KOLLAREDDY, J. ŘEHULKA, R. ČÁBELKA, D. CAHOVÁ, V. HAVLÍČEK a M. HAJDÚCH. Evaluation of complex proteomic samples in case of platinum drugs treated CEM cell line (Proteomické profilování platinových léčiv). In: VII. DNY DIAGNOSTICKÉ, PREDIKTIVNÍ A EXPERIMENTÁLNÍ ONKOLOGIE. Onkologie. 2011, 5, Suppl. B, B15-B16. ISSN 1802-4475. BEZ IF. DEDIKACE: LC07017, MSM6198959216, GAČR (H-084), Biomedreg CZ.1.05/2.1.00/01.0030.

OŽDIAN, T., D. HOLUB, P. DŽUBÁK, G. RYLOVÁ, M. KOLLAREDDY, J. ŘEHULKA, V. HAVLÍČEK a M. HAJDÚCH. Proteomic profile of platinum based drug treatment of CEM cells. In: Book of Abstracts: 5th Central and Eastern European Proteomic Conference. 2011, 62. BEZ IF. DEDIKACE: LC07017, MSM6198959216, GAČR H-084, BIOMEDREG CZ.1.05/2.1.00/01.0030.

OŽDIAN, T., D. HOLUB, P. DŽUBÁK, G. RYLOVÁ, M. KOLLAREDDY, J. ŘEHULKA, V. HAVLÍČEK a M. HAJDÚCH. Proteomic Profile CEM cells Treated by Platinum Based Drugs. In: Sborník abstrakt konference: Nová léčiva závažných lidských onemocnění, 2011, 16. BEZ IF. DEDIKACE: LC07017, MSM6198959216, Czech Science Foundation (H-084), BIOMEDREG CZ.1.05/2.1.00/01.0030.

OŽDIAN, T., D. HOLUB, P. DŽUBÁK, G. RYLOVÁ, M. KOLLAREDDY REDDY, J. ŘEHULKA, V. HAVLÍČEK a M. HAJDÚCH. Proteomic profile of oxaliplatin treatment in CEM lymphoblastoid cells. In: VI. DNY DIAGNOSTICKÉ, PREDIKTIVNÍ A EXPERIMENTÁLNÍ ONKOLOGIE. Onkologie. 2010, 4, Suppl.A.,14-15. ISSN 1803-5922. BEZ IF. DEDIKACE: LC07017, MSM6198959216, Czech Science Foundation (H-084), (BIOMEDREG) CZ.1.05/2.1.00/01.0030.

OŽDIAN, T., D. HOLUB, G. RYLOVÁ, P. DŽUBÁK, P. NOVÁK, M. KOLLAREDDY REDDY, P. MAN, V. HAVLÍČEK a M. HAJDÚCH. Expressional Proteomic Profilig of Cisplatin Treated Leukaemia Derived Cell Line. In: Sborník abstrakt: Biotechnologie v medicíně. Vílanec u Jihlavy, 27.-29. května 2010, 9. BEZ IF. DEDIKACE: MŠM6198959216, LC07017, GAČR H-084.

8   APPENDIX - FULL TEXT PUBLICATIONS RELATED TO THE THESIS

8.1 Appendix A

RYLOVÁ, G.*, **T. OŽDIAN***, L. VARANASI, M. SOURAL, J. HLAVÁČ, D. HOLUB, P. DŽUBÁK a M. HAJDÚCH. Affinity-Based Methods in Drug-Target Discovery. Current Drug Targets. 2015, 16(1), 60-76. ISSN 1389-4501. IF: 3.597. PMID: 25410410 (*both authors contributed equally to the study)

1

Affinity-Based Methods in Drug-Target Discovery

Gabriela Rylova[1,#], Tomas Ozdian[1,#], Lakshman Varanasi[1,#], Miroslav Soural[1,2], Jan Hlavac[1,2], Dusan Holub[1], Petr Dzubak[1] and Marian Hajduch[1,*]

[1]*Laboratory of Experimental Medicine, Institute of Molecular and Translational Medicine, Faculty of Medicine and Dentistry, Palacky University and University Hospital in Olomouc, Hnevotinska 5, 77515 Olomouc, The Czech Republic;* [2]*Department of Medicinal Chemistry, Institute of Molecular and Translational Medicine, Faculty of Science, Palacky University, 17. Listopadu 12, 771 46 Olomouc, The Czech Republic*

Abstract: Target discovery using the molecular approach, as opposed to the more traditional systems approach requires the study of the cellular or biological process underlying a condition or disease. The approaches that are employed by the "bench" scientist may be genetic, genomic or proteomic and each has its rightful place in the drug-target discovery process. Affinity-based proteomic techniques currently used in drug-discovery draw upon several disciplines, synthetic chemistry, cell-biology, biochemistry and mass spectrometry. An important component of such techniques is the probe that is specifically designed to pick out a protein or set of proteins from amongst the varied thousands in a cell lysate. A second component, that is just as important, is liquid-chromatography tandem mass-spectrometry (LC-MS/MS). LC-MS/MS and the supporting theoretical framework has come of age and is the tool of choice for protein identification and quantification. These proteomic tools are critical to maintaining the drug-candidate supply, in the larger context of drug discovery.

Keywords: Affinity purification, drug-target discovery, immobilization, mass-spectrometry, probe, quantification.

1. INTRODUCTION

The identification of one or more "druggable" components of a biological pathway or network is often the starting point for the de-novo development of a drug although it has not always been so. The use of many therapeutic agents has traditionally been on the basis of their observed effects (phenotypic screening) and not on the basis of their cellular mechanism of action (rational drug design) [1-3]. Overington *et al.* have reported that only 82% of Food and Drug Administration (FDA) drugs approved between 1989 and 2000 had one or more defined cellular target [4, 5]. Approved drugs with unknown mechanisms of action include Bexarotene and arsenic trioxide, for the treatment of cutaneous T-cell lymphoma and acute promyelocytic leukemia (PML) respectively [6]. Such drugs may have off-target effects that are the result of drug-binding to targets that are not known or not intended, and may be detrimental or beneficial to the patient [6-8]. Drugs having poly-pharmacological effects that are beneficial, especially in diseases like cancer, are at a premium [9]. Conversely, drugs with adverse off-target effects can sometimes be repurposed. Thalidomide is an illustrative example. Originally prescribed to expectant women to ease pregnancy-associated nausea and "morning sickness"

in 1957, it was discontinued when discovered to cause horrific congenital abnormalities. Five decades later, the drug has found new use in leprosy and cancer treatment [10]. Celecoxib is another example. It was designed to treat osteoarthritis but is now prescribed for colorectal cancer prevention [11]. The small-molecule Histone deacetylase 6 (HDAC6) inhibitor, tubacin also inhibits Serine Palmitoyl transferase (SPT), the rate-limiting enzyme of sphingolipid biosynthesis [12].

A clearer understanding of the in-vivo interaction behavior of the drug-candidate will facilitate the design of a drug that has a more defined target or targets. Such a drug-candidate is likely to have more predictable effects and can better inform clinical trials. In any case, the rationale for more complete knowledge of drug-targets is considerable and must be heeded.

Our knowledge of cellular and physiological processes of the human body has gone hand-in-hand with the development of analytical tools and procedures in the chemical and biomedical sciences. Conventional techniques such as electrophoresis and chromatography still form the mainstay of protein analyses, but now work alongside newer instruments such as microarrays and mass spectrometers [13].

The following is a discussion of some proteomic analytical techniques frequently used in drug-target isolation and identification. They are affinity-based and occur at the confluence of synthetic organic chemistry, cell biology, biochemistry and mass spectrometry. A unifying feature in the techniques is the use of a probe molecule to isolate a target-

*Address correspondence to this author at the Institute of Molecular and Translational Medicine Faculty of Medicine and Dentistry Palacky University in Olomouc, Hněvotínská 5, 779 00 Olomouc, Czech Republic; Tel: +420 585 632 082; Fax:+420 585 632 180; E-mail: marian.hajduch@upol.cz

[#]*These authors contributed equally to the paper.*

protein(s). Another such feature is the mass-spectrometric identification and quantification of isolated targets. Some of the techniques discussed here have been previously grouped under the descriptor Chemical Proteomics (a technique in which the probe bonds covalently with the target) [14, 15]. The common issues that accompany each technique are presented, as are the efforts made to get around them.

2. AFFINITY PURIFICATION

Affinity purification is a generic technique based on interaction between a ligand (probe) and protein of interest (target or target molecule). The technique was first developed in the early 1950s and has evolved since then to include a number of variants that have been successfully applied to molecular target identification [16].

A typical affinity purification experiment (Fig. 1) begins with the modification of the probe (a drug, drug-candidate or an inhibitor) and its immobilization on a solid phase matrix. To ensure that the probe's target specificity is retained after derivatization, the structure and activity of the modified probe is compared with that of the unmodified parent compound. The immobilized probe is then incubated with the

sample of interest, usually cell-lysate. Biomolecules in the sample that interact with the probe are captured during incubation. The resin is then washed free of non-specifically bound proteins and other "contaminant" molecules. Captured target proteins are eluted by a buffer that disrupts the ligand-protein interaction. These proteins are analyzed and identified by different techniques such as 1D or 2D gel electrophoresis, Western blotting and mass spectrometry. This approach is frequently used for studying protein interaction networks and in identification of drug targets [17]. The type of tag used for the purification varies with the purpose, with the scale of the purification and on the ease of preparation of the affinity matrix.

2.1.The Issue of the Non-Specific Interaction

An affinity purification or a protein-protein interaction is characterized by at least two parameters, sensitivity and specificity. These parameters indicate the "goodness" of the affinity purification technique and are necessary for the purpose of comparison of two or more techniques.

Sensitivity is a feature by virtue of which a true interaction is detected. It is denoted as a ratio of the number of true
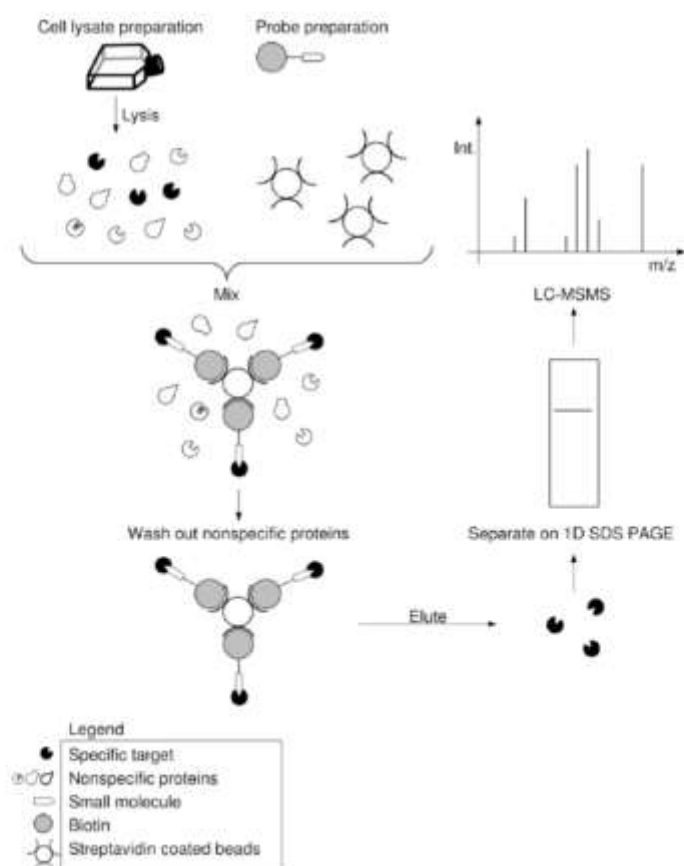
Fig. (1). Typical workflow of affinity purification: Modified (e.g. biotinylated) compound is immobilized onto solid-phase (streptavidin coated beads). Cell lysate is incubated with immobilized compound on the solid-phase. Specific proteins are captured by compound and non-specific ones are washed out. Target proteins are then eluted and subjected to 1D SDS PAGE and LC-MS/MS.

interactions that are detected to the total number of true interactions. Thus,

Sensitivity = TP/ (TP + FN)   (1)

Where, TP - True-positives are true interactions that are detected by the technique. FN - False-negatives are true interactions that are not detected by the technique

Specificity is a feature by virtue of which a false interaction does not happen and is not detected. It is expressed as a ratio of the truly non-interacting proteins to the total number of non-interacting proteins. Thus,

Specificity = TN/(TN + FP)   (2)

Where, TN = True-negatives are interactions that are not supposed to happen and are not detected. FP = False-positives are interactions that are not supposed to happen, but are detected

Equation 2 suggests that the specificity of a technique is an inverse function of the number of false-positives. The specificity of a protein interaction may also be indicated by a dissociation constant, K_D. The higher the K_D, the higher is the probability that the binding-site of the protein will be populated by its true binding-partner or ligand.

The primary challenge in affinity purification lies in isolating "real" target proteins while excluding non-specifically bound "contaminants", the false-positives. Interactions which escape detection because of the experimental conditions or because of a limitation of the technique are false-negatives. For example, probes that target post-translational modifications will be rendered ineffective if the modification

is absent or is lost in sample preparation. Steric hindrance of the probe or target protein by another molecular entity will also preclude a fruitful interaction. From equation 1 above, it is evident that as the number of false negatives decreases, there is a proportional increase in the sensitivity i.e. sensitivity is inversely proportional to the number of false negatives.

An intelligent choice of the affinity matrix can go some distance in alleviating the problem of non-specific interactions, as the matrix and probe linkers can be a source of such interactions [18, 19]. Matrices that are incompletely derivatized may also bind contaminants with high affinity. The proteins that bind non-specifically to the common matrix types (agarose, Sepharose etc) have been identified and listed for reference [20-22].

There are different ways of separating valid targets (true-positives) from non-specifically bound proteins (false-positives) and improving the confidence of discovery. Pre-incubating the sample with non-derivatized and inactive matrix before the actual experiment improves the purity of the isolated product. Serial incubation of cell-lysate with the affinity resin is yet another option (serial affinity-chromatography) [23]. Cell-lysate is incubated with two successive batches of affinity-resin. The first batch captures most of the true targets and the second, most of the non-specifically bound contaminants.

There are several ways of immobilizing a compound to a matrix (Fig. 3). The probe may be chemically bonded to the matrix via its functional group, for eg. a sulfhydryl, a hydroxyl, amino or carboxyl. This approach has been used for
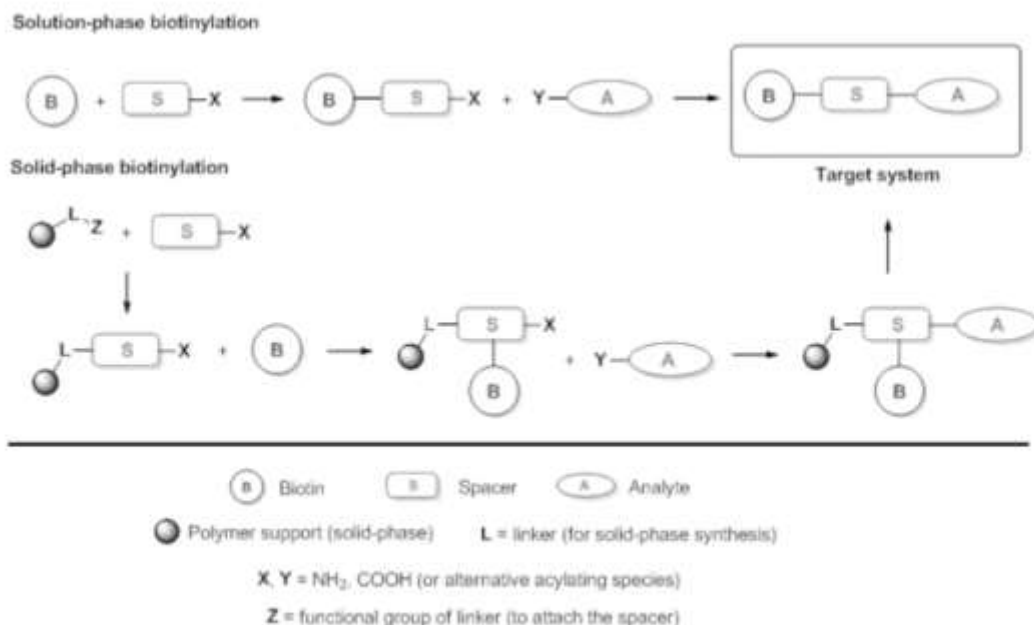
Fig. (2). Two alternative approaches to modify compound **A**: The selection of more suitable method is based on specific requirements such as required quantity of the target system, its stability and availability of analyte (for low quantities the solution-phase method is preferable). A typical polymer support is divinyl-styrene based resin. Except for amides also ester or ether groups are frequently used to attach individual parts of the target system.

PROBE	STRUCTURE	IMMOBILISING RESIN	PROPERTIES
Compound modified by NH₂, OH, SH or COOH moieties		Agarose, sepharose	Inexpensive, specificity is sometimes an issue
Biotinylated compound		Streptavidin coated matrix	Specificity of biotin-straptavidin bond is high and requires stringent elution conditions, spacer design is complicated
Photoaffinity label		Streptavidin coated matrix	Drug binds covalently to target, Photoaffinity probe needs to be carefully selected and derivatized to drug
Compound modified for click chemistry reaction		Streptavidin coated matrix	The reaction can be performed in a variety of conditions, Long exposures to Cuᴵ catalyst may be cytotoxic
Immuno-chemo-proteomics probe		Antiflag beads	Sulfation of FLAG polypeptide can alter antibody specificity
Kinobeads		Agarose beads	For kinases only. Some kinases do not bind to ATP structural analogs and cannot be enriched by Kinobeads
Activity based probe		Streptavidin coated beads, fluorescent or radiolabell scanning	For enzymes only, warheads aren't available for all enzyme types

Fig. (3). Examples of compound immobilization strategies.

the tyrosine kinase inhibitors imatinib, nilotinib and dasatinib. The inhibitors are first chemically treated to introduce amino and acetyl groups on their surface and are then immobilized on N-hydroxysuccinimide activated Sepharose via the same groups [7]. Commonly used matrices are polysaccharide-derived (agarose, Affi-Gel®, Sepharose), methacrylate-derived (Toyopearl®) or magnetic beads (Dynabeads®, SG beads) [24]. Affinity matrices are constantly being developed and improved, to enhance binding of the specific target and to minimize non-specific interactions. Their applications, strength and limitations are reviewed in detail by Sakamoto *et al.* [25].

A competitive-binding experiment, where the free probe competes with immobilized probe for the target, is another approach for determining spurious interactions [26, 27]. The Stable Isotope Labeling of amino Acids in Cell-culture, or SILAC, is an isotope labeling technique that is used for quantifying proteins in a sample by mass-spectrometry (described in section 3.1 below) labeled (heavy, H) and unlabeled (light, L) cell-lysates are incubated with the immobilized small-molecule probe, and the unlabeled lysate simultaneously incubated with non-immobilized probe. Following incubation, the probe-bearing beads are washed, mixed and the probe-protein complexes eluted. These are separated by SDS-PAGE and analyzed by LC-MS/MS. Specifically bound target proteins have a H/L ratio greater than 1, while the non-specific ones have an H/L ratio of 1 (Fig. **4**) [28]. The principle of competitive binding has been applied to the investigation of cellular kinases too. Kinases and phosphatases are frequently involved in regulation of different cellular processes because they phosphorylate and dephosphory-

late proteins. Because blocking a single kinase can perturb one or more cellular pathways kinases are considered to be potentially valuable drug-targets [29, 30]. The ATP binding domain of a kinase binds to ATP to hydrolyze it to the corresponding mono-phosphate form. Multiple ATP structural analogs fused to synthetic beads (KinobeadsTM) exploit this property to isolate specific tyrosine kinases involved in CML that are targets of BCR/ABL inhibitors [31, 32]. Cell-lysate is incubated with Kinobeads in the presence and absence of inhibitors. Tyrosine kinase binding to Kinobeads is competitively inhibited by the inhibitors. The difference in the levels of bound kinases in the two samples, as measured by quantitative mass-spectrometry, distinguishes the target kinases from the false-positives. The method must be used with caution though, as some kinases do not bind to Kinobeads, even though they bear an ATP binding domain. A similar approach has identified novel targets of the multi-kinase inhibitor E-3810 [33]. As with other proteins, protein kinases may also be isolated using their respective inhibitors as probes, for example, bisindolylmaleimide X, AX14596, PP58 and purvalanol [34].

There is another less evident problem that results in false positives. It is encountered in samples where the dynamic range of protein concentrations is high [14]. A protein that is not a true target may still bind to the probe by virtue of its high concentration, even though it has a low affinity (for the probe). As a result the probe captures some non-specific high-abundance proteins (HAPs) in addition to the true (low-abundance) target. One of the strategies for overcoming this problem is called Proteominer™ (BioRad) [35-37]. The technique is based on an exhaustive combinatorial library of
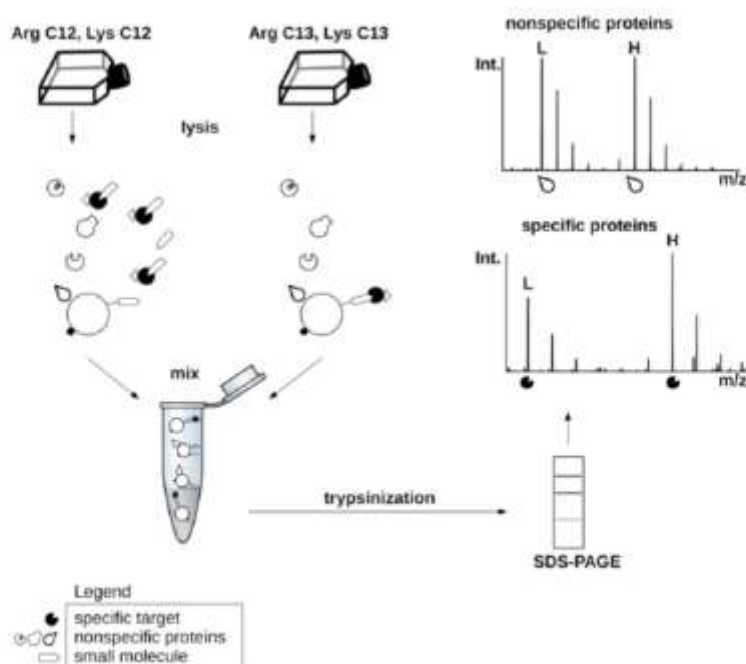
Fig. (4). Identification of protein targets by quantitative proteomics: Cells of interest are incubated with labeled (heavy, H) and non-labeled (light, L) aminoacids (SILAC). Non-labeled cell extract is simultaneously incubated with immobilized and free small-molecule. Heavy cell lysate is incubated with immobilized small-molecule alone. Afterwards, both bead-types are washed and mixed in equimolar quantities. Captured proteins are eluted, separated by SDS PAGE, analyzed by MS and quantified. Target protein has heavy to light ratio (H/L) above one and nonspecific protein has H/L ratio approximately one.

hexapeptides that can theoretically bind every possible protein in a sample. When sample is incubated with the ProteominerTM matrix, HAPs quickly saturate their binding sites and the surplus flows off, while the low-abundance proteins (LAPs) bind almost completely. This results in the depletion of HAPs and in the enrichment of LAPs, thereby effectively decreasing the dynamic range of protein concentrations in the sample by a few orders of magnitude.

Other, more conventional strategies for reducing the dynamic range of protein concentrations in serum utilize antibodies that target the most abundant proteins [38-40]. The relative rate of trypsin digestion of high- and low-abundance proteins is also a basis for improving the dynamic range of detection of proteins [41]. This technique, christened DigDeAPr, has been used with cell-lysates but can be extended to complex mixtures that have a wide range of protein concentrations [41].

2.2. Affinity Reagents that Bind Non-Covalently

2.2.1. Biotin and Streptavidin

The strong non-covalent interaction between biotin and streptavidin (or avidin) is the basis for a common type of affinity purification [42]. These interactions are the strongest known, non-covalent interactions occurring in nature ($K_D = 1*10^{-14}$). The probe is tagged with biotin and used to "fish" for interaction partners in the sample. The probe and the captured proteins can then be isolated on a streptavidin coated matrix via the biotin tag. In drug-target identification studies,

it is common practice to biotinylate the drug or inhibitor, as for instance, biotinylated 2-hydroxycinnamic acid (HCA). This drug binds to proteasome subunits (target) in lysates of SW620 colon cancer cell-line and inhibits the L3-like activity of the proteasome [43]. Biotinylated diazonamid can likewise isolate the mitochondrial enzyme ornithine delta-amino transferase from HeLa cell-lysate [44]. The primary downside to this purification strategy is that the strong streptavidin-biotin interaction requires stringent elution conditions that may negatively affect protein structure or function.

Key to the synthesis of the biotin label is the design of the spacer that connects the target molecule to the biotin moiety [18, 45-47]. Spacer design and its impact on the affinity purification has been intensively studied and some rules formulated [48-51]. The design must conform to these principles to perform optimally: (i) it must be sufficiently long to prevent steric repulsion between avidin and the target biomolecule; ii. it should not interact with the target biomolecule to cause a false-positive result ; iii. it must also not alter the solubility of the biotin probe in water.

Aliphatic spacers such as caproic acid derivatives have been used frequently in the past although the newer ethylene glycol (EG) spacers are now preferred, as they satisfy the above criteria better [52]. The compound of interest is typically attached to the biotin-EG system via a stable amide bond. It can be done either by acylation of the biotin-EGs-NH2 system by carboxy group-containing compounds or acylation of an amino group-containing compound with the biotin-EGs-COOH system. Derivatization has tradition-

ally been done using solution-phase synthesis and more recently using solid-phase synthesis. The substrate for biotinylation (or conversely the biotin-EG-NH2 system) is pre-immobilized or directly synthesized on an insoluble polymer backbone, and the resulting preloaded resin then used for reaction with an appropriate compound for e.g. the Biotin-PEG-NovaTag™ (Novabiochem) [53-56]. Conversely, the biotin-EG-NH2 system may be immobilized in place of the biotinylated substrate. Preparation of a similar system for immobilization of carboxy-group containing analytes via spacers of different length has also been described [57] (Fig. **2**).

2.2.2. The Recombinant Tag

Several commercially available tags can be cloned into the protein of interest, expressed in eukaryotic of prokaryotic cells and isolated by affinity-methods. The FLAG tag is a small hydrophilic octopeptide (DYKDDDDK) that works in this manner [58]. Recombinant proteins containing this tag may be over-expressed and isolated using anti-FLAG antibodies immobilized on a suitable support resin. Isomeric forms of protein kinase C have been identified using FLAG-tagged bisindolylmaleimide III (Bis III), an inhibitor of protein kinase C [59]. The FLAG tag was considered generally applicable till a few years ago, when it was shown that a sulfation event (a post-translational modification) eliminates the interaction of the FLAG tag and anti-FLAG antibodies [60]. Such a modification lowers the yield of the purification without ever becoming evident. The sulfation was observed in insect cells and is not known to occur in the CHO and HEK293T human cell-lines. The His-tag exhibits this problem to a degree and it is likely that other tag-systems suffer from it too. A systematic, comparative investigation of the various clonal tag-systems has not been reported so far, but potential savings in time and resources may make such a study worth investing in [61].

The His-tag is a sequence of 6-histidine residues that is also cloned into the protein of interest. Histidine residues bind with high affinity to divalent metal ions on a solid matrix (Ni^{2+}, or Co^{2+}) [62-64]. Bound proteins may be eluted by a high concentration of free histidine. The GST-tag works similarly [65]. GST is a 26-amino acid enzyme called Glutathione-S-Transferase and its binding ligand is its substrate, glutathione. Glutathione is a three-residue moiety (Glu-Cys-Gly) that is immobilized on a solid matrix via the reduced cysteine. Concentrated, reduced glutathione constitutes the elution reagent. Yet another clonal tag-system is the eight-amino acid Strep-tag system [66]. It is based on the streptavidin-biotin principle and binds to streptavidin or an engineered streptavidin matrix. However, it requires much milder elution conditions than the streptavidin-biotin interaction.

The tag-systems mentioned above are archetypal of the class of clonal affinity-based tags. Each system has its strengths, and because no single tag is appropriate for all purposes a protocol may utilize two or more in a purification process. The tag may sometimes not be specific enough or its affinity of interaction may require stringent elution conditions that are harmful to proteins. Larger tags impose a larger metabolic burden on the culture organism than the smaller ones and may not be ideal for use in eukaryotic cells. Addi-

tionally, an expensive resin that cannot be recycled is also not likely to be widely used [61].

2.2.3. Click Chemistry Based Probes

The structure of the drug/inhibitor in question may sometimes be amenable to modification by click-chemistry [67]. "Click" reactions are near-ideal chemical reactions that are typically easy to perform in different conditions, have high yields that are pure, and are not affected by the reactant groups. The Huisgen's cyclo-addition reaction is the classical example of this type of chemical reactions [68]. It is the 1,3- dipolar cycloaddition reaction between an azide and a terminal or internal alkyne that yields a 1,2,3-triazole. For instance, after incubation with the sample, alkyne derivatized Orlistat is made to react with rhodamine-azide and visualized on an electrophoretic gel by fluorimetry [69]. Alternatively, if the protein-drug complex is to be recovered, Orlistat-alkyne is coupled to an affinity-tag-azide. Click-chemistry may be used in exploring drug-binding sites in proteins as well and to investigate the off-target effects of drugs [69]. Because click-reaction conditions are flexible, they are easily performed under "native" conditions to avoid harming the biological sample. An alkyne group engineered into an ABPP or photoaffinity label can expand their downstream analytical and purification options [69].

2.2.4. Identification of Drug-Targets without Isolation

Target protein identification may sometimes be performed without isolating the drug-target complex [70-73]. One technique is based on the relative conformational stabilities of free- and small molecule-bound protein and is called Drug Affinity Responsive Target Stability (DARTS). Small molecule binding may stabilize the conformational stability of the protein, with the result that it becomes more resistant to denaturation and less susceptible to proteolysis [72]. Small molecule binding may also sterically protect a protease cleavage site. This property has been made use of in the identification of the FK506-binding protein 12 (FKBP12) and the mammalian target of rapamycin (mTOR) as targets of kinase inhibitor E4 and rapamycin respectively [74]. Cellular targets of grape-seed extracts have also been successfully screened by this technique [75]. Cell-lysate is incubated with the desired small molecule or an inactive structural analog. The mixtures are subject to protease digestion and the proteins/peptides subsequently resolved by electrophoresis and identified by LC-MS/MS. A difference in the intensity of a band on an electrophoretic gel is indicative of a drug-target interaction (Fig. **6**). There are other variants of this technique and these have been discussed at length in Lomenick *et al.* [70]. DARTS is conceptually simple and the technique does not involve derivatization of the drug or molecule-of-interest. However, as with other techniques, it is not without issues. The probability of non-specific interactions increases with the complexity of the protein mixture and cell-lysates are very complex mixtures. Moreover, the cell solubilization conditions are artificial. Protease cleavage sites are occasionally missed by the proteolytic enzyme because proteolytic digestion doesn't always go to completion and because some proteins are inherently resistant to digestion [72]. The proteolytic enzyme may also act on proteins complexed with the target-protein to further complicate results [31].
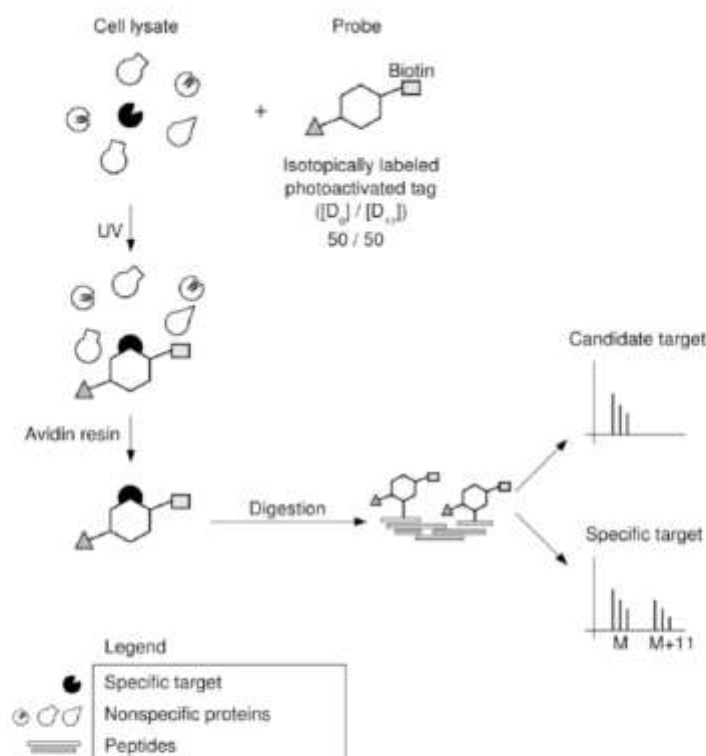
Fig. (5). Photoaffinity labeling with two isotopes: PAL probes consisting of an affinity tag (biotin) and an isotopically labeled photoreactive tag (non-deuterized, D_0 and deuterized benzophenone isotope label, D_{11}) are mixed in 1:1 ratio. Probes are incubated with cell lysate and subsequently irradiated by UV-light. Proteins are then purified via avidin resin and identified by MS. Identification of the unique isotopic pattern (M, mass of peptide with non-deuterized tag; M+11, mass of peptide with deuterized tag) of some peptides helps to distinguish valid protein targets from frequent contaminants.

2.2.5. Decreasing Sample Complexity: Capturing the Sub-Proteome

The affinity purification procedure may be designed to capture not just one protein but an entire class of proteins. The N-glycoproteome, phosphoproteome or the SUMOy-lated group of proteins are important druggable PTMs in cancer and have suitable purification procedures [76-79]. The probe, in this case, binds non-specifically to all proteins bearing a particular post-translational modification (PTM).

Antibodies against phosphate group epitopes are commonly used reagents for isolation of phosphorylated proteins [80]. However, because they are specific to an epitope, they allow enrichment of only one kind of phosphorylated peptide at a time. This problem is circumvented when the purification is based on phosphotyrosine, phosphoserine or phosphothreonine antibodies or a non-specific electrostatic affinity between the negatively charged phosphate groups and positively charged metal cations (Immobilized Metal Affinity Chromatography IMAC and by Metal Oxide Affinity Chromatography, MOAC) [80-82]. The negative charge on the free carboxyl groups and acidic amino acids must be neutralized for the isolation of peptides with phosphate groups by appropriate affinity resin or by conventional ion-exchange chromatography (hydroxyapatite column enrichment or Hydrophobic Interaction Liquid Chro-

matography, or HILIC) [83]. Conversely, the phosphate group itself may be derivatized to improve its binding to the affinity resin. The derivatization step is not always desirable because it introduces an additional step in the sample preparation workflow, and because the reaction does not always go to completion. Furthermore, phophorylated proteins or peptides can be visualized on a polyacrylamide gel using a fluorescent stain, Pro-Q Diamond. Fluorescent bands are then cut out and prepared for identification by mass-spectrometry [84].

In some cases, it is possible to reverse the affinity purification technique such that the purified target-protein is given the role of a probe to screen for active compounds. Desorption/Ionization on Silicon, or DIOS, is one such method. DIOS is now known by its commercial name NALDI™ (Nano-Assisted Laser Desorption and Ionization) [85-87]. It is a matrix-free method and uses laser desorption and ionization of compounds from a porous silicon surface to directly analyze probe-target interactions by mass-spectrometry. This is illustrated in the investigation of Bovine Serum Albumin (BSA) binding partners by Zou *et al.* [88]. BSA was immobilized on a silicon surface (wafer) and incubated with a mixture of two molecules, ketoprofen and sulpiride. The silicon surface was then washed to remove unbound compound and subject to direct NALDI

analysis. The analysis yielded a signal for ketoprofen but none for sulpiride, confirming that ketoprofen binds to BSA and that sulpiride does not.
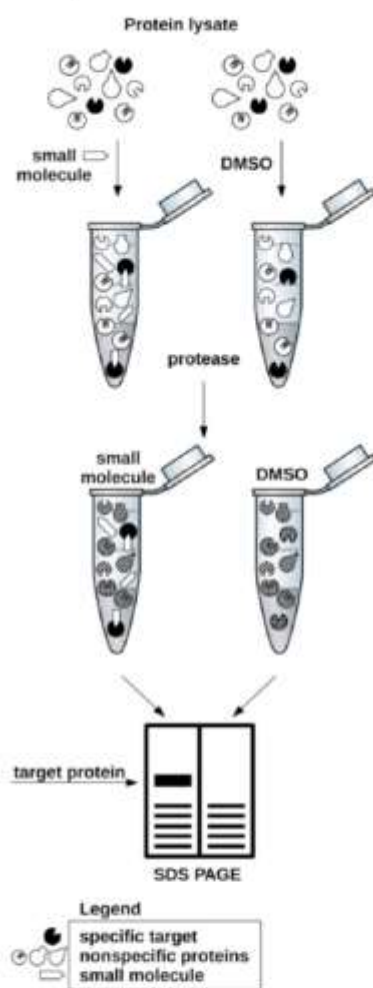
Fig. (6). DARTS. Protein lysate is incubated with a small-molecule compound and then subjected to protease digestion. Target proteins interacting with the small-molecule compound are enriched and protected against protease digestion while nonspecific proteins are digested.

2.3. Affinity Reagents that Bind Covalently

The affinity purification techniques described so far depend on a physical interaction between the probe and its target-protein. However probes may also be engineered to form chemical bond(s) with their targets to increase their specificity and sensitivity.

2.3.1. Photo-Affinity Labeling (PAL)

The photoaffinity probe (PAP) typically consists of three components: a drug/ligand that interacts specifically with the target, a photoreactive group that forms the chemical bond with the target, and at least one reporter group (Fig. **5**) [89,

90]. The probe is incubated with the target and then irradiated with light of appropriate wavelength. The probe binds to its target and upon irradiation, the photoreactive group forms a chemical bond with the same. Photoaffinity probes are classified into roughly three types on the basis of their photoreactive group, although there are probes that fall outside this classification: benzophenones, aryl azides and diazirine. All three kinds of PAPs are used in the (high-throughput) identification of drug-leads and for the investigation of the drug-target interaction [91]. Ovalicin and fumagilin tagged with a bifunctional-PAP have been used in cellular target screens [92-94]. Labeling ideally requires that (i) the PAPs be inert in the absence of actinic light, (ii) the activating light not be of a harmful wavelength, lest the protein-target be damaged, (iii) the lifetime of the activated photoreactive moiety be smaller than the lifetime of the drug-target interaction (iv) the photoreactive and the reporter groups not get in the way of the drug-target binding and (v) that the adduct formed during the reaction be stable enough to withstand the rigors of the downstream purification or analytical process. The caveat with this technique is that the binding profile of a PAP is different in live cells and in cell-lysates.

In a variation on this principle, the photo-reactive group can be made with one (3H) or two (^{1}H, ^{2}H) different isotopes [95, 96]. When the two isotopic-PAPs are used in equimolar quantities in the same purification, they yield unique isotopic patterns in a tandem mass spectrum that can help discern valid protein targets from false-positives.

2.3.2. Activity-Based Protein Profiling (ABPP)

Activity-based protein profiling relies on chemical bond-formation for target enzyme discovery. The probe here consists of three parts, a "warhead" which can bind covalently with a functional group in the active-site of the target enzyme, a tag for the visualization or purification of target proteins, and a linker joining these two parts [97-99]. The tag is a fluorophore, a biotin group or a click-chemistry group (described in section 2.2.3). On incubation with cell-lysate the probe binds to target proteins and bound enzymes can be purified using the biotin-group, visualized by electrophoresis and identified by mass spectrometry. ABPP can be scaled for profiling proteins, and by extension their source tissues, on the basis of their biological activities. This isn't possible with other proteomic techniques such as PAGE or LC-MS. Multiple ABPP probes have been used to comprehensively map the binding-surface of a known target in order to determine a greater number of specific binding sites [100]. The different variants of orlistat (tetrahydrolipstatin, THL), THL-R, THL-L and THL-T have been used to determine orlistat's possible off-target effects [69].

2.3.3. Cell-Permeable Probes

In some cases it is desirable to have a drug/probe that is cell permeable, to monitor and identify drug-target interactions in the native state [101, 102]. Orlistat and several other ABPP, photo-affinity and click-chemistry probes have this property. The HIV Trans-activating transcriptional activator (TAT) is a nine-residue peptide (RKKRRQRRR) with a preponderance of positively charged amino acids, and can transport molecular cargo into eukaryotic cells [101, 103]. The cargo could be metabolites, peptides, proteins or nucleic

acids. The TAT peptide could potentially also transport some of the above mentioned affinity tags into the cell in studies of drug-target interactions. A composite probe consisting of the drug (Bis III), a fluorescent group and the TAT peptide has been used, for instance, for this purpose [104]. Other cell-penetrating peptides are now available and may be similarly used in future for the entire class of affinity-tags [105]. A provision for tag cleavage and removal, as by incorporation of a thrombin cleavage site is another desirable feature of a tag-based system [106]. This is important when the isolated target is to be analyzed by mass-spectrometry without interference from the affinity tag.

3. PROTEIN QUANTITATION IN DRUG-TARGET DISCOVERY

It is evident from the preceding section that drug-target identification frequently requires the relative or absolute quantification of one or more proteins in multiple samples. Protein identification and quantification is performed downstream of an affinity purification. The sample preparation protocol, the mass-spectrometer and the quantitation procedure (labeled samples or label-free; Fig. **7**) is dictated by the purpose of the experiment. Proteins isolated from the affinity purification stage may directly be used for preparation of MS samples, or may be further resolved by 1- or 2-dimensional electrophoretic separation before being processed for MS. Procedures for isolating proteins from Western blot transfer membranes have also been developed [107].

3.1. Quantification Using Isotopic Labels

Quantification has traditionally been done using isotope tags. Common tagging protocols include the Stable Isotope Labeling with Amino acids in Cell culture (SILAC), Isotope Coded Affinity Tags (ICAT), Tandem Mass Tags (TMT) and

the Isobaric Tag for Relative and Absolute quantification (iTRAQ) [108-111]. SILAC and ICAT are isotopic labeling protocols that use isotopes of hydrogen, carbon and nitrogen (^{2}H, ^{13}C, ^{15}N). They are based on a mass-difference between the light and the heavy isotope labels. The mass difference causes a shift in the peaks of the labeled peptides relative to the unlabeled peptide. The intensities of the two peaks indicates the relative quantities of the two peptides.

SILAC involves the labeling of cells in culture in minimal media. Cells are grown in media containing normal or labeled amino acids. Labeled amino acids are incorporated into cellular proteins via natural biosynthetic pathways. Five-ten cell-doublings later, the two cell samples are harvested, mixed, processed and trypsinized before being analyzed by mass spectrometry. Because samples are labeled during cell culture the method is suitable only for samples derived from culture and not for samples from other sources. Samples that cannot be labeled in culture may be labeled after collection using the ICAT, TMT or iTRAQ labeling procedures. The extent of labeling can be more precisely controlled in these post collection labeling techniques since it is not dependent on the *in vivo* protein metabolic rate or the cell doubling-time. The SILAC mass-difference concept, in most setups, restricts the number of samples that can be simultaneously analyzed to two. For this reason, the method cannot be used in experiments with multiple samples. Different ionization states of a peptide will alter the observed mass difference between the heavy and light peptide. Additionally, the minimal medium itself may perturb the cellular physiology such that the observed biological effect is not completely true. Deuterium-labeled amino acids create another issue. The deuterium labeled peptide migrates ahead of the unlabeled peptide in the reverse-phase chromatographic column and changes the peptide's elution profile. This causes a split in the MS1 peak of the peptide and makes quantification less
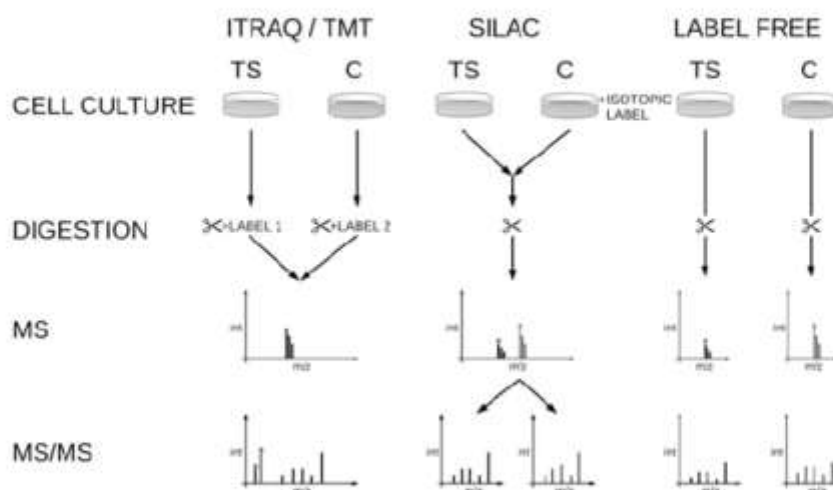
Fig. (7). A principle of mass-spectrometry peptide-labeling techniques: iTRAQ/TMT labeling is based on isobaric tag and quantification is done at the MS/MS level. SILAC labeling is based on metabolic incorporation of isotopic label into cell culture and quantification is done at the MS level. This is ideal for high resolution mass spectrometry where MS has usually higher accuracy than MS/MS. However, SILAC is more or less limited to cell cultures. Label free techniques rely on peak-intensity, spectral-count, peptide-count and/or fragment-ion intensity for quantification of protein.

accurate. Isotopes of ^{13}C and ^{15}N do not cause this problem and are therefore most commonly used. These isotopes also give rise to a larger mass-difference that can be more accurately measured.

ICAT resolves some of the issues of SILAC [109]. The ICAT tag consists of three groups- a thiol-reactive group, an isotope coded tag and a biotin moiety. A cleavable linker is sometimes included between the isotope tag and the biotin to facilitate purification [112]. The two collected samples are reduced and trypsinized and incubated with ICAT labels to allow thiol reactive groups to react with the reduced thiols on cysteine residues. Differentially labeled samples are then mixed and subject to mass-spectrometry. The property of thiol specificity is at once useful and problematic- useful because tagged peptides can be purified, and problematic because peptides lacking cysteines are lost [113]. Furthermore, the accuracy of the protocol depends on the labeling efficiency. Less than optimal labeling will bias the quantification and make the results unreliable.

3.2. Quantification Using Isobaric Labels

Labeled quantification evolved further with the introduction of isobaric tags called Tandem Mass Tags (TMTs) and isobaric tags for relative and absolute quantification (iTRAQ) [110,111]. Isobaric tagging allows the simultaneous analysis of more than two samples and does not require the isolation of tagged peptides.

TMTs consist of at least three groups, a reporter group, a mass normalization group and a protein reactive group. These tags are amine reactive and are therefore more likely to have a higher efficiency of labeling than ICAT tags. TMTs are available in sets of two, six or ten labeling reagents (2-plex, 6-plex and 10-plex respectively) [110, 114]. Tags of a particular set have identical overall mass. For this reason, two identical peptides tagged with different TMTs will co-elute from a reverse-phase column and have the same MS^1 spectral peaks. Depending on the labeling reagent, two, six or ten samples can be simultaneously analyzed by MS. The control and treated sample sets are trypsinized, labeled, pooled and analyzed. Upon peptide fragmentation of selected MS^1 precursor-ions by collision-induced dissociation in the mass spectrometer, the MS^2 spectrum exhibits the reporter group peaks in addition to the characteristic peptide fragmentation peaks. The intensities of the different reporter group ions indicate the relative quantities of the peptide in the different sample sets. Untagged peptides will not have reporter-group peaks in the MS^2 and will therefore be disregarded in the quantification process. This makes isobaric tagging intrinsically more accurate than the ICAT labeling procedure described above. TMT has been used in the determination of targets of antidepressant drugs [115].

One assumption here is that the applied m/z window selects only a single desired precursor-ion for fragmentation. However, in reality, additional precursor-ions within a 2kDa range of the desired ion are also fragmented. Reporter groups from these additional precursors then lower the quantification accuracy [116]. Another assumption is that the different isobaric tags label proteins in the different treatment groups (samples) with equal efficiency. But it may not necessarily be so. The nature of the peptides in the samples and possibly their concentration, influences the derivatization reaction and affects quantification [117]. The above description of tandem mass tags is also true of the iTRAQ tags, except that iTRAQ is also meant for absolute quantification. iTRAQ tags come in 4-plex and 8-plex reagent kits and enable the use of up to 8 samples [111, 118].

3.3. Quantification by Selected Reaction Monitoring (SRM)

Absolute quantification is now also possible with AQUA peptides (acronym for Absolute Quantification and with SRM (Selected Reaction Monitoring) [119-122]. SRM is the detection and quantification of specific proteins in a sample. It is based on the prior knowledge of one or more proteotypic peptides from the protein of interest. A proteotypic peptide is one which is always formed upon protein digestion and has complete or near complete ionization. The mass-to-charge ratio of a peptide's precursor- and fragment-ions and its chromatographic retention time are characteristic of the peptide and together constitute a "transition" or "assay". There may be more than one transition per peptide. The mass-spectrometer is directed to follow pre-specified transitions for a specified duration and to measure their intensities. The quantity of the peptide may be calculated from a standard curve of an identical labeled peptide (intensity versus concentration). Reliable SRM assays for most of the human, yeast and bacterial proteins are now publicly available in the SRMAtlas, an online repository [123].

SRM can precisely measure multiple proteins in a mixture of proteins on the basis of pre-supplied transition information. Transitions, of the order of 10^3, can be reliably measured in a single SRM run, boosting throughput incrementally and solving a long-standing problem in proteomics [124]. Statistical design of experiments employing SRM may be guided by a statistical software devoted to this purpose called SRMstats [125].

3.4. Label-Free Quantification

The label-free approach is a more recent quantification approach that dispenses with the labeled standard. This approach has evolved sufficiently for routine use in different studies [126-131]. It is generally agreed that the theoretical basis for such quantification is a "work-in-progress" but the convenience of the label-free approach outweighs any small loss in accuracy. The technique also does away with the labor and cost of the labeling step and can be applied to multiple datasets and to datasets retrospectively.

Bottom-up proteomics consists of piecing together protein sequence information from constituent peptides. Protein samples are processed and digested into peptides. The peptide mixture is resolved by reverse-phase liquid chromatography, ionized and injected into the mass-spectrometer. These peptide-ions, also called precursor-ions, are scanned and their intensities plotted as a function of chromatographic retention-time (Fig. S1). The plot is called the total ion chromatogram. A peak or a specific retention time range can be selected and the intensity of the constituent peptides plotted as a function of their mass-to-charge ratio (m/z). This graph is the MS^1 spectrum and each peak in the MS^1 spectrum represents a precursor-ion. Desired precursor-ions are selected for frag-

mentation. The fragmentation process breaks up precursor-ions into smaller peptides called fragment-ions and their intensities are graphed as a function of their m/z values in the MS^2 spectrum. For a given energy applied to the fragmentation, a precursor-ion will always yield the same fragment-ion(s). The different fragment-ions in the MS^2 spectrum are compared with peptides in a database using their masses to derive their sequence and to identify the protein.

The initial plot of peptide-ion intensity versus retention time may be used to determine the quantity of peptide in a particular mixture. The Area Under Curve (AUC) and Percentage Area (PA) are two techniques that use this parameter [122, 129, 132-136]. The latter is called by different names depending on the normalization algorithm it uses: Total AUC, PA, AUC_{STD}, or R_{PA}. Chromatographic peak intensity and peak areas exhibit a good correspondence with protein abundance but the calculation is made difficult by noise and by peptides that co-elute or elute close to each other.

A tally of MS^2 spectra from all peptides of a protein is also a measure of a protein's abundance [126, 129, 131, 137, 138]. The tally is called the spectral count. The Normalized Spectral Abundance Factor (NSAF), $\log_2$(protein ratio), and R_{SC} all use this feature. Appropriate normalization must be applied to correct for differences in ionization efficiency though, which is frequently not ideal. Neither is detection linear across an indefinite range of protein concentrations. Additionally, because identification and sequencing is done using peptide masses, an artefact can skew the measurement. The artefact is any peptide that has the same mass as the peptide of interest but not the same sequence.

Alternatively, the number of unique peptides identified from a protein, or the peptide count, also indicates its abundance [126, 138]. Peptide and spectral counts show a good correlation with protein abundance but may tend to overestimate the abundance, especially when low spectral counts are observed.

The above quantification techniques use a single mass-spectral parameter for measuring protein quantity in a sample, either the spectral count (an MS^2 feature) or the chromatographic peak area. A more recent approach for measuring quantity uses a Normalized Spectral Index (SI_N) that is derived from the peptide count, the spectral count and the fragment ion intensity (an MS^2 feature). In comparisons of the different label-free techniques, the SI_N outperforms the others indicating that multiple features yield a more consistent and reliable measure of abundance than a feature alone [139, 140]. One or more of the above methods are now integrated into computational workflows that analyze tandem mass-spectra [140].

It is important to note that inadequacies in data acquisition can bias label-free quantification such that a biological effect is observed. The effect may not be real. It is imperative that certain standards be rigorously applied to data acquisition before the data can be considered worthy of label-free analysis. Individual (shotgun-proteomics) runs on a mass-spectrometer typically have only a 30% overlap. For this reason, multiple (5-10) runs are required (regardless of the instrument) to achieve a 95% analytical completeness of the data before the data can be used [139, 141].

4. IN PERSPECTIVE

It is important to view the drug-target discovery process in the larger context of rational drug-development. This involves the identification of one or more components of the relevant dysfunctional process. Such a component is often expressed at different levels in the normal and diseased tissue, or exhibits different levels of activity and constitutes a druggable target [5, 142-144]. The target may be a protein, nucleic acid or metabolite. Of these, proteins constitute the majority and an estimate puts the number of druggable protein targets at a few hundred [5].

4.1. The Genomic Approach, or the Proteomic?

Knowledge of the aberrant component can be obtained by examining its DNA, RNA or protein. These are complementary sources of information and the aberration can be explained more completely be considering them all. For instance, protein sequence alterations may be accurately determined by mRNA sequencing [145]. However, RNA-sequencing is of little use when it comes to investigating protein modifications that occur post-translationally. These require protein identification tools.

The information from nucleic acid is sometimes at odds with information derived from protein. In other words, there isn't always a correspondence between the RNA and protein expression profiles and the relation is yet to be completely understood [146].

The speed of the transcriptional process, the nature of the mRNA transcript, the changes wrought on the transcript post-transcription, its stability and translation into protein, the modifications subsequently made to the protein, the protein's stability and its turnover are all determinants of the relative levels of RNA and protein [146]. These RNA and protein profiles are "read" by the genomic and proteomic technique respectively, and any lack of correspondence between the two species may be attributed to one or more of the above reasons. For instance, a gene whose mRNA is unstable but is copiously translated into stable protein during its lifetime, will exhibit a large discrepancy between its RNA and protein levels. Conversely, an aberrant gene that is constitutively expressed is likely to have high levels of both mRNA and protein.

Technique has a role too in the incongruity. The method of tissue collection and storage, the sample processing protocol and the instrument influence the quantification. A large cross-platform study, much like the first NCI-CPTAC conducted study for proteomics, would go a long way in standardizing experimental conditions for the measurement of RNA and protein [147].

4.2. Validating the Target

The drug-target candidates from the different discovery approaches are collated and put through a rigorous validation procedure. Target validation is meant to confirm the role of the discovered target in the disease phenotype. It must be demonstrated in this phase that the drug-candidate effects the intended change and that it has the desired therapeutic effect [148]. The drug-candidate may cause a change in the expression level of different proteins, in a modulation of enzyme

activity, in a disruption or stabilization of protein-protein or protein-DNA interaction(s) or in the incorporation of post-translational modifications in proteins [5, 143, 144, 149]. The validation process is guided by the technique of target discovery and is normally performed in the cell- or animal models [148]. Validation is followed by lead-discovery and optimization, and then by preclinical tests, before the drug-candidate can graduate to a clinical trial.

The validation and subsequent lead development and optimization process is lengthy and expensive [150]. Only a small percentage of drug-targets pass the validation stage. It has been estimated that only three out of sixty new drug-candidates identified qualify as a marketable drug [151]. The total cost of drug-development- from target identification to market can now cost anywhere between USD4 billion and USD14 billion [151-154]. It does not help that drug-targets reported in the literature are not reliable candidates for development. Upto 60% of the relevant biomedical literature may have results that are irreproducible [155-158]. The problem is acute and some pharmaceutical companies now routinely conduct preliminary checks of reported targets before committing valuable resources to their validation and to drug-development. Discovery appears to influence overall research-and-development productivity more than anything else [153]. It makes sense therefore to improve the confidence of target-discovery such that valuable time and resources are not wasted in the validation and post-validation stages.

Increasing the specificity of the affinity-purification method such that the number of false-positives is lowered would go some distance in increasing the confidence of target identification. More specific affinity-capture reagents would help here, as would a greater number of them. To this end, the Human Proteome Project is tasked with developing antibodies for each of the 20,300 proteins coded by the human genome [159]. More rigorous standards and guidelines for investigating and reporting drug-targets would stem the decline in experimental reproducibility and facilitate drug development. Such standards have already been proposed for biomarker development and have helped [160, 161].

4.3. The Target as Part of a System

As stated at the outset, a comprehensive knowledge of the in-vivo interaction behavior of the drug-candidate can guide the design of a drug that has a more defined target or targets. There is a growing realization that biological systems, networks and modules within the cell must be considered in their entirety and that it is not enough to examine components in isolation. The observed effect of a drug is the sum total of interaction of multiple cellular and physiological systems. Systems biology has contributed to this realization. This is a relatively new discipline but has become indispensable in the drug-development enterprise in recent years. It deals with the relationships between the different components of a system to determine how they interact and how they function as a whole [162, 163]. It enables the integration of genomic and proteomic data with environmental data to better describe the organism's phenotype and to predict its response to a stimulus.

Systems biology has spawned computational tools for viewing a potential target in the larger cellular or physiologi-

cal context and to increase the confidence of target-identification. They include MetaCore[TM] (Thomson Reuters), Pathway Studio (Elsevier), IPA (Ingenuity Systems), the Software Tool for Researching Annotations of Proteins (STRAP), Gene Map Annotator and Pathway Profiler (GeneMAPP), Pathway Tools and the Database for Annotation, Visualization and Integrated Discovery (DAVID) [87, 164-166]. Various companies also offer in-house software for drug-target discovery, lead discovery, drug-repositioning etc. on a contractual basis.

4.4. The "Promiscuous" Drug

A review of all FDA approved therapeutics/imaging agents from the period 1998-2009 shows that most of the first-in-class drugs were those discovered by conventional phenotypic screening, and not by target-based drug-discovery [3]. Clearly, there is scope for improvement in the target-discovery approach.

Drugs frequently interact with unintended cellular components and have unpredicted and unexpected effects [167]. The polypharmacological effects of drugs were discussed in an earlier section as a rationale for a "more complete knowledge of drug-targets". Adverse side-effects are usually detected in late-stage clinical trials when the drug is administered to the patient. As the pharmacokinetics and dosage of the drug has already been tested prior to this, it makes sense to find a new use for the drug and salvage some of the developmental effort, than start all over again [168]. Ever-increasing drug-development costs can be partly offset by such "repurposing" of failed drugs [169]. Repurposing would also supplement the dwindling supply of drugs to the market [169].

Even if the drug has beneficial, if off-target effects, it is worthwhile to define and characterize the target-set. Living systems have redundant signaling pathways for any given task and this is especially true of cancer cells. The "plasticity" in the signaling network enables cancer cells to switch to alternative signaling routes when a particular component or pathway is blocked, and is the cause of their resistance to therapy [170]. Overcoming signaling "plasticity" requires that the redundant pathways be simultaneously blocked and is the reason polytherapy is preferred to monotherapy in cancer treatment. Likewise, a drug that binds to multiple cellular targets, with low affinity, may have a greater potential for safely curing a disease than one that binds "tightly" to a single target [171].

4.5. Advancing the Science

A road-map for development of proteomic tools and reagents has been laid out in an NIH-conducted workshop and seeks to partially address this question [172]. The workshop put forward feasible goals in proteomics that include (i) increasing the sensitivity of mass-spectrometry by 100- or a 1000-fold (ii) development of novel affinity-capture reagents that are more specific (in their interaction with bait) and easy and inexpensive to prepare (i.e, a general improvement in affinity-resins) and (iii) software that is better able to separate signal from noise in mass-spectral data, and that is less susceptible to over-fitting. Recommendations for advancing the cutting edge in systems-

biology-enabled medicine have also been made in workshops and seminars in related areas [173].

4.6. Big Data

With the coming of next-generation sequencing tools and precision mass-spectrometry has come an information deluge and the need for computational tools for storage, mining and analysis of this data. The need has been recognized but the development of software is still not commensurate with the advancements in instrumentation. The ideal software would analyze mass-spectrometric data, collate related information from scientific literature and databases and prepare a report that is of value to the scientist and the medic. This is not feasible yet, although a vision of information-enabled medical science that is tailored to an individual has been articulated [174].

It is sobering to recall that the rational drug-discovery process yielded only a minority of the FDA approved drugs in the decade gone by [175]. The biomedical community must step up to the challenge of improving that score, that the promise of rational drug-design may be realized.

ABBREVIATIONS

FDA	=	Food and drug administration
PML	=	promyelocytic leukemia
HDAC6	=	Histone deacetylase 6
SPT	=	Serine Palmitoyl transferase
SG	=	Styrene and glycidyl methacrylate
SILAC	=	Stable Isotope labelling of amino acids in cell-culture
H	=	Heavy isotopic label
L	=	Light isotopic label
SDS-PAGE	=	Sodium dodecylsulphate polyacrylamide gel electrophoresis
LC	=	liquid chromatography
MS	=	Mass spectrometry
ATP	=	Adenosine triphosphate
CML	=	Chronic myeloid leukemia
HAPs	=	high-abundance proteins
LAPs	=	low-abundance proteins
HCA	=	2'-hydroxycinnamaldehyde
EGs	=	Ethyleneglycols
PEG	=	Polyethyleneglycol
BIS III	=	bisindolylmaleimide III
GST	=	Glutathione S tranfrease
DARTS	=	Drug affinity responsive target stability
FKBP12	=	FK506 binding protein 12
mTOR	=	Mammalian target of rapamycin
PTM	=	Posttranslational modification

SUMO	=	Small ubiquitin-like modifier
IMAC	=	Immobilized metal affinity chromatography
MOAC	=	Metal oxide affinity chromatography
HILIC	=	Hydrophobic interaction liquid chromatography
DIOS	=	Desorption/ionization on silicon
NALDI	=	Nano-assisted laser desorption/ionization
BSA	=	Bovine serum albumin
PAP	=	Photoaffinity probe
PAL	=	Photoaffinity label
ABPP	=	Activity based protein profiling
THL	=	Tetrahydrolipostatin
HIV	=	Human immunodeficiency virus
TAT	=	Trans-activator of transcription
ICAT	=	Isotope code
TMT	=	Tandem Mass Tag
iTRAQ	=	Isobaric tag for relative and absolute quantitation
SRM	=	Slected reaction monitoring
AQUA	=	Absolute quantification
AUC	=	Area under curve
PA	=	Percentage area
NSAF	=	Normalized spectral abundance factor
Rsc	=	Ratio of spectral counts
SI_N	=	Normalized spectral index
DNA	=	Deoxyribonucleic acid
RNA	=	Ribonucleic acid
mRNA	=	Mediator RNA
NCI-CPTAC	=	National cancer institute - Clinical proteomics tumor analysis consortium
USD	=	United states dollars
IPA	=	Ingenuity Pathway Analysis
STRAP	=	Software tool for researching annotations of proteins
GeneMAPP	=	Gene Map Annotator and Pathway Profiller
DAVID	=	Database for Annotation, Visualization and Integrated Discovery
NIH	=	National Institutes of Health

CONFLICT OF INTEREST

The authors confirm that this article content has no conflict of interest.

ACKNOWLEDGEMENTS

AUTHOR CONTRIBUTIONS

Conception and design of the article (MH, PD), drafting (GR, TO, DH, LV, JH, MS) and/or critical revision of the content (LV, MH, PD), final approval to be published (MH).

SUPPLEMENTARY MATERIALS

Supplementary material is available on the publishers web site along with the published article.

REFERENCES

[1]   Drews J. Drug discovery: a historical perspective. Science 2000; 287(5460): 1960-4.
[2]   Lindsay MA. Target discovery. Nat Rev Drug Discov 2003; 2(10): 831-8.
[3]   Swinney DC, Anthony J. How were new medicines discovered? Nat Rev Drug Discov 2011; 10(7): 507-19.
[4]   Overington JP, Al-Lazikani B, Hopkins AL. How many drug targets are there? Nat Rev Drug Discov 2006; 5(12): 993-6.
[5]   Imming P, Sinning C, Meyer A. Drugs, their targets and the nature and number of drug targets. Nat Rev Drug Discov 2006; 5(10): 821-34.
[6]   Yang L, Wang K-J, Wang L-S, et al. Chemical-protein interactome and its application in off-target identification. Interdiscip Sci Comput Life Sci 2011; 3(1): 22-30.
[7]   Rix U, Hantschel O, Dürnberger G, et al. Chemical proteomic profiles of the BCR-ABL inhibitors imatinib, nilotinib, and dasatinib reveal novel kinase and nonkinase targets. Blood 2007; 110(12): 4055-63.
[8]   Ximenes JCM, de Oliveira Gonçalves D, Siqueira RMP, et al. Valproic acid: an anticonvulsant drug with potent antinociceptive and anti-inflammatory properties. Naunyn Schmiedebergs Arch Pharmacol 2013; 386(7): 575-87.
[9]   Petrelli A, Giordano S. From single- to multi-target drugs in cancer therapy: when aspecificity becomes an advantage. Curr Med Chem 2008; 15(5): 422-32.
[10]   Sheskin J. Thalidomide in the treatment of Lepra reactions. Clin Pharmacol Ther 1965; 6: 303-6.
[11]   Kawamori T, Rao CV, Seibert K, Reddy BS. Chemopreventive activity of celecoxib, a specific cyclooxygenase-2 inhibitor, against colon carcinogenesis. Cancer Res 1998 1; 58(3): 409-12.
[12]   Siow D, Wattenberg B. The histone deacetylase-6 inhibitor tubacin directly inhibits de novo sphingolipid biosynthesis as an off-target effect. Biochem Biophys Res Commun 2014 4; 449(3): 268-71.
[13]   Washburn MP, Wolters D, Yates JR. Large-scale analysis of the yeast proteome by multidimensional protein identification technology. Nat Biotechnol 2001; 19(3): 242-7.
[14]   Rix U, Superti-Furga G. Target profiling of small molecules by chemical proteomics. Nat Chem Biol 2009; 5(9): 616-24.
[15]   Jeffery DA, Bogyo M. Chemical proteomics and its application to drug discovery. Curr Opin Biotechnol 2003; 14(1): 87-95.
[16]   Campbell DH, Luescher E, Lerman LS. Immunologic Adsorbents: I. Isolation of Antibody by Means of a Cellulose-Protein Antigen. Proc Natl Acad Sci USA 1951; 37(9): 575-8.
[17]   Azarkan M, Huet J, Baeyens-Volant D, Looze Y, Vandenbussche G. Affinity chromatography: a useful tool in proteomics studies. J Chromatogr B Analyt Technol Biomed Life Sci 2007; 849(1-2): 81-90.
[18]   Shiyama T, Furuya M, Yamazaki A, Terada T, Tanaka A. Design and synthesis of novel hydrophilic spacers for the reduction of non-specific binding proteins on affinity resins. Bioorg Med Chem 2004; 12(11): 2831-41.
[19]   Von Rechenberg M, Blake BK, Ho Y-SJ, et al. Ampicillin/penicillin-binding protein interactions as a model drug-target system to optimize affinity pull-down and mass spectrometric strategies for target and pathway identification. Proteomics 2005; 5(7): 1764-73.
[20]   Trinkle-Mulcahy L, Boulon S, Lam YW, et al. Identifying specific protein interaction partners using quantitative mass spectrometry and bead proteomes. J Cell Biol 2008; 183(2): 223-39.
[21]   Chen GI, Gingras A-C. Affinity-purification mass spectrometry (AP-MS) of serine/threonine phosphatases. Methods 2007; 42(3): 298-305.
[22]   Gingras A-C, Caballero M, Zarske M, et al. A Novel, Evolutionarily Conserved Protein Phosphatase Complex Involved in Cisplatin Sensitivity. Mol Cell Proteomics 2005; 4(11): 1725-40.
[23]   Yamamoto K, Yamazaki A, Takeuchi M, Tanaka A. A versatile method of identifying specific binding proteins on affinity resins. Anal Biochem 2006; 352(1): 15-23.
[24]   Sakamoto S, Kabe Y, Hatakeyama M, Yamaguchi Y, Handa H. Development and application of high-performance affinity beads: toward chemical biology and drug discovery. Chem Rec N Y N 2009; 9(1): 66-85. http://onlinelibrary.wiley.com/doi/10.1002/tcr.20170/full
[25]   Sakamoto S, Hatakeyama M, Ito T, Handa H. Tools and methodologies capable of isolating and identifying a target molecule for a bioactive compound. Bioorg Med Chem 2012; 20(6): 1990-2001.
[26]   Zhang S, Gerhard GS. Heme mediates cytotoxicity from artemisinin and serves as a general anti-proliferation target. PloS One 2009; 4(10): e7472.
[27]   Ong S-E, Schenone M, Margolin AA, et al. Identifying the proteins to which small-molecule probes and drugs bind in cells. Proc Natl Acad Sci USA 2009; 106(12): 4617-22.
[28]   Ong S-E, Li X, Schenone M, Schreiber SL, Carr SA. Identifying cellular targets of small-molecule probes and drugs with biochemical enrichment and SILAC. Methods Mol Biol 2012; 803: 129-40.
[29]   Cohen P. Protein kinases- the major drug targets of the twenty-first century? Nat Rev Drug Discov 2002; 1(4): 309-15.
[30]   Zhang J, Yang PL, Gray NS. Targeting cancer with small molecule kinase inhibitors. Nat Rev Cancer 2009; 9(1): 28-39.
[31]   Bantscheff M, Eberhard D, Abraham Y, et al. Quantitative chemical proteomics reveals mechanisms of action of clinical ABL kinase inhibitors. Nat Biotechnol 2007; 25(9): 1035-44.
[32]   Kruse U, Pallasch CP, Bantscheff M, et al. Chemoproteomics-based kinome profiling and target deconvolution of clinical multikinase inhibitors in primary chronic lymphocytic leukemia cells. Leukemia 2011; 25(1): 89-100.
[33]   Colzani M, Noberini R, Romanenghi M, et al. Quantitative chemical proteomics identifies novel targets of the anti-cancer multikinase inhibitor E-3810. Mol Cell Proteomics 2014; 13(6): 1495-509.
[34]   Wissing J, Jänsch L, Nimtz M, et al. Proteomics analysis of protein kinases by target class-selective prefractionation and tandem mass spectrometry. Mol Cell Proteomics 2007; 6(3): 537-47.
[35]   Thulasiraman V, Lin S, Gheorghiu L, et al. Reduction of the concentration difference of proteins in biological liquids using a library of combinatorial ligands. Electrophoresis 2005; 26(18): 3561-71.
[36]   Righetti PG, Boschetti E. The ProteoMiner and the FortyNiners: searching for gold nuggets in the proteomic arena. Mass Spectrom Rev 2008; 27(6): 596-608.
[37]   Boschetti E, Righetti PG. The ProteoMiner in the proteomic arena: a non-depleting tool for discovering low-abundance species. J Proteomics 2008; 71(3): 255-64.
[38]   Yadav AK, Bhardwaj G, Basak T, et al. A Systematic Analysis of Eluted Fraction of Plasma Post Immunoaffinity Depletion: Implications in Biomarker Discovery. PLoS ONE 2011; 6(9): e24442. http://www.plosone.org/article/info%3Adoi%2F10.1371%2Fjournal.pone.0024442
[39]   Chromy BA, Gonzales AD, Perkins J, et al. Proteomic analysis of human serum by two-dimensional differential gel electrophoresis

after depletion of high-abundant proteins. J Proteome Res 2004; 3(6): 1120-7.

[40]   Björhall K, Miliotis T, Davidsson P. Comparison of different depletion strategies for improved resolution in proteomic analysis of human serum samples. Proteomics 2005; 5(1): 307-17.

[41]   Fonslow BR, Stein BD, Webb KJ, *et al.* Digestion and depletion of abundant proteins improves proteomic coverage. Nat Methods 2013; 10(1): 54-6.

[42]   Stayton PS, Freitag S, Klumb LA, *et al.* Streptavidin-biotin binding energetics. Biomol Eng 1999; 16(1-4): 39-44.

[43]   Hong SH, Kim J, Kim J-M, *et al.* Apoptosis induction of 2'-hydroxycinnamaldehyde as a proteasome inhibitor is associated with ER stress and mitochondrial perturbation in cancer cells. Biochem Pharmacol 2007; 74(4): 557-65.

[44]   Wang G, Shang L, Burgett AWG, Harran PG, Wang X. Diazonamide toxins reveal an unexpected function for ornithine delta-amino transferase in mitotic cell division. Proc Natl Acad Sci USA 2007; 104(7): 2068-73.

[45]   O'Carra P, Barry S, Corcoran E. Affinity chromatographic differentiation of lactate dehydrogenase isoenzymes on the basis of differential abortive complex formation. FEBS Lett 1974; 43(2): 163-8.

[46]   Terry MP. Affinity Chromatography. Encyclopedia of Chromatography, 3rd ed. CRC Press; 2009.

[47]   Fee CJ, Van Alstine JM. Purification of pegylated proteins. Methods Biochem Anal 2011; 54: 339-62.

[48]   Fasoli E, Reyes YR, Guzman OM, *et al.* Para-aminobenzamidine linked regenerated cellulose membranes for plasminogen activator purification: effect of spacer arm length and ligand density. J Chromatogr B Analyt Technol Biomed Life Sci 2013; 930: 13-21.

[49]   Zamolo L, Salvalaglio M, Cavallotti C, *et al.* Experimental and theoretical investigation of effect of spacer arm and support matrix of synthetic affinity chromatographic materials for the purification of monoclonal antibodies. J Phys Chem B 2010; 114(29): 9367-80.

[50]   Busini V, Moiani D, Moscatelli D, Zamolo L, Cavallotti C. Investigation of the influence of spacer arm on the structural evolution of affinity ligands supported on agarose. J Phys Chem B 2006; 110(46): 23564-77.

[51]   Martín del Valle EM, Galán MA. Effect of the Spacer Arm in Affinity Chromatography: Determination of Adsorption Characteristics and Flow Rate Effect. Ind Eng Chem Res 2002; 41(9): 2296-304.

[52]   Diamandis EP, Christopoulos TK. The biotin-(strept)avidin system: principles and applications in biotechnology. Clin Chem 1991; 37(5): 625-36.

[53]   Gilmore BF, Quinn DJ, Duff T, Cathcart GR, Scott CJ, Walker B. Expedited solid-phase synthesis of fluorescently labeled and biotinylated aminoalkane diphenyl phosphonate affinity probes for chymotrypsin- and elastase-like serine proteases. Bioconjug Chem 2009; 20(11): 2098-105.

[54]   Dang THT, de la Riva L, Fagan RP, *et al.* Chemical probes of surface layer biogenesis in Clostridium difficile. ACS Chem Biol 2010; 5(3): 279-85.

[55]   Berry AFH, Heal WP, Tarafder AK, *et al.* Rapid multilabel detection of geranylgeranylated proteins by using bioorthogonal ligation chemistry. Chembiochem Eur J Chem Biol 2010; 11(6): 771-3.

[56]   Baumeister B, Beythien J, Ryf J, Schneeberger P, White PD. Evaluation of Biotin-OSu and Biotin-ONp in the Solid Phase Biotinylation of Peptides. Int J Pept Res Ther 2005; 11(2): 139-41. http://link.springer.com/article/10.1007%2Fs10989-004-4707-2

[57]   Cankarova N, Funk P, Hlavac J, Soural M. Novel preloaded resins for solid-phase biotinylation of carboxylic acids. Tetrahedron Lett 2011; 52(44): 5782-8.

[58]   Hopp TP, Prickett KS, Price VL, *et al.* A Short Polypeptide Marker Sequence Useful for Recombinant Protein Identification and Purification. Nat Biotechnol 1988; 6(10): 1204-10.

[59]   Saxena C, Zhen E, Higgs RE, Hale JE. An immuno-chemo-proteomics method for drug target deconvolution. J Proteome Res 2008; 7(8): 3490-7.

[60]   Schmidt PM, Sparrow LG, Attwood RM, Xiao X, Adams TE, McKimm-Breschkin JL. Taking down the FLAG! How Insect Cell Expression Challenges an Established Tag-System. PLoS ONE 2012; 7(6): e37779.

[61]   Waugh DS. Making the most of affinity tags. Trends Biotechnol 2005; 23(6): 316-20.

[62]   Chaga G, Hopp J, Nelson P. Immobilized metal ion affinity chromatography on Co2+-carboxymethylaspartate-agarose Superflow, as demonstrated by one-step purification of lactate dehydrogenase from chicken breast muscle. Biotechnol Appl Biochem 1999; 29 (1): 19-24.

[63]   Chaga G, Bochkariov DE, Jokhadze GG, Hopp J, Nelson P. Natural poly-histidine affinity tag for purification of recombinant proteins on cobalt(II)-carboxymethylaspartate crosslinked agarose. J Chromatogr A 1999; 864(2): 247-56.

[64]   Hochuli E, Döbeli H, Schacher A. New metal chelate adsorbent selective for proteins and peptides containing neighbouring histidine residues. J Chromatogr 1987; 411: 177-84.

[65]   Smith DB, Johnson KS. Single-step purification of polypeptides expressed in Escherichia coli as fusions with glutathione S-transferase. Gene 1988; 67(1): 31-40.

[66]   Schmidt TG, Skerra A. The Strep-tag system for one-step purification and high-affinity detection or capturing of proteins. Nat Protoc 2007; 2(6): 1528-35.

[67]   Kolb HC, Finn MG, Sharpless KB. Click Chemistry: Diverse Chemical Function from a Few Good Reactions. Angew Chem Int Ed Engl 2001; 40(11): 2004-21.

[68]   Rolf Huisgen. In: Centenary lecture, Proceedings of the Chemical Society. October 1961: 357-96.

[69]   Yang P-Y, Liu K, Ngai MH, Lear MJ, Wenk MR, Yao SQ. Activity-Based Proteome Profiling of Potential Cellular Targets of Orlistat − An FDA-Approved Drug with Anti-Tumor Activities. J Am Chem Soc 2010; 132(2): 656-66.

[70]   Lomenick B, Olsen RW, Huang J. Identification of direct protein targets of small molecules. ACS Chem Biol 2011; 6(1): 34-46.

[71]   West GM, Tucker CL, Xu T, *et al.* Quantitative proteomics approach for identifying protein-drug interactions in complex mixtures using protein stability measurements. Proc Natl Acad Sci USA 2010; 107(20): 9078-82.

[72]   Liu P-F, Kihara D, Park C. Energetics-based discovery of protein-ligand interactions on a proteomic scale. J Mol Biol 2011; 408(1): 147-62.

[73]   Chang Y, Schlebach JP, VerHeul RA, Park C. Simplified proteomics approach to discover protein-ligand interactions. Protein Sci 2012; 21(9): 1280-7.

[74]   Lomenick B, Jung G, Wohlschlegel JA, Huang J. Target identification using drug affinity responsive target stability (DARTS). Curr Protoc Chem Biol 2011; 3(4): 163-80.

[75]   Derry MM, Somasagara RR, Raina K, *et al.* Target identification of grape seed extract in colorectal cancer using drug affinity responsive target stability (DARTS) technique: role of endoplasmic reticulum stress response proteins. Curr Cancer Drug Targets 2014; 14(4): 323-36.

[76]   Hajduch M, Skalnikova H, Halada P, *et al.* Cyclin-Dependent Kinase Inhibitors and Cancer: Usefulness of Proteomic Approaches in Assessment of the Molecular Mechanisms and Efficacy of Novel Therapeutics. In: Clinical Proteomics: From diagnosis to therapy; Eyk JEV, Dunn MJ, eds. Wiley-VCH Verlag GmbH & Co. KGaA; 2007; p. 177-202.

[77]   Huang W, He T, Chai C, *et al.* Triptolide inhibits the proliferation of prostate cancer cells and down-regulates SUMO-specific protease 1 expression. PloS One 2012; 7(5): e37693. http://www.plosone.org/article/info%3Adoi%2F10.1371%2Fjournal.pone.0037693

[78]   Driscoll JJ, Dechowdhury R. Therapeutically targeting the SUMOylation, Ubiquitination and Proteasome pathways as a novel anticancer strategy. Target Oncol 2010; 5(4): 281-9.

[79]   Fila J, Honys D. Enrichment techniques employed in phosphoproteomics. Amino Acids 2012; 43(3): 1025-47.

[80]   Andersson L, Porath J. Isolation of phosphoproteins by immobilized metal (Fe3+) affinity chromatography. Anal Biochem 1986; 154(1): 250-4.

[81]   Matsuda H, Nakamura H, Nakajima T. New ceramic titania: selective adsorbent for organic phosphates. Anal Sci 1990; 6(6): 911-2.

[82]   Pan X, Whitten DA, Wu M, Chan C, Wilkerson CG, Pestka JJ. Global protein phosphorylation dynamics during deoxynivalenol-induced ribotoxic stress response in the macrophage. Toxicol Appl Pharmacol 2013; 268(2): 201-11.

[83]   Alpert AJ. Hydrophilic-interaction chromatography for the separation of peptides, nucleic acids and other polar compounds. J Chromatogr 1990; 499: 177-96.

[84]   Mezhoud K, Baucher al., Château-Joubert S, *et al.* Proteomic and phosphoproteomic analysis of cellular responses in medaka fish (Oryzias latipes) following oral gavage with microcystin-LR. Toxicon Off J Int Soc Toxinology 2008; 51(8): 1431-9.

[85]   Shen Z, Thomas JJ, Averbuj C, *et al.* Porous silicon as a versatile platform for laser desorption/ionization mass spectrometry. Anal Chem 2001; 73(3): 612-9.

[86]   Silva JC, Gorenstein MV, Li G-Z, Vissers JPC, Geromanos SJ. Absolute quantification of proteins by LCMSE: a virtue of parallel MS acquisition. Mol Cell Proteomics 2006; 5(1): 144-56.

[87]   Bhatia VN, Perlman DH, Costello CE, McComb ME. Software tool for researching annotations of proteins: open-source protein annotation software with data visualization. Anal Chem 2009; 81(23): 9819-23.

[88]   Zou H, Zhang Q, Guo Z, Guo B, Zhang Q, Chen X. A Mass Spectrometry Based Direct-Binding Assay for Screening Binding Partners of Proteins. Angew Chem 2002; 114(4): 668-70.

[89]   Sumranjit J, Chung SJ. Recent advances in target characterization and identification by photoaffinity probes. Mol Basel Switz 2013; 18(9): 10425-51.

[90]   Vodovozova EL. Photoaffinity labeling and its application in structural biology. Biochem Mosc 2007; 72(1): 1-20.

[91]   Hatanaka Y, Sadakane Y. Photoaffinity labeling in drug discovery and developments: chemical gateway for entering proteomic frontier. Curr Top Med Chem 2002; 2(3): 271-88.

[92]   Griffith EC, Su Z, Turk BE, *et al.* Methionine aminopeptidase (type 2) is the common target for angiogenesis inhibitors AGM-1470 and ovalicin. Chem Biol 1997; 4(6): 461-71.

[93]   Dormán G, Prestwich GD. Using photolabile ligands in drug discovery and development. Trends Biotechnol 2000; 18(2): 64-77.

[94]   Han S-Y, Hwan Kim S. Introduction to chemical proteomics for drug discovery and development. Arch Pharm (Weinheim) 2007; 340(4): 169-77.

[95]   Kotake Y, Sagane K, Owa T, *et al.* Splicing factor SF3b as a target of the antitumor natural product pladienolide. Nat Chem Biol 2007; 3(9): 570-5.

[96]   Lamos SM, Krusemark CJ, McGee CJ, Scalf M, Smith LM, Belshaw PJ. Mixed isotope photoaffinity reagents for identification of small-molecule targets by mass spectrometry. Angew Chem Int Ed Engl 2006; 45(26): 4329-33.

[97]   Cravatt BF, Wright AT, Kozarich JW. Activity-based protein profiling: from enzyme chemistry to proteomic chemistry. Annu Rev Biochem 2008; 77: 383-414.

[98]   Verhelst SHL, Bogyo M. Chemical proteomics applied to target identification and drug discovery. BioTechniques 2005; 38(2): 175-7.

[99]   Sadaghiani AM, Verhelst SH, Bogyo M. Tagging and detection strategies for activity-based proteomics. Curr Opin Chem Biol 2007; 11(1): 20-8.

[100]   Adam GC, Burbaum J, Kozarich JW, Patricelli MP, Cravatt BF. Mapping enzyme active sites in complex proteomes. J Am Chem Soc 2004; 126(5): 1363-8.

[101]   Brooks H, Lebleu B, Vivès E. Tat peptide-mediated cellular delivery: back to basics. Adv Drug Deliv Rev 2005; 57(4): 559-77.

[102]   Ruben S, Perkins A, Purcell R, *et al.* Structural and functional characterization of human immunodeficiency virus tat protein. J Virol 1989; 63(1): 1-8.

[103]   Fawell S, Seery J, Daikh Y, *et al.* Tat-mediated delivery of heterologous proteins into cells. Proc Natl Acad Sci USA 1994; 91(2): 664-8.

[104]   Saxena C, Bonacci TM, Huss KL, Bloem LJ, Higgs RE, Hale JE. Capture of drug targets from live cells using a multipurpose immuno-chemo-proteomics tool. J Proteome Res 2009; 8(8): 3951-7.

[105]   Foerg C, Merkle HP. On the biomedical promise of cell penetrating peptides: limits versus prospects. J Pharm Sci 2008; 97(1): 144-62. http://onlinelibrary.wiley.com/doi/10.1002/jps.21117/full

[106]   Zheng C-F, Simcox T, Xu L, Vaillancourt P. A new expression vector for high level protein production, one step purification and direct isotopic labeling of calmodulin-binding peptide fusion proteins. Gene 1997; 186(1): 55-60.

[107]   Luque-Garcia JL, Zhou G, Spellman DS, Sun T-T, Neubert TA. Analysis of electroblotted proteins by mass spectrometry: protein identification after Western blotting. Mol Cell Proteomics 2008; 7(2): 308-14.

[108]   Ong S-E, Mann M. A practical recipe for stable isotope labeling by amino acids in cell culture (SILAC). Nat Protoc 2006; 1(6): 2650-60.

[109]   Gygi SP, Rist B, Gerber SA, Turecek F, Gelb MH, Aebersold R. Quantitative analysis of complex protein mixtures using isotope-coded affinity tags. Nat Biotechnol 1999; 17(10): 994-9.

[110]   Thompson A, Schäfer J, Kuhn K, *et al.* Tandem mass tags: a novel quantification strategy for comparative analysis of complex protein mixtures by MS/MS. Anal Chem 2003; 75(8): 1895-904.

[111]   Ross PL, Huang YN, Marchese JN, *et al.* Multiplexed protein quantitation in Saccharomyces cerevisiae using amine-reactive isobaric tagging reagents. Mol Cell Proteomics 2004; 3(12): 1154-69.

[112]   Vaughn CP, Crockett DK, Lim MS, Elenitoba-Johnson KSJ. Analytical Characteristics of Cleavable Isotope-Coded Affinity Tag-LC-Tandem Mass Spectrometry for Quantitative Proteomic Studies. J Mol Diagn 2006; 8(4): 513-20.

[113]   Shiio Y, Aebersold R. Quantitative proteome analysis using isotope-coded affinity tags and mass spectrometry. Nat Protoc 2006; 1(1): 139-45.

[114]   Werner T, Sweetman G, Savitski MF, Mathieson T, Bantscheff M, Savitski MM. Ion coalescence of neutron encoded TMT 10-plex reporter ions. Anal Chem 2014; 86(7): 3594-601.

[115]   Malki K, Campbell J, Davies M, *et al.* Pharmacoproteomic investigation into antidepressant response in two mouse inbred strains. Proteomics 2012; 12(14): 2355-65.

[116]   Sandberg A, Branca RMM, Lehtiö J, Forshed J. Quantitative accuracy in mass spectrometry based proteomics of complex samples: The impact of labeling and precursor interference. J Proteomics 2014; 96: 133-44.

[117]   Quaglia M, Pritchard C, Hall Z, O'Connor G. Amine-reactive isobaric tagging reagents: Requirements for absolute quantification of proteins and peptides. Anal Biochem 2008; 379(2): 164-9.

[118]   Christoforou A, Arias AM, Lilley KS. Determining protein subcellular localization in mammalian cell culture with biochemical fractionation and iTRAQ 8-plex quantification. Methods Mol Biol Clifton NJ 2014; 1156: 157-74.

[119]   Gerber SA, Rush J, Stemman O, Kirschner MW, Gygi SP. Absolute quantification of proteins and phosphoproteins from cell lysates by tandem MS. Proc Natl Acad Sci 2003; 100(12): 6940-5.

[120]   Kuhn E, Wu J, Karl J, Liao H, Zolg W, Guild B. Quantification of C-reactive protein in the serum of patients with rheumatoid arthritis using multiple reaction monitoring mass spectrometry and 13C-labeled peptide standards. Proteomics 2004; 4(4): 1175-86.

[121]   Picotti P, Bodenmiller B, Mueller LN, Domon B, Aebersold R. Full dynamic range proteome analysis of S. cerevisiae by targeted proteomics. Cell 2009; 138(4): 795-806.

[122]   Andersen JS, Wilkinson CJ, Mayor T, Mortensen P, Nigg EA, Mann M. Proteomic characterization of the human centrosome by protein correlation profiling. Nature 2003; 426(6966): 570-4.

[123]   Hüttenhain R, Surinova S, Ossola R, *et al.* N-glycoprotein SRMAtlas: a resource of mass spectrometric assays for N-glycosites enabling consistent and multiplexed protein quantification for clinical applications. Mol Cell Proteomics 2013; 12(4): 1005-16.

[124]   Picotti P, Aebersold R. Selected reaction monitoring-based proteomics: workflows, potential, pitfalls and future directions. Nat Methods 2012; 9(6): 555-66.

[125]   Chang C-Y, Picotti P, Hüttenhain R, *et al.* Protein significance analysis in selected reaction monitoring (SRM) measurements. Mol Cell Proteomics 2012; 11(4): M111.014662.

[126]   Gilchrist A, Au CE, Hiding J, *et al.* Quantitative proteomics analysis of the secretory pathway. Cell 2006; 127(6): 1265-81.

[127]   Takamori S, Holt M, Stenius K, *et al.* Molecular anatomy of a trafficking organelle. Cell 2006; 127(4): 831-46.

[128]   Kislinger T, Cox B, Kannan A, *et al.* Global survey of organ and organelle protein expression in mouse: combined proteomic and transcriptomic profiling. Cell 2006; 125(1): 173-86.

[129]   Old WM, Meyer-Arendt K, Aveline-Wolf L, *et al.* Comparison of label-free methods for quantifying human proteins by shotgun proteomics. Mol Cell Proteomics 2005; 4(10): 1487-502.

[130]   Zhang B, VerBerkmoes NC, Langston MA, Uberbacher E, Hettich RL, Samatova NF. Detecting differential and correlated protein expression in label-free shotgun proteomics. J Proteome Res 2006; 5(11): 2909-18.

[131]   Andersen JS, Wilkinson CJ, Mayor T, Mortensen P, Nigg EA, Mann M. Proteomic characterization of the human centrosome by protein correlation profiling. Nature 2003; 426(6966): 570-4.

[132] Cutillas PR, Vanhaesebroeck B. Quantitative profile of five murine core proteomes using label-free functional proteomics. Mol Cell Proteomics 2007; 6(9): 1560-73.

[133] Bondarenko PV, Chelius D, Shaler TA. Identification and relative quantitation of protein mixtures by enzymatic digestion followed by capillary reversed-phase liquid chromatography-tandem mass spectrometry. Anal Chem 2002; 74(18): 4741-9.

[134] Chelius D, Bondarenko PV. Quantitative profiling of proteins in complex mixtures using liquid chromatography and mass spectrometry. J Proteome Res 2002; 1(4): 317-23.

[135] Silva JC, Denny R, Dorschel C, et al. Simultaneous qualitative and quantitative analysis of the Escherichia coli proteome: a sweet tale. Mol Cell Proteomics 2006; 5(4): 589-607.

[136] Gao B-B, Stuart L, Feener EP. Label-free quantitative analysis of one-dimensional PAGE LC/MS/MS proteome: application on angiotensin II-stimulated smooth muscle cells secretome. Mol Cell Proteomics 2008; 7(12): 2399-409.

[137] Liu H, Sadygov RG, Yates JR. A model for random sampling and estimation of relative protein abundance in shotgun proteomics. Anal Chem 2004; 76(14): 4193-201.

[138] Ishihama Y, Oda Y, Tabata T, et al. Exponentially modified protein abundance index (emPAI) for estimation of absolute protein amount in proteomics by the number of sequenced peptides per protein. Mol Cell Proteomics 2005; 4(9): 1265-72.

[139] Griffin NM, Yu J, Long F, et al. Label-free, normalized quantification of complex mass spectrometry data for proteomic analysis. Nat Biotechnol 2010; 28(1): 83-9.

[140] Trudgian DC, Ridlova G, Fischer R, et al. Comparative evaluation of label-free SINQ normalized spectral index quantitation in the central proteomics facilities pipeline. Proteomics 2011; 11(14): 2790-7.

[141] Liu H, Sadygov RG, Yates JR. A Model for Random Sampling and Estimation of Relative Protein Abundance in Shotgun Proteomics. Anal Chem 2004; 76(14): 4193-201.

[142] Wells JA, McClendon CL. Reaching for high-hanging fruit in drug discovery at protein-protein interfaces. Nature 2007; 450(7172): 1001-9.

[143] Zorn JA, Wells JA. Turning enzymes ON with small molecules. Nat Chem Biol 2010; 6(3): 179-88.

[144] Mullard A. Protein-protein interaction inhibitors get into the groove. Nat Rev Drug Discov 2012; 11(3): 173-5.

[145] Wang Z, Gerstein M, Snyder M. RNA-Seq: a revolutionary tool for transcriptomics. Nat Rev Genet 2009; 10(1): 57-63.

[146] Maier T, Güell M, Serrano L. Correlation of mRNA and protein in complex biological samples. FEBS Lett 2009; 583(24): 3966-73.

[147] Ellis MJ, Gillette M, Carr SA, et al. Connecting Genomic Alterations to Cancer Biology with Proteomics: The NCI Clinical Proteomic Tumor Analysis Consortium. Cancer Discov 2013; 3(10): 1108-12.

[148] Benson JD, Chen Y-NP, Cornell-Kennon SA, et al. Validating cancer drug targets. Nature 2006; 441(7092): 451-6.

[149] Thiel P, Kaiser M, Ottmann C. Small-molecule stabilization of protein-protein interactions: an underestimated concept in drug discovery? Angew Chem Int Ed Engl 2012; 51(9): 2012-8.

[150] Petsko GA. For medicinal purposes. Nature 1996; 384(6604 Suppl):7-9.

[151] Cockett M, Dracopoli N, Sigal E. Applied genomics: integration of the technology within pharmaceutical research and development. Curr Opin Biotechnol 2000; 11(6): 602-9.

[152] Munos B. Lessons from 60 years of pharmaceutical innovation. Nat Rev Drug Discov 2009; 8(12): 959-68.

[153] Paul SM, Mytelka DS, Dunwiddie CT, et al. How to improve R&D productivity: the pharmaceutical industry's grand challenge. Nat Rev Drug Discov 2010; 9(3): 203-14.

[154] Groner B, Weber A, Mack L. Increasing the range of drug targets: Interacting peptides provide leads for the development of oncoprotein inhibitors. Bioengineered 2012; 3(6): 320-5.

[155] Mullard A. Reliability of "new drug target" claims called into question. Nat Rev Drug Discov 2011; 10(9): 643-4.

[156] Prinz F, Schlange T, Asadullah K. Believe it or not: how much can we rely on published data on potential drug targets? Nat Rev Drug Discov 2011; 10(9): 712-712.

[157] Ioannidis JPA. Why Most Published Research Findings Are False. PLoS Med 2005; 2(8): e124. http://www.plosmedicine.org/article/info%3Adoi%2F10.1371%2Fjournal.pmed.0020124

[158] Announcement: Reducing our irreproducibility. Nature 2013; 496(7446): 398-398.

[159] Legrain P, Aebersold R, Archakov A, et al. The human proteome project: current state and future direction. Mol Cell Proteomics 2011; 10(7): M111.009993.

[160] McShane LM, Altman DG, Sauerbrei W, Taube SE, Gion M, Clark GM. REporting recommendations for tumor MARKer prognostic studies (REMARK). Nat Clin Pract Oncol 2005; 2(8): 416-22.

[161] Altman DG, McShane LM, Sauerbrei W, Taube SE. Reporting Recommendations for Tumor Marker Prognostic Studies (REMARK): explanation and elaboration. PLoS Med 2012; 9(5): e1001216. http://www.plosmedicine.org/article/info%3Adoi%2F10.1371%2Fjournal.pmed.1001216

[162] Hood L, Perlmutter RM. The impact of systems approaches on biological problems in drug discovery. Nat Biotechnol 2004; 22(10): 1215-7.

[163] Hood L, Friend SH. Predictive, personalized, preventive, participatory (P4) cancer medicine. Nat Rev Clin Oncol 2011; 8(3): 184-7.

[164] Dahlquist KD, Salomonis N, Vranizan K, Lawlor SC, Conklin BR. GenMAPP, a new tool for viewing and analyzing microarray data on biological pathways. Nat Genet 2002; 31(1): 19-20.

[165] Karp PD, Paley S, Romero P. The Pathway Tools software. Bioinforma Oxf Engl 2002; 18 Suppl 1: S225-32.

[166] Huang DW, Sherman BT, Lempicki RA. Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources. Nat Protoc 2009; 4(1): 44-57.

[167] Mencher SK, Wang LG. Promiscuous drugs compared to selective drugs (promiscuity can be a virtue). BMC Pharmacol Toxicol 2005; 5(1): 3. http://www.biomedcentral.com/1472-6904/5/3

[168] Ashburn TT, Thor KB. Drug repositioning: identifying and developing new uses for existing drugs. Nat Rev Drug Discov 2004; 3(8): 673-83.

[169] Rebuilding Big Pharma's Business Model. In vivo. Pharma & Medtech Business Intelligence 2003. https://www.pharmamedtechbi.com/publications/in-vivo/21/10/rebuilding-big-pharmas-business-model

[170] Hölzel M, Bovier A, Tüting T. Plasticity of tumour and immune cells: a source of heterogeneity and a cause for therapy resistance? Nat Rev Cancer 2013; 13(5): 365-76.

[171] Gujral TS, Peshkin L, Kirschner MW. Exploiting polypharmacology for drug target deconvolution. Proc Natl Acad Sci 2014; 111(13): 5048-53.

[172] Hood LE, Omenn GS, Moritz RL, et al. New and improved proteomics technologies for understanding complex biological systems: Addressing a grand challenge in the life sciences. Proteomics 2012; 12(18): 2773-83.

[173] Vidal M, Chan DW, Gerstein M, et al. The human proteome - a scientific opportunity for transforming diagnostics, therapeutics, and healthcare. Clin Proteomics 2012; 9(1): 6.

[174] Hood L, Flores M. A personal view on systems medicine and the emergence of proactive P4 medicine: predictive, preventive, personalized and participatory. New Biotechnol 2012; 29(6): 613-24.

[175] Swinney DC. Biochemical mechanisms of drug action: what does it take for success? Nat Rev Drug Discov 2004; 3(9): 801-8.

8.2 Appendix B

OŽDIAN, T.*, D. HOLUB, G. RYLOVÁ, J. VÁCLAVKOVÁ, M. HAJDÚCH a P. DŽUBÁK. Porovnání hmotnostně spektrometrických přístupů v proteomickém profilování léčiv. Chemagazín. 2016, XXVI(5),8-11. ISSN 1210-7409.

POROVNÁNÍ HMOTNOSTNĚ SPEKTROMETRICKÝCH PŘÍSTUPŮ V PROTEOMICKÉM PROFILOVÁNÍ LÉČIV

OŽDIAN T., HOLUB D., RYLOVÁ G., VÁCLAVKOVÁ J., HAJDÚCH M., DŽUBÁK P.

Ústav molekulární a translační medicíny, Lékařská Fakulta, Univerzita Palackého v Olomouci, petr.dzubak@upol.cz

Při analýze proteomu se využívají dva hlavní iontové zdroje, elektrosprej (ESI) a matricí asistovaná laserová desorpce/ionizace (MALDI). Pro účely jejich porovnání byla SILAC značená linie CCRF-CEM ošetřena třemi platinovými léčivy a následně byl po frakcionaci a digesci analyzován celobuněčný proteinový lyzát. Vzorky byly paralelně měřeny na třech hmotnostních spektrometrech s různými analyzátory – iontovou pastí (ESI-IT), analyzátorem doby letu (MALDI-TOF) a orbitální iontovou pastí (nESI-Orbitrap). Každý spektrometr byl spojen s vysokoúčinnou kapalinovou chromatografií a byl nezávisle optimalizován pro nejvyšší výkon. Data byla nezávisle analyzována, přičemž průměrný počet proteinů identifikovaných ESI-IT byl 660 ± 124 s 66% překryvem ve všech třech replikátech, u MALDI-TOF 355 ± 68 proteinů s 41% překryvem a u nESI-Orbitrap 3 430 ± 306 proteinů s 76% překryvem. Kvantifikační přesnost vyjádřená jako R^2 byla u ESI-IT 0,454 ± 0,047, u MALDI-TOF 0,524 ± 0,134 a u nESI-Orbitrap 0,692 ± 0,063. Při vzájemném porovnání se proteiny identifikované ESI-IT a MALDI-TOF lišily minimálně a byly beze zbytku identifikovány také pomocí nESI-Orbitrap.

Úvod

Hmotnostní spektrometrie (MS) je jednou ze základních metod identifikace proteinů. Dalo by se říci, že rozvoj hmotnostní spektrometrie stojí do značné míry i za rozvojem proteomiky. Ve svých počátcích byla využívána k identifikacím jednotlivých proteinů, například metodou protein mass fingerprinting, přičemž v současnosti je schopná identifikovat tisíce proteinů vedle sebe. Běžné je propojení hmotnostního spektrometru s vysokoúčinnou kapalinovou chromatografií (HPLC), jsou ale možná i jiná spojení, například s elektroforetickou separací, nebo s iontově výměnnou chromatografií. Časté je i kombinování metod například v prvním kroku separace elektroforézou, digesce a následná HPLC-MS analýza. Tím se dosáhne lepší separace jednotlivých peptidů a zvýší se tak jejich záchyt. Cílem tohoto příspěvku je porovnat tři metody hmotnostně spektrometrické analýzy používané v proteomice.

První metodou identifikace proteinů je on-line spojení HPLC s elektrosprejem a iontovou pastí (ESI-IT). Tato kombinace vyniká rychlostí měření jednotlivých spekter a možností rychlé fragmentace peptidů. Díky vysoké rychlosti měření jednotlivých spekter je možné fragmentovat větší množství peptidů vytékajících v jednom okamžiku z kolony. Nevýhodou je nižší rozlišení hmotnostních spekter.

Druhou metodou je spojení HPLC s matricí asistovanou laserovou desorpcí/ionizací s analyzátorem doby letu (MALDI-TOF). Tento způsob ionizace vyžaduje přítomnost matrice, a proto musí toto spojení probíhat off-line za pomoci spotovacího zařízení. Při spotování dochází k přimíchávání matrice do HPLC eluátu, přičemž vzniklá směs je dávkována na MALDI terčík. Hlavní výhodou HPLC–MALDI-TOF analýzy je časové oddělení separace od samotného měření. Hmotnostní spektrometr má tedy čas fragmentovat všechny píky ve spektru. Časové oddělení měření od separace je současně největší nevýhodou tohoto přístupu. Měření může trvat v závislosti na bohatosti vzorku až šestinásobek doby HPLC běhu.

Posledním přístupem v tomto porovnání je spojení nanoHPLC s hybridním hmotnostním spektrometrem s orbitální iontovou pastí a nanoelektrosprejem (nESI-Orbitrap). Výhodou tohoto spektrometru je vysoká přesnost analýzy peptidů v orbitální iontové pasti, která trvá delší dobu než měření v lineární iontové pasti. Proto se s výhodou uplatňuje uspořádání, kdy po změření peptidů v pasti orbitální je možné je dále fragmentovat v paralelní iontové pasti. Dochází tak k získání dvou druhů informací. První informací je velmi přesná hmotnost dostačující k identifikaci peptidu a druhou informací je jeho sekvence získaná z fragmentačního spektra. Jako hlavní nevýhodu lze uvést v současnosti vysokou cenu přístroje.

Toto srovnání je také srovnání dvou generací proteomických přístupů. V prvním, starším, jsme vycházeli z komplementarity ESI a MALDI, kdy využíváme skutečnosti, že některé peptidy se lépe ionizují pomocí jednoho zdroje a jiné pomocí druhého [1]. Za účelem dalšího zvýšení identifikace proteinů jsme použili vyhledávání pomocí dvou mechanismů (Mascot [2] a Phenyx [3]) a dvou databází (NCBI a Uniprot). Novější přístup spočívá v analýze nanoHPLC s dlouhým gradientem a přesnou analýzou nESI-Orbitrapem. Data z této metody byla zpracována a vyhledána v akademickém softwaru MaxQuant [4], který využívá databázi Uniprot. Z tohoto popisu je tedy jasné, že novější přístup je mnohem jednodušší a zejména ve srovnání s HPLC-MALDI-TOF také mnohem rychlejší.

Materiál a metody

Chemosenzitivní buněčnou linii CCRF-CEM jsme označili metabolickým izotopickým značením metodou SILAC [5], která je založena na dlouhodobé kultivaci buněk s „těžkým" $^{12}C^{13}C_6$ argininem a $^{13}C_6$ lyzinem. Takto značenou „těžkou" linii jsme použili ve všech případech jako kontrolu. Izotopicky neznačená, „lehká", buněčná linie byla naproti tomu ošetřena cisplatinou (koncentrace léčiva 5 x IC50 = 12,6 μM, doba inkubace 2,5 h), karboplatinou (5 x IC50 = 50,7 μM, doba inkubace 3 h) a oxaliplatinou (5 x IC50 = 29,3 μM, doba inkubace 4 h). Tyto koncentrace byly stanoveny na základě předchozích experimentů jako pětinásobek cytotoxické koncentrace, která usmrtí po 72 hodinách 50 % buněk (IC50). Časy ošetření léčivy byly stanoveny jako polovina času do aktivace kaspáz, jejichž aktivita je známkou nezvratně zahájených procesů vedoucích k buněčné smrti (apoptóze). Pro jejich stanovení jsme použili kit MagicRed monitorující aktivitu kaspáz [6]. Tímto jsme v experimentech použili ekviaktivní koncentraci a srovnatelný čas ošetření, ve kterém jsme zohlednili rychlost nástupu buněčné smrti po ošetření různými léčivy. Zároveň jsme se vyvarovali hodnocení pozdních apoptotických změn, které by v pozdějších intervalech převládly. Po inkubaci byla ošetřená buněčná linie smíchána v poměru 1:1 s neošetřenou linií. Směs byla lyzována pufrem obsahujícím 20 mM Tris-HCl, 7 M močovinu, 10 mM DTT, 1% Triton X-100, 0,5% SDS a 2U benzonázy. Lyzát byl separován pomocí preparativní SDS-PAGE elektroforézy. Elektroforetický gel byl rozřezán na 20 frakcí a proteiny v těchto frakcích redukovány Tris(2-karboxyethyl)fosfinem, alkylovány jodacetamidem a naštěpeny trypsinem. Směs peptidů byla po štěpení a před LC/MS analýzou přečištěna pomocí kolonek MiChrom Macrotrap s C18 sorbentem.

Hmotnostně spektrometrická analýza probíhala celkem na třech přístrojích. Nejprve byly vzorky analyzovány na iontové pasti Bruker HCTultra spojené s kapalinovým chromatografem Agilent 1200 a kolonou Agilent Zorbax 300SB-C18 o rozměrech 150 mm x 75 μm a velikostí částic 3,5 μm. Mobilní fáze byly voda s 0,1% kyselinou mravenčí a acetonitril s 0,1% kyselinou mravenčí. Gradient byl nastaven s nárůstem od 5 % do 30 % organické fáze po dobu 70 minut s průtokem 0,4 μl/min. Změřená spektra byla odeslána do programu

Bruker ProteinScape 2.1.0573, kde byla vyhledána dvěma paralelními přístupy. Prvním byl algoritmus Mascot s databází NCBI aktuální k 21.1.2010 a nastavením: taxonomie *Homo sapiens*, enzym trypsin, 2 povolená chybná štěpení, modifikace: karbamidomethylace cysteinu, oxidace methioninu a značení $^{13}C_6$ argininu a $^{13}C_6$ lysinu, hmotnostní tolerance peptidů ± 0,5 Da, MS/MS tolerance ± 0,5 Da, náboj peptidů 2+ a 3+. Peptid byl akceptován při minimálním Mascot skóre 12 a protein při minimálním skóre 40. Druhý algoritmus, Phenyx, měl tato nastavení: databáze UniProt (aktuální k 25.1.2010), taxonomie *Homo sapiens*, výchozí náboj iontů 2+, 3+, 4+, skórovací model HCTultra, modifikace byly totožné jako v algoritmu Mascot, enzym trypsin, hmotnostní tolerance 0,5 Da, maximálně povolená 2 chybná štěpení, mód štěpení „half cleaved", b-y ionty, sekvenční pokrytí 20 %. Výsledky z obou vyhledávacích přístupů byly spojeny a zpracovány pomocí nástroje Protein extractor a proteiny byly kvantifikovány nástrojem Bruker Warp-LC verze 1.2. Všechna data pak byla zpracována pomocí programu ProteinScape.

Paralelní analýzou k prvnímu přístupu byla analýza LC-MALDI. Separace probíhala off-line na kapalinovém chromatografu Agilent Capillary 1200 s kolonou MiChrom Magic C18AQ 0,2x150 mm a velikostí částic 5 μm spojeným se spotovacím robotem Bruker Proteineer fc. Mobilní fáze byly voda s 0,1% trifluoroctovou kyselinou a acetonitril s 0,1% trifluoroctovou kyselinou, průtok 3 μl/min, gradient byl 5 % do 5. minuty a pak se procento organické fáze lineárně zvyšovalo až na 35 % po dobu 55 minut. Poté následoval proplach 95% organickou fází po dobu 15 minut a ekvilibrace na výchozí podmínky trvala 15 minut. Frakce byly sbírány mezi 6. a 70. minutou. Eluát byl pomocí spotovacího robota rozdělen na 384 frakcí s délkou jedné frakce 8 s a ke každé frakci byl přidáván 1 μl α kyanohydroxyskořicové kyseliny (1 μg/ml) jako matrice. Samotné měření probíhalo na přístroji Bruker Autoflex III. V první fázi byla změřena MS spektra v rozsahu m/z 600–3500 v reflektronovém pozitivním módu při síle laseru 67 % a napětí na detektoru 2426 V. Na jeden spot bylo naměřeno 2000 spekter. Po proměření celého terčíku byl v programu WARP-LC vygenerován seznam píků vhodných pro fragmentační analýzu. Ta probíhala automaticky v reflektronovém pozitivním módu s izolačním oknem 10 Da a se zvýšením síly laseru o 65 % a napětí na detektoru o 100 % vůči měření prekurzorových iontů. Hmotnostní spektra byla zpracována v programu ProteinScape a vyhledána pomocí algoritmů Mascot a Phenyx podobným způsobem jako u dat z iontové pasti s rozdíly v nastavení: náboj 1+, MS tolerance 100 ppm a MS/MS tolerance 0,5 Da.

Analýza nESI-Orbitrap byla provedena na přístroji Orbitrap Elite (Thermo) spojeným s iontovým zdrojem Proxeon Easy-Spray a chromatografem Ultimate 3000 RSLCnano. Jeden mikrolitr vzorku byl nanesen na odsolovací kolonu PepMap 100 (75 μm x 2 cm, 3 μm, velikost pórů 100 Å; Thermo) zapojenou „in-line" s analytickou kolonou PepMap RSLC (75 μm x 15 cm, 3 μm, velikost pórů 100 Å; Thermo) vyhřívanou na 35 °C. Peptidy byly separovány na analytické koloně s lineárně rostoucím gradientem organické fáze z 5 % na 35 % s celkovou délkou běhu 150 minut. Vodná fáze byla voda s 0,1% kyselinou mravenčí a organická fáze acetonitril s 0,1% kyselinou mravenčí. Rozlišení orbitální pasti bylo nastaveno na 120000 a prekurzorové ionty byly skenovány v hmotnostním rozsahu 300–1950 m/z. V každém cyklu bylo vybráno dvacet nejintenzivnějších iontů pro fragmentaci kolizí indukovanou disociací a analyzováno v lineární pasti s kolizní energií 35 eV.

Výběr píků a vyhledání peptidů u nESI-Orbitrap proběhly s pomocí programu MaxQuant v. 1.3.0.5[4] používající databázi UniProt (aktuální k 4.4.2013) omezenou na lidské proteiny. Všechny frakce a replikáty jednoho léčiva byly vyhledány společně. Karbamidomethylace cysteinů byla zvolena jako fixní modifikace a jako variabilní modifikace oxidace methioninů a acetylace proteinů na N-konci. Bylo použito značení pomocí $^{13}C_6$ argininu a $^{13}C_6$ lysinu. Minimální počet peptidů a razor peptidů byl nastaven na 1 a minimální délka peptidu na 6 aminokyselin. Hmotnostní tolerance prekurzorového iontu byla nastavena na 20 ppm a tolerance fragmentů na 0,5 Da. Počet falešných identifikací (False discovery rate, FDR) byl pro proteiny a peptidy nastaven na 1 %.

Data z hmotnostních spektrometrů ESI-IT a MALDI-TOF byla exportována z programu ProteinScape ve formě excelových tabulek, katalogová čísla proteinů byla sjednocena na standard UniProt pomocí nástroje Retrieve / ID mapping (http://www.uniprot.org/uploadlists/, aktuální ke 2. 3. 2016) a převedena na textový soubor. Data z nESI-Orbitrap vyhledaná pomocí programu MaxQuant jsou v textovém formátu standardně. Vyexportovaná data byla zpracována v programu Perseus 1.4.0 s využitím funkcí Combine runs a Numerical Venn Diagram. Vennovy diagramy byly vizualizovány pomocí programu Venn diagram plotter 1.5 a Inkscape 0.48. Pro porovnávání metod byly použity peptidy a proteiny, které nebyly jen identifikovány, ale také kvantifikovány. Kvantitativní informace má při biologické interpretaci mnohem větší význam než pouhá identifikace.

Jako první porovnávanou informaci jsme zvolili počet MS/MS spekter, která byla danou metodou změřena a identifikována. Úspěšná identifikace MS/MS spekter vede k určení peptidů. Následným skládáním peptidů se identifikují příslušné proteinové rodiny a konkrétní proteiny. Peptidy a proteiny jsme porovnávali pomocí Vennových diagramů pro jednotlivé metody a replikáty, ale také pro všechny metody dohromady. V tomto zobrazení jsou započteny peptidy a proteiny přítomné alespoň v jednom replikátu na metodu.

Pro porovnání kvantifikační účinnosti byly vybrány pouze proteiny, které se pro dané léčivo a metodu vyskytovaly ve všech třech opakováních. Hodnota poměru těžkého a lehkého peptidu (H/L) byla zprůměrována a do grafu byly vyneseny vztahy jednotlivých analýz k jejich průměru. Hodnota kvantifikační přesnosti byla určena jako R^2 a je udána pro každý replikát a metodu zvlášť.

Výsledky a diskuze

Při porovnávání hmotnostních spektrometrů je nutné vzít do úvahy řadu faktorů jako je rozlišení, přesnost, citlivost a rychlost. Tyto parametry jsou důležité a přispívají ke konečnému výsledku proteomického experimentu. Jak již bylo řečeno v úvodu, v tomto příspěvku porovnáváme technologicky rozdílné HPLC-MS přístupy a spíše než abychom porovnávali samotné hmotnostní spektrometry, hodnotíme komplexní účinnost tří odlišných experimentálních uspořádání, z nichž každé má své výhody a nevýhody. Jednotlivé platformy byly totiž nezávisle optimalizovány pro co nejvyšší výkon, a tudíž nebylo možné porovnávat měření při stejných podmínkách (např. HPLC gradient, velikost kolony a další). Data pocházející z ESI-IT a MALDI-TOF byla zpracovávána a uchovávána v softwaru Bruker ProteinScape. Data z třetího přístroje, nESI-Orbitrapu, byla zpracována v programu MaxQuant vytvořeném přímo pro analýzu spekter z Orbitrapu. Ačkoliv existují univerzální formáty prezentace MS spekter [7], považovali jsme specializovaný software dodávaný výrobcem, respektive vyvinutý přímo pro konkrétní přístroj, za lepší řešení.

Prvním parametrem, který v tomto přehledu porovnáváme, je počet MS/MS spekter. Je to počet píků, které daný software vyhodnotil jako významné a zvolil pro fragmentaci. Fragmentace prekurzorových iontů má pro identifikaci peptidů veliký význam, protože v ideálním případě jsme z fragmentačního spektra schopni určit celou sekvenci peptidu. Těchto ideálních spekter je však v běžné analýze málo. Spektra, která neobsahují kompletní sekvenční informaci, jsou však přesto identifikována na základě pravděpodobnostního modelu [2]. V případě ESI-IT byl průměrný počet MS/MS spekter 70 283 ± 10 137, z toho bylo identifikováno 10 862 ± 2 052. U MALDI-TOF bylo naměřeno 73 547 ± 14 990 spekter, z toho 3 838 ± 812 bylo identifikováno. Analýza pomocí nESI-Orbitrap poskytla 596 296 ± 32 521 naměřených a 109 260 ± 22 419 identifikovaných spekter (obr. 1).

Počet identifikovaných peptidů je prvním důležitým parametrem. Stejně důležitá je ale i jeho reprodukovatelnost. Stav, kdy sice máme dostatek peptidů, ale nízkou opakovatelnost, se nazývá undersampling a je nežádoucím efektem „large scale" proteomických a metabolomických experimentů [8]. Na obrázku 2 je překrytí jednotlivých peptidů, kdy v případě ESI-IT bylo identifikováno 6 177 ± 1 218 peptidů překrývajících se ve všech třech replikátech, u MALDI-TOF 1 778 ± 92 peptidů a nESI-Orbitrap měl překryv 12 296 ± 2 028 peptidů.

Obr. 1 – Počet MS/MS spekter, které byly jednotlivými metodami změřeny a spekter, které byly identifikovány a kvantifikovány

Počet změřených spekter

Počet identifikovaných spekter

■ ESI-IT ■ MALDI-TOF ■ nESI-Orbitrap
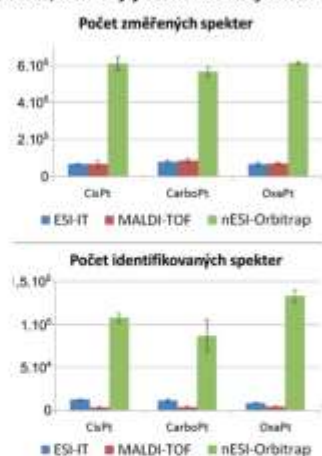
Obr. 2 – Míra překryvu identifikovaných a kvantifikovaných peptidů pro jednotlivá měření. Suma byla počítána z počtu peptidů, které se v dané metodě vyskytovaly alespoň v jednom replikátu. Grafické znázornění sumárního diagramu je následující
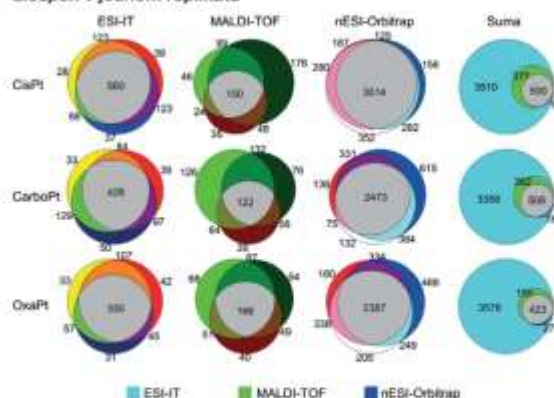

■ ESI-IT ■ MALDI-TOF ■ nESI-Orbitrap
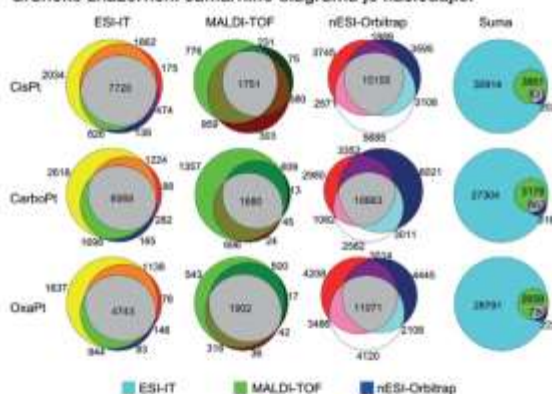
Počet identifikovaných peptidů je důležitý při hodnocení změn na úrovni proteinových rodin a jednotlivých proteinů. Čím více peptidů bylo přiřazeno konkrétnímu proteinu, tím je vyšší jeho sekvenční pokrytí a tím přesnější je kvantifikace a další údaje o studovaném proteinu. Průměrný počet peptidů na protein byl $12,0 \pm 1,6$ u ESI-IT, $7,4 \pm 1,5$ u MALDI-TOF a $5,5 \pm 0,4$ u nESI-Orbitrap. Pokrytí sekvence bylo (v procentech) $28,9 \pm 3,0$ u ESI-IT, $15,6 \pm 2,8$ u MALDI-TOF a $15,8 \pm 1,6$ u nESI-Orbitrap. Počet proteinů identifikovaných ESI-IT byl 660 ± 124 s 66% překryvem ve všech třech replikátech, u MALDI-TOF 355 ± 68 proteinů s 41% překryvem a u nESI-Orbitrap $3 430 \pm 306$ proteinů s 76% překryvem (obr. 3). Nejméně proteinů bylo tedy identifikováno metodou MALDI-TOF a zároveň byla v této metodě nejvyšší úroveň undersamplingu. Metody ESI-IT a nESI-Orbitrap se s problémem undersamplingu vyrovnaly lépe, přesto lze vidět rozdíl jednoho řádu v počtu identifikovaných proteinů.

Při porovnání ESI-IT a MALDI-TOF je nutné vzít v potaz, že z principu metody je rozlišení a přesnost MALDI-TOF oproti ESI-IT mnohem vyšší. Z toho plyne rozdíl v kvantifikační přesnosti znázorněný na obrázku 4, kdy ESI-IT mělo průměrnou přesnost R^2 $0,454 \pm 0,047$ a MALDI-TOF R^2 $0,569 \pm 0,081$. Kvantifikační přesnost nESI-Orbitrap je v rámci porovnávaných metod nejvyšší, a to $R^2 = 0,692 \pm 0,063$.

Z porovnání na obrázku 2 a 3 je zřejmé, že udávaná komplementarita MALDI a ESI přístupů se ve zkoumaném vzorku příliš neprojevila. Jejich komplementaritu jsme v naší laboratoři pozorovali až při použití MALDI-TOF spektrometru novější generace (nepublikovaná data). Význam komplementarity MALDI-TOF a ESI-IT ovšem zaniká

Obr. 3 – Míra překryvu pro proteiny identifikované a kvantifikované pomocí rozdílných hmotnostně spektrometrických přístupů. Každé léčivo bylo měřeno ve třech biologických replikátech. Suma byla počítána z počtu peptidů, které se v dané metodě vyskytovaly alespoň v jednom replikátu

■ ESI-IT ■ MALDI-TOF ■ nESI-Orbitrap

v porovnání s daty získanými z nESI-Orbitrap, kdy jsme identifikovali jak všechny proteiny z ESI-IT tak všechny proteiny z MALDI-TOF.

Větší počet identifikovaných proteinů nám umožňuje popsat více detailů v buněčné odpovědi. Naše data odráží podobný trend, jako byl popsán v článku Hammerové et al. [1] z roku 2010, kdy autoři kvantifikovali 778 proteinů pomocí komplementárních přístupů MALDI-TOF a ESI-FTICR. Signifikantně změněné proteiny zde byly pouze zmíněné, případně členěné do jednotlivých skupin. V článku o pět let novějším, od autorů Petroviče et al. [9], bylo kvantifikováno pomocí novější generace iontové pasti 2064 proteinů a již se neuvádí pouze signifikantní změny jednotlivých proteinů, ale popisují se i ucelené buněčné procesy ovlivněné doxorubicinem. Za tímto posunem stojí nejen rozvoj počtu identifikovaných proteinů, ale také rozvoj bioinformatiky umožňující zařazení proteinů do jednotlivých drah. Pomocí zlepšení v těchto oblastech jsme schopni získat ucelenější informace o buněčné odpovědi na ošetření léčivem.

Závěr

Porovnali jsme tři nezávislé proteomické přístupy, ESI-IT, MALDI-TOF a nESI-Orbitrap, z hlediska jejich účinnosti v LC-MS uspořádání. Ve většině sledovaných parametrů měl největší úspěšnost identifikace nESI-Orbitrap. Pomocí ESI-IT bylo identifikováno více proteinů než v případě MALDI TOF, MALDI-TOF naproti tomu poskytlo přesnější kvantifikaci. U přístupu ESI-IT se vyskytoval nejvyšší počet peptidů na protein a pozorovali jsme nejvyšší sekvenční pokrytí v rámci celého experimentu.

Závěrem je nutné říci, že ve srovnání použitých přístupů je Orbitrap ze všech tří přístrojů nejnovější a toto srovnání je díky tomu i odrazem rychlého vývoje přístrojových a analytických technologií v oblasti proteomiky a hmotnostní spektrometrie za několik posledních let.

Literatura

[1] E. Hammer, S. Bien, M.G. Salazar, L. Steil, C. Scharf, P. Hildebrandt, et al., Proteomic analysis of doxorubicin-induced changes in the proteome of HepG2cells combining 2-D DIGE and LC-MS/MS approaches, *Proteomics*. 10 (2010) 99–114. doi:10.1002/pmic.200800626.

[2] D.N. Perkins, D.J.C. Pappin, D.M. Creasy, J.S. Cottrell, Probability-based protein identification by searching sequence databases using mass spectrometry data, *ELECTROPHORESIS*. 20 (1999) 3551–3567. doi:10.1002/(SICI)1522-2683(19991201)20:18<3551::AID-ELPS3551>3.0.CO;2-2.

Obr. 4 – Kvantifikační přesnost. Poměry jednotlivých proteinů byly vyneseny proti průměrnému poměru pro daný protein a metodu. Pro každý replikát byla vypočítána samostatná hodnota R^2. R1, R2 a R3 je označení pro jednotlivé replikáty
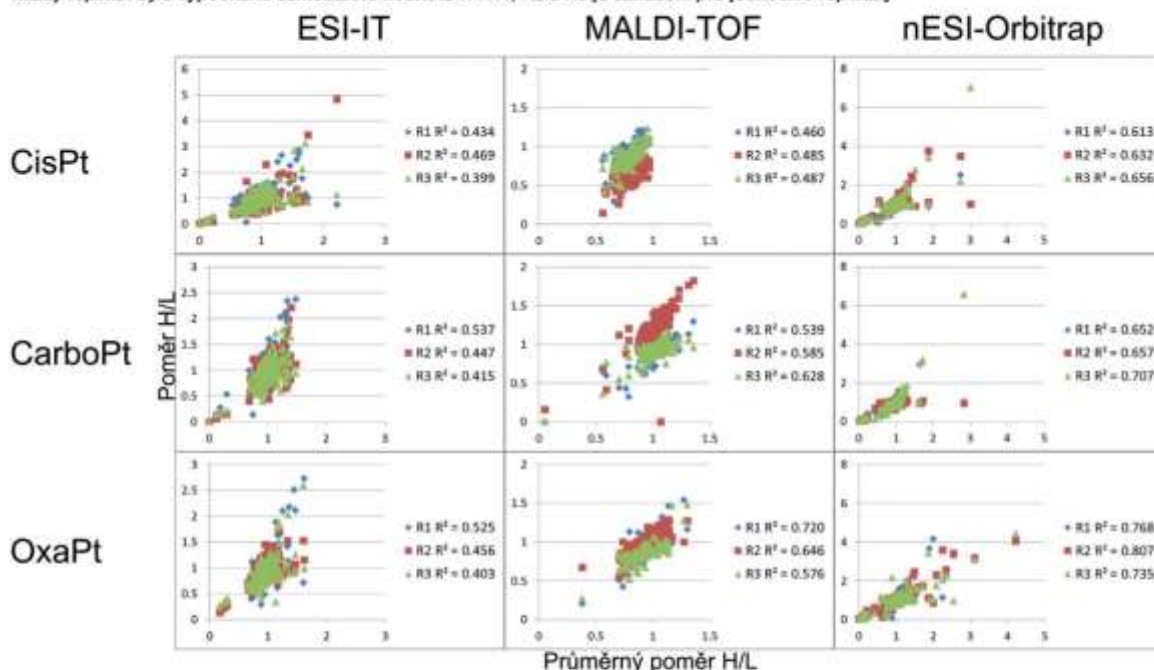
[3]J. Colinge, A. Masselot, M. Giron, T. Dessingy, J. Magnin, OLAV: towards high-throughput tandem mass spectrometry data identification, *Proteomics*. 3 (2003) 1454–1463. doi:10.1002/pmic.200300485.

[4]J. Cox, M. Mann, MaxQuant enables high peptide identification rates, individualized p.p.b.-range mass accuracies and proteome-wide protein quantification, *Nat. Biotechnol.* 26 (2008) 1367–1372. doi:10.1038/nbt.1511.

[5]S.-E. Ong, B. Blagoev, I. Kratchmarova, D.B. Kristensen, H. Steen, A. Pandey, et al., Stable isotope labeling by amino acids in cell culture, SILAC, as a simple and accurate approach to expression proteomics, *Mol. Cell. Proteomics* MCP. 1 (2002) 376–386.

[6]B.W. Lee, G.L. Johnson, S.A. Hed, Z. Darzynkiewicz, J.W. Talhouk, S. Mehrotra, DEVDase detection in intact apoptotic cells using the cell permeant fluorogenic substrate, (z-DEVD)2-cresyl violet, *BioTechniques*. 35 (2003) 1080–1085.

[7]L. Martens, M. Chambers, M. Sturm, D. Kessner, F. Levander, J. Shofstahl, et al., mzML--a community standard for mass spectrometry data, *Mol. Cell. Proteomics* MCP. 10 (2011) R110.000133. doi:10.1074/mcp.R110.000133.

[8]H.C.J. Hoefsloot, S. Smit, A.K. Smilde, A classification model for the Leiden proteomics competition, *Stat. Appl. Genet. Mol. Biol.* 7 (2008) Article8. doi:10.2202/1544-6115.1351.

[9]M. Petrovic, C. Simillion, P. Kruzliak, J. Sabo, M. Heller, Doxorubicin Affects Expression of Proteins of Neuronal Pathways in MCF-7 Breast Cancer Cells, *Cancer Genomics - Proteomics*. 12 (2015) 347–358.

Abstract

THE COMPARISON OF MASS SPECTROMETRY APPROACHES IN PROTEOMIC PROFILING OF DRUG RESPONSES.

Summary: *There are two main mass spectrometry approaches in proteome analysis: electrospray (ESI) and matrix assisted laser desorption/ionization (MALDI). For purposes of their comparison, the SILAC labeled CCRF-CEM cell line was treated by three platinum drugs and whole cell lysate was analyzed. Samples were measured in parallel by three mass spectrometers – ion trap (ESI-IT), time-of-flight (MALDI-TOF) and orbital ion trap (nESI-Orbitrap). Each spectrometer was coupled with high performance liquid chromatography and independently optimized for best performance. Data were analyzed independently and average count of identified proteins was 660 ± 124 with 66% coverage in all three replicates, MALDI-TOF with 355 ± 68 proteins and 41% coverage and nESI-Orbitrap 3430 ± 306 proteins with 76% coverage. Quantification accuracy of ESI-IT determined as R^2 was 0,454 ± 0,047, MALDI-TOF 0,524 ± 0,134 and nESI-Orbitrap 0,692 ± 0,063. Proteins identified by ESI-IT and MALDI-TOF in at least one replicate was nearly the same and all of them were identified in nESI-Orbitrap as well.*

Key words: Proteomics, mass spectrometry, LC-MS

157

8.3 Appendix C

ŠIMKOVÁ, D., G. KHARAISHVILI, G. KOŘÍNKOVÁ, **T. OŽDIAN**, T. SUCHANKOVA-KLEPLOVA, T. SOUKUP, M. KRUPKA, A. GALANDÁKOVÁ, P. DŽUBÁK, M. JANÍKOVÁ, J. NAVRATIL, Z. KAHOUNOVA, K. SOUCEK a J. BOUCHAL. CHYBI DEDI a PDF_The dual role of asporin in breast cancer progression. Oncotarget. 2016, -, -. ISSN 1949-2553. IF: 5.008. PMID: 27409832

This appendix also includes supplementary file 1 containing MS data.

The dual role of asporin in breast cancer progression

Dana Simkova[1], Gvantsa Kharaishvili[1], Gabriela Korinkova[1], Tomas Ozdian[2], Tereza Suchánková-Kleplová[3], Tomas Soukup[3], Michal Krupka[4], Adela Galandakova[5], Petr Dzubak[2], Maria Janikova[1], Jiri Navratil[6], Zuzana Kahounova[7,8], Karel Soucek[7,8,9], Jan Bouchal[1]

[1]Department of Clinical and Molecular Pathology, Institute of Molecular and Translational Medicine, Faculty of Medicine and Dentistry, Palacky University, Olomouc, Czech Republic

[2]Laboratory of Experimental Medicine, Institute of Molecular and Translational Medicine, Faculty of Medicine and Dentistry, Palacky University, Olomouc, Czech Republic

[3]Department of Histology and Embryology, Faculty of Medicine in Hradec Kralove, Charles University, Hradec Kralove, Czech Republic

[4]Department of Immunology, Faculty of Medicine and Dentistry, Palacky University, Olomouc, Czech Republic

[5]Department of Medical Chemistry and Biochemistry, Institute of Molecular and Translational Medicine, Faculty of Medicine and Dentistry, Palacky University, Olomouc, Czech Republic

[6]Department of Comprehensive Cancer Care, Masaryk Memorial Cancer Institute, Brno, Czech Republic

[7]Department of Cytokinetics, Institute of Biophysics, Academy of Sciences of the Czech Republic, v.v.i., Brno, Czech Republic

[8]Center of Biomolecular and Cellular Engineering, International Clinical Research Center, St. Anne's University Hospital Brno, Brno, Czech Republic

[9]Department of Experimental Biology, Faculty of Science, Masaryk University, Brno, Czech Republic

Correspondence to: Jan Bouchal, *email:* jan.bouchal@upol.cz

ABSTRACT

Asporin has been reported as a tumor suppressor in breast cancer, while asporin-activated invasion has been described in gastric cancer. According to our in silico search, high asporin expresion associates with significantly better relapse free survival (RFS) in patients with low-grade tumors but RFS is significantly worse in patients with grade 3 tumors. In line with other studies, we have confirmed asporin expression by RNA scope *in situ* hybridization in cancer associated fibroblasts. We have also found asporin expression in the Hs578T breast cancer cell line which we confirmed by quantitative RT-PCR and western blotting. From multiple testing, we found that asporin can be downregulated by bone morphogenetic protein 4 while upregulation may be facilited by serum-free cultivation or by three dimensional growth in stiff Alvetex scaffold. Downregulation by shRNA inhibited invasion of Hs578T as well as of CAFs and T47D cells. Invasion of asporin-negative MDA-MB-231 and BT549 breast cancer cells through collagen type I was enhanced by recombinant asporin. Besides other investigations, large scale analysis of aspartic acid repeat polymorphism will be needed for clarification of the asporin dual role in progression of breast cancer.

INTRODUCTION

Tumor progression is partly a result of evolving crosstalk between different cell types within the tumor and its surrounding supportive tissue/tumor stroma [1]. The extracellular matrix (ECM) is a dynamic structure providing constructional support, organisation and orientation into tissues [2] as well as supplying crucial elements for cell survival i.e. growth factors, inflammatory molecules and immune soluble mediators [3]. ECM consists of a plethora of proteins of varying structure and function [4]. Inter alia, the small leucine rich proteoglycan family (SLRPs) constitutes an important group of ECM proteins, widely found in most extracellular matrices [5].

Asporin was identified by three research groups in 2001 [6, 7, 8]. The aspartic acid rich N-terminal region and

159

central part of the asporin molecule bind type II collagen [9] while collagen I is bound by the central region. Asporin inhibits collagen fibrilogenesis and competes with decorin in binding the same sites. For this reason, this competition may have a role in regulating the development of ECM. The asporin N-terminal polyaspartate domain also binds calcium, and works in concert with other domains in order to initiate the mineralization of collagen [10]. Asporin plays an important role in normal development, in particular of cartilage, bone and teeth, while its genetic polymorphisms have been associated with various bone and joint diseases, including osteoarthritis, rheumatoid arthritis and lumbar disc disease [11].

We identified asporin by microdissection and expression profiling as a novel breast cancer related protein [12]. It was upregulated in invasive carcinomas, in particular lobular ones, together with other extracellular matrix proteins such as periostin or versican [13, 14]. Asporin can also be found in gene lists from other expression microarray analyses both in breast and other cancer types [15–19]. Association of asporin with different types of cancer has also been confirmed by other methods such as tag profiling and mass spectrometry [20–23]. Important findings have been recently reported by Satoyoshi et al. [24] who found strong asporin expression in cancer associated fibroblasts and its importance in coordinated invasion of gastric cancer. On the other hand, Maris et al. [25] have described the tumor suppressive potential of asporin in breast cancer. With respect to these contradictory results, the role of asporin in cancer progression and the tumor microenvironment deserves further investigation.

RESULTS

Asporin associates with better prognosis in low-grade tumors but not in high-grade breast cancer

To acquire novel clues for the role of asporin in breast cancer we performed data mining using the Kaplan-Meier Plotter database comprising 4142 breast cancer patients. Detailed results on relapse free survival (RFS), overall survival and distant metastases free survival are summarized in Table 1 and Supplementary Table 1A. Although high asporin expression may serve as a good prognostic marker for tumor grades 1 and 2, the opposite can be observed in high-grade breast cancer (Figure 1A). High asporin expression is associated with significantly worse RFS both in estrogen receptor positive and negative grade 3 tumors, in particular with metastasis to lymph nodes (Figure 1B, Table 1). Asporin predicts good response to endocrine therapy in estrogen receptor positive breast cancer, in particular the luminal A molecular subtype. However, it is associated with worse RFS in the chemotherapy treated basal subtype

or in the untreated luminal B subtype (Supplementary Table S1A).

Analysis of other cancer types showed significantly worse overall survival of ovarian and gastric cancer with high asporin expression. On the other hand, lung adenocarcinoma with high asporin expression had better outcome than patients with low asporin expression (Supplementary Table S1B).

Identification of asporin positive cells *in-vitro* and antibody validation

We have searched public repositories (Gene Expression Omnibus and Array Express) and found Hs578T as the only breast cancer cell line with asporin expression (Supplementary Figure S1A), which we confirmed by qRT-PCR (see below). Western blot results with several antibodies were not clear in relation to mRNA expression of positive gingival fibroblasts and negative breast cancer cell lines BT-549 and MDA-MB-231 (Supplementary Figure S2, Supplementary Table S2). Given the urgent need for a reliable antibody for subsequent experiments we saught another positive control. Asporin plays an important role in collagen mineralisation [10] and it has recently been found to be strongly upregulated during odontogenic differentiation of human dental pulp stem cells (hDPSC) [26]. We reproduced this finding (three independent samples with mean $\Delta\Delta Ct$ 4.93 which corresponds to approximately 30 fold mRNA upregulation) and western blot results were in concordance with the mRNA levels (Figure 2, Supplementary Figure S2). Presence of asporin in the indicated double-band was further confirmed by mass spectrometry (Supplementary Figure S3, Supplementary Table S3).

Asporin is regulated by BMP4, serum-starvation and 3D cultivation in Hs578T cells

We have tested multiple treatments which could modify expression of asporin in Hs578T. Expression of asporin is increased by TGF-beta in chondrocytes [27, 28], however, we observed no consistent upregulation of asporin after TGF-beta (Figure 3A, 1.3 mean fold upregulation). Other members of the TGF-beta family, BMP2 and 4, have been reported to increase asporin levels in periodontal ligament cells [29]. In breast cancer Hs578T cells, BMP4 slightly downregulated asporin and phosphorylation of focal adhesion kinase (FAK) at tyrosine 397 (2.4 and 2.1 mean fold downregulation, respectively), while BMP2 did not cause any consistent changes (Figure 3A, 1.2 mean fold downregulation). BMP4 also downregulated the asporin mRNA level (three independent samples with mean $\Delta\Delta Ct$ 3.77 which corresponds to approximately 14 fold mRNA downregulation).

Interestingly, asporin expression was markedly enhanced in serum free medium in comparison to a medium containing 10% FBS (Figure 3A and 3D,

160

Table 1: Prognostic value of high asporin expression in different breast cancer categories

subtype*	status**	prognostic value***	RFS	patients	status	prognostic value	OS	patients	status	prognostic value	DMFS	patients
any	any		0.43	3554	any	↑↑	0.046	1117	any		0.42	1609
	ER+	↑	0.067	1802	ER+	↑↑↑	0.0012	377	ER+	↑↑↑	0.0014	577
	ER-	↓↓	0.041	671	ER-		0.16	142	ER-		0.57	170
	Her2+	↓	0.081	168	Her2+		n.d.	28	Her2+		0.5	111
	Her2-	↑↑	0.048	756	Her2-		0.25	62	Her2-		0.18	82
	LN+		0.14	945	LN+		0.33	197	LN+		0.96	337
	LN-		0.61	1813	LN-	↑↑	0.034	425	LN-	↑↑	0.035	896
	G1	↑↑↑	0.0013	308	G1		0.13	135	G1		0.48	172
	G2	↑↑↑	0.0039	724	G2	↑↑↑	0.00053	287	G2	↑↑↑	0.0021	495
	G3	↓↓	0.014	723	G3		0.33	347	G3		0.21	391
	G3 ER+	↓↓	0.028	235	G3 ER+		0.58	96	G3 ER+		0.51	132
	G3 ER-	↓↓↓	0.0035	293	G3 ER-	↓	0.13	109	G3 ER-	↓	0.068	125
	G3 Her2+		0.67	87	G3 Her2+		n.d.	25	G3 Her2+		0.93	54
	G3 Her2-	↓↓↓	0.0025	239	G3 Her2-	↓	0.066	46	G3 Her2-		0.19	37
	G3 LN+	↓↓↓	0.00015	271	G3 LN+		0.32	122	G3 LN+		0.35	123
	G3 LN-		0.57	381	G3 LN-		0.38	160	G3 LN-		0.35	269
	G3 ER+ LN+	↓↓	0.041	117	G3 ER+ LN+		0.22	40	G3 ER+ LN+		0.9	48
	G3 ER+ LN-		0.85	114	G3 ER+ LN-		0.36	54	G3 ER+ LN-		0.75	81
	G3 ER- LN+	↓↓↓	0.0067	106	G3 ER- LN+		0.21	35	G3 ER- LN+		n.d.	27
	G3 ER- LN-		0.3	182	G3 ER- LN-		0.31	72	G3 ER- LN-		0.26	96
	G3 Her2- LN+	↓↓↓	0.0055	115	G3 Her2- LN+		0.18	30	G3 Her2- LN+		n.d.	14
	G3 Her2- LN-		0.48	123	G3 Her2- LN-		n.d.	16	G3 Her2- LN-		n.d.	23
luminal A	any		0.12	1764	any	↑↑	0.043	504	any	↑	0.074	918
	ER+	↑↑	0.027	1205	ER+	↑↑↑	0.0034	262	ER+	↑	0.048	419
	ER-		0.57	83	ER-		0.77	28	ER-		0.91	20
	LN+	↑↑	0.022	453	LN+	↑↑↑	0.00033	71	LN+		0.4	167
	LN-	↑↑	0.042	999	LN-		0.29	221	LN-	↑	0.065	546
	G3	↓	0.12	169	G3		0.5	73	G3		0.39	97
luminal B	any	↓↓	0.02	1002	any		0.52	320	any		0.77	361
	ER+	↓↓↓	0.0039	556	ER+		0.69	99	ER+		0.91	146
	ER-		0.88	128	ER-		0.28	22	ER-		0.81	28
	LN+	↓↓	0.017	275	LN+		0.5	43	LN+		0.61	87
	LN-		0.63	446	LN-	↓	0.091	80	LN-		0.38	171
	G3	↓	0.12	209	G3		0.24	96	G3		0.69	95
basal	any	↓	0.079	580	any		0.5	204	any		0.23	219
subtype	ER+		0.41	36	ER+		n.d.	14	ER+		n.d.	10

(Continued)

161

subtype*	status**	prognostic value***	RFS	patients	status	prognostic value	OS	patients	status	prognostic value	DMFS	patients
	ER-	↓↓	0.038	339	ER-		0.19	60	ER-		0.39	74
	LN+	↓↓	0.042	144	LN+		0.31	52	LN+		0.46	48
	LN-		0.69	291	LN-		0.67	96	LN-		0.74	134
	G3		0.17	263	G3		0.59	129	G3		0.57	136

* Molecular subtypes were derived from gene expression profiling at KMPLOT (http://kmplot.com/analysis/). With respect to the available Her2 status by FISH (see column status), the Her2 molecular subtype is not shown.

** Protein expression of estrogen receptor (ER+, ER-) may differ from mRNA expression which is used for molecular classification.

*** Arrows indicate positive (↑) or negative (↓) prognostic value of high asporin expression. Significance $p < 0.01$, $p < 0.05$ and trend $p < 0.15$ are highlighted by three, two or one arrow, respectively.

n.d., not done for less than 30 patients; RFS, relapse free survival; OS, overall survival; DMFS, distant metastasis free survival.
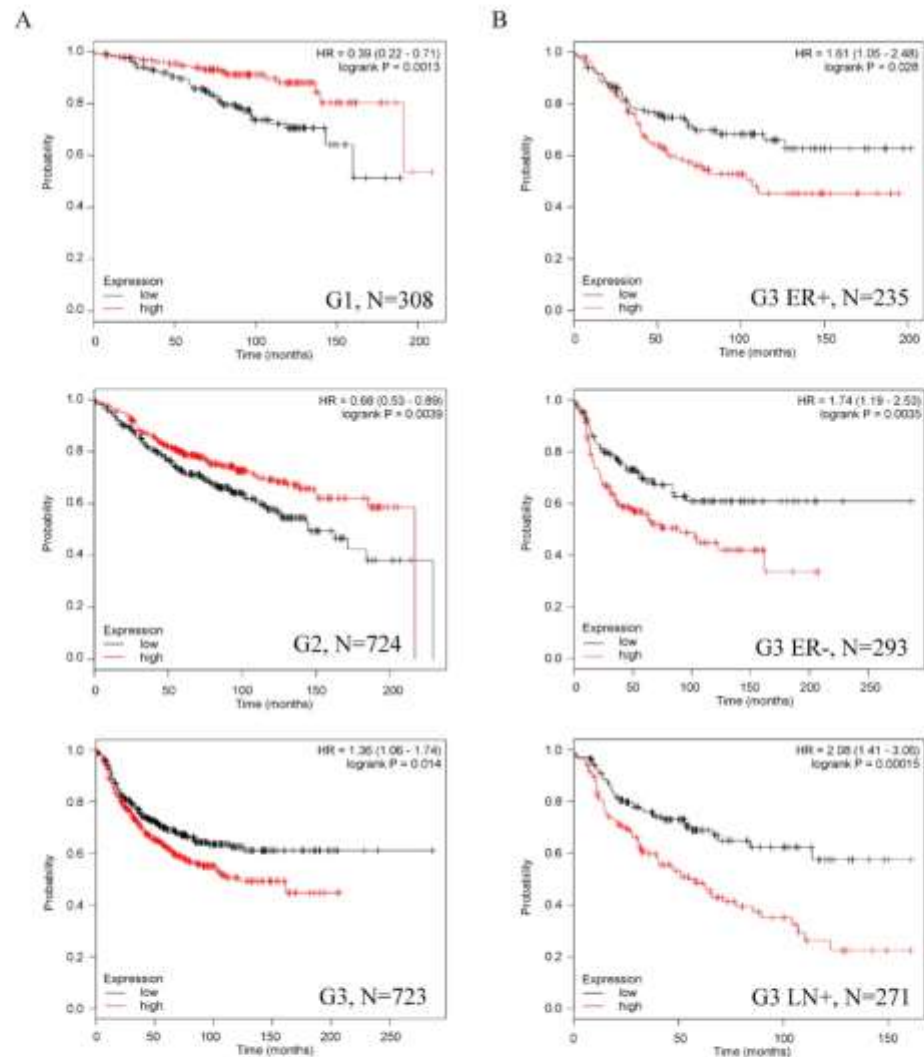
Figure 1: Asporin expression has different prognostic value in low-grade and advanced breast cancer. A. High asporin expression associates with significantly better relapse free survival (RFS) in patients with low-grade tumors but RFS is significantly worse in patients with grade 3 tumors. **B.** High asporin expression is associated with significantly worse RFS both in estrogen receptor positive and negative grade 3 tumors, in particular with metastasis to lymph nodes. HR, hazard ratio; N, number of patients in the KMPLOT analysis.

22.5 mean fold protein upregulation; mean $\Delta\Delta$Ct 4.03 which corresponds to approximately 16 fold mRNA upregulation). Lack of nutrients may correspond to stress conditions for cancer cells *in vivo* where hypoxia also often occurs. We have tested cultivation of Hs578T in hypoxic conditions both in serum-free medium and with 10% serum, but we did not observe any consistent influence of hypoxia on asporin expression (Figure 3D, 1.1 mean fold upregulation in serum-free medium).

Asporin binds to collagen type I, inhibits its fibrillogenesis and may affect the extracellular matrix [10]. In this sense, we tested whether different 3D matrices can modulate asporin expression in Hs578T. The matrix of collagen type I (800 µg/ml) and Geltrex had no effect but 3D polystyrene Alvetex increased asporin expression (Figures 3B-3C; 10.0 mean fold protein upregulation; mean $\Delta\Delta$Ct 2.8 which corresponds to approximately 8 fold mRNA upregulation). Neither upregulation of asporin by serum starvation nor Alvetex matrix was associated with changes in phospho-FAK (Figure 3A-3B).

Downregulation of asporin inhibits invasion of Hs578T through collagen matrix

In order to further test the importance of asporin expression in Hs578T we prepared three stable clones with shRNA against asporin. We confirmed downregulation of asporin expression (Figure 3E; mean $\Delta\Delta$Ct 5.3 which corresponds to approximately 40 fold mRNA downregulation) which did not affect proliferation, adhesion, migration or spheroid growth (data not shown).

Importantly, downregulation of asporin inhibited invasion through the collagen type I matrix (Figure 3F).

We also generated stable clones of MDA-MB-231 and BT-459 cells with asporin ORF (open reading frame with myc and DKK tag) expression vectors. Expression of mRNA was comparable to Hs578T while immunoblotting detected a band with slower mobility (Figure 2; immunoblotting for MDA-MB-231 is not shown). Expression of asporin ORF did not affect proliferation, adhesion, migration or invasion of stable clones of BT-549 or MDA-MB-231 (data not shown). Interestingly, we were not able to prepare stable clones with expression vector with full asporin sequence (including 5´ and 3´UTR).

Recombinant asporin enhances invasion through collagen type I matrix

Asporin binds to collagen type I, inhibits its fibrillogenesis and may affect extracellular matrix [7, 10]. Prof. Oldberg kindly provided us with the recombinant asporin which we used in invasion assays. First, we tested invasion of the three breast cancer cell lines mentioned above through various concentrations of collagen type I matrix (50, 200 and 800 µg/ml; data not shown) and selected a concentration of 200 µg/ml which was also used by Sodek et al. [30]. MDA-MB-231 and BT-549 cells invaded faster through the collagen matrix which was prepared with the recombinant asporin in comparison with the matrix without asporin (Figure 4A-4B). Invasion of Hs578T was comparable in both matrices (data not shown).
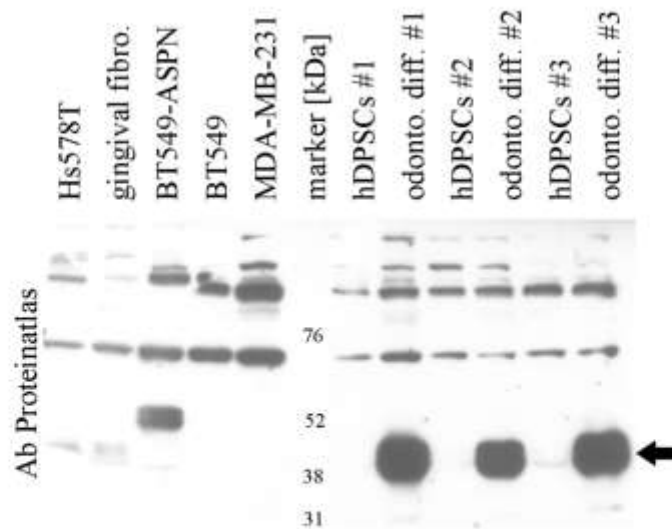

Figure 2: Validation of asporin antibody (Proteinatlas #HPA008435) in breast cancer cells, gingival fibroblasts and human dental pulp stem cells (hDPSCs). Black arrow indicates strong upregulation of asporin in hDPSCs after odontogenic differentiation. BT-549-ASPN with stable transfected asporin open reading frame sequence display a band with higher molecular weight (a similar band was observed also in MDA-MB-231-ASPN cells, data not shown). Validation of other antibodies is provided in Supplementary Figure S2.

163
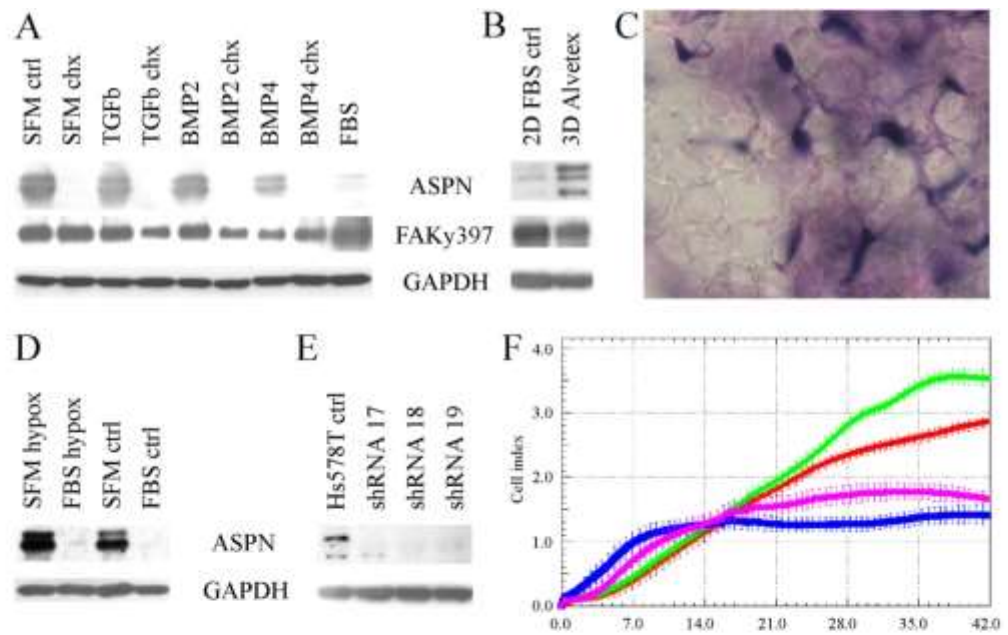
Figure 3: Glycoprotein asporin in Hs578T breast cancer cells. A. Asporin is consistently upregulated by serum starvation (SFM-ctrl, serum-free medium control) and downregulated by two-day treatment with 100 ng/ml bone morphogenetic protein 4 (BMP4). Mild asporin downregulation by TGF-beta was not reproduced in other experiments. BMP4 also decreased phosphorylation of focal adhesion at tyrosine 397 (FAKy397). Inhibition of translation by cotreatment with 10 μg/ml cycloheximide led to downregulation of asporin but not of FAK. All treatments were performed in serum free media for 48 hours. **B.** Expression of asporin is upregulated by 12-day cultivation of Hs578T cells in Alvetex polystyrene scaffold (3D) in comparison to normal 2D conditions. **C.** Formallin-fixed paraffin-embedded Hs578T cells (stained with hematoxylin-eosin) in polystyrene Alvetex scaffold (transparent stellar shapes; magnification 1000x). **D.** Four-day cultivation of Hs578T cells in hypoxic conditions does not consistently change asporin expression. **E.** Expression of asporin is successfully downregulated with three different shRNAs. **F.** Invasion of Hs578T cells (parental cells are displayed in green, control cells with scrambled shRNA are in red) through collagen matrix is inhibited by shRNAs 17 and 19 (blue and purple, respectively). Similar inhibition was observed also for the shRNA 18 (data not shown). Reading of cell index was every 10 minutes up to 42 hours with the real time cell monitoring xCELLigence instrument (error bars indicate standard deviations from quadruplicate measurement). All experiments were performed at least three times and representative blots/chart are shown.
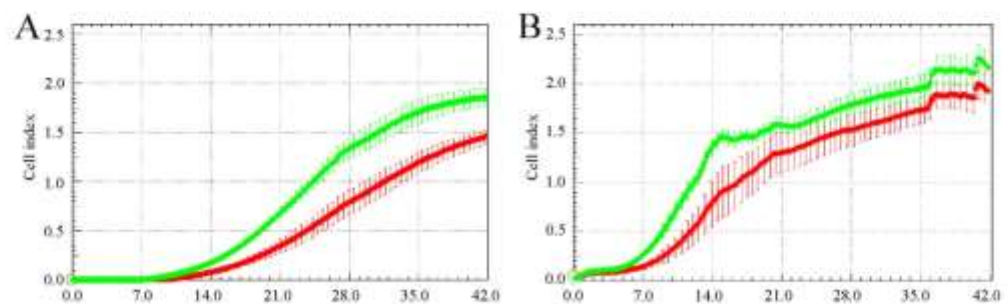

Figure 4: Recombinant asporin enhances invasion of MDA-MB-231 A. and BT-549 B. breast cancer cells through collagen matrix. Both models were seeded at a density of 75000 cells per well into CIM plates coated by collagen stiffened in presence (green) or absence (red) of 10 nM recombinant asporin. The bottom chambers were filled with complete medium containing 10% FCS (fetal calf serum) as a chemoattractant. Reading of cell index was every 10 minutes up to 42 hours with the real time cell monitoring xCELLigence instrument (error bars indicate standard deviations from quadruplicate measurement). All experiments were performed at least three times and representative charts are shown.

164

RNA scope *in situ* hybridization detects asporin in cancer associated fibroblasts

We have identified asporin as a novel breast cancer related protein by laser microdissection and microarray analysis [12], in particular in invasive lobular carcinomas. However, asporin has been repeatedly reported to be expressed by cancer associated fibroblasts (CAFs) or other stromal cell types [15, 23, 24, 25]. Although laser microdissection is designed for single cell isolation we cannot exclude the possible capturing of adjacent cells, in particular in invasive lobular cancers which may grow in thin lines. As there is no reliable antibody for immunohistochemistry, we decided to use RNA scope technology with gene-specific probe pairs for detection of asporin in breast cancer tissues. As a positive control (Figure 5) we used endometrium (Supplementary Figure S4) and the above mentioned hDPSCs after odontogenic
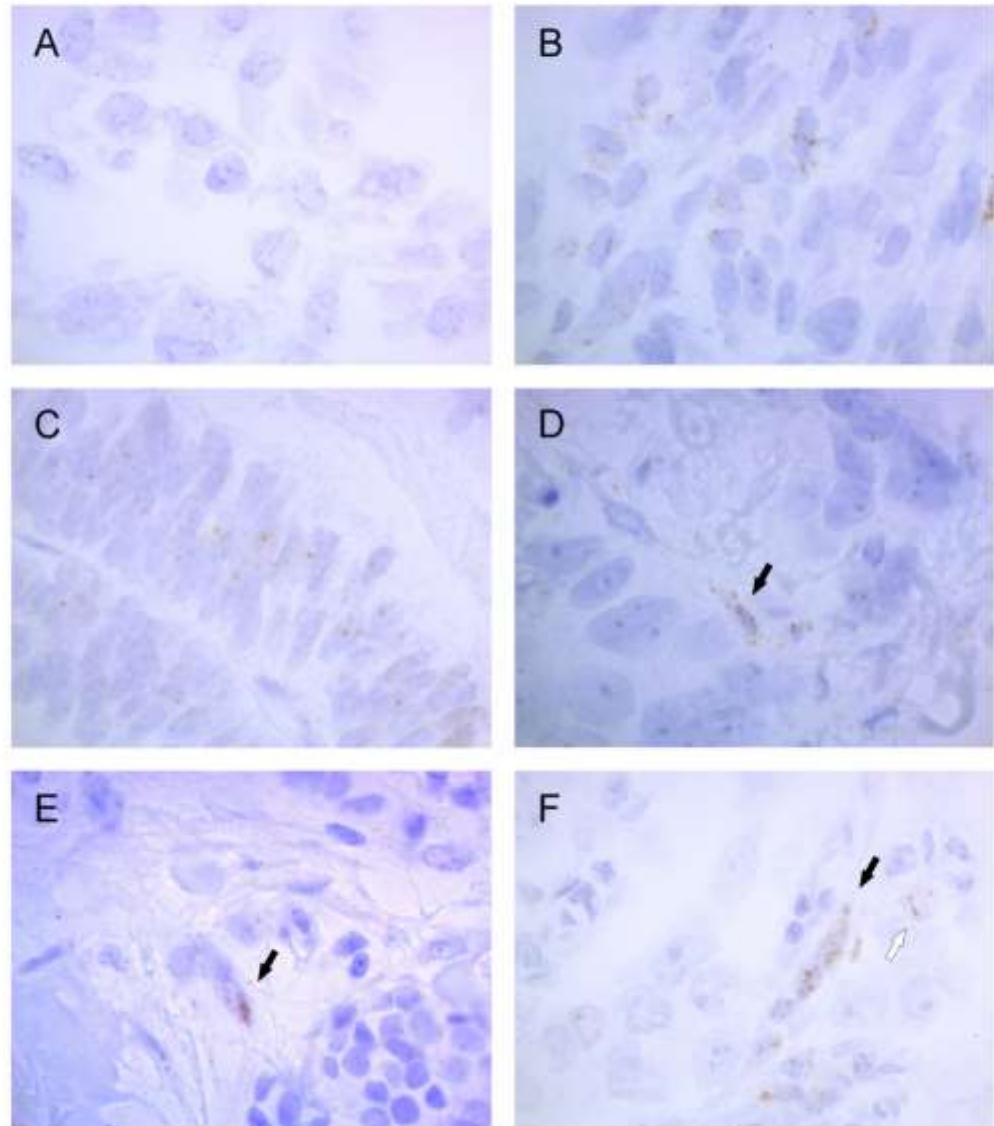
Figure 5: Asporin is detected by RNA scope *in situ* hybridization in invasive breast cancer. Human dental pulp stem cells before **A.** and after **B.** odontogenic differentiation, as well as endometrium **C.**, were used as positive controls. Specific dot-like positivity was found in cancer associated fibroblasts both in ductal **D.** and lobular **E-F.** invasive carcinomas (black arrows). Asporin positivity was rarely observed also in cancer cells (F, white arrow). Magnification is 1000x while lower magnification (400x) of the same samples is provided in the Supplementary Figure S5.

165

differentiation. Next, we hybridized 43 lobular and 7 ductal invasive breast cancer tissues. However, frequencies of specific dot-like signals were very low which precluded statistical analysis. Asporin dot-like positivity was found in CAFs both in ductal and lobular invasive carcinomas (Figure 5 and Supplementary Figure S5).

Downregulation of asporin in cancer associated fibroblasts attenuates their coordinated invasion with breast cancer cells

We have tested three different CAFs from patients with grade 3 breast cancer (full details are in the Methods

section) and selected those with the fastest invasion and proliferation for knock-down and co-culture experiments with T47D cells. This breast cancer cell line has low invasive potential alone and represents luminal B subtype [31] (please note poor RFS of luminal B patients with high asporin expression in Table 1). Transwell experiments indicated coordinated invasion of CAFs with T47D cells which was attenuated by asporin knock down in CAFs (Figure 6A) (Figure 6B and 6D).

Similarly to Hs578T, asporin expression was enhanced in CAFs by their 3D culture in Alvetex scaffold (Figure 6C). Analysis of aspartic acid (D) repeat length of asporin revealed that our breast CAFs have D13/
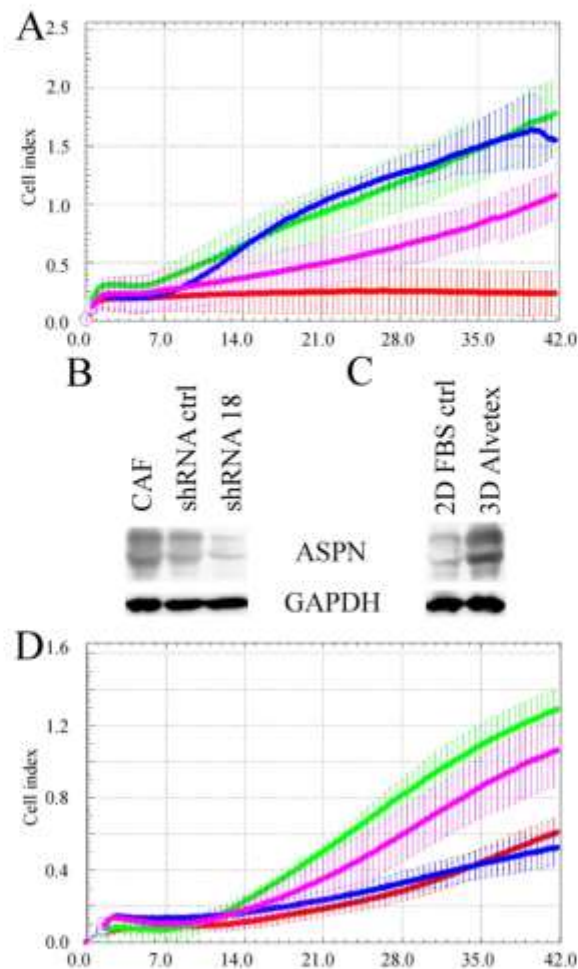
Figure 6: Invasion of breast cancer associated fibroblasts and T47D cells is attenuated by asporin downregulation. **A.** Transwell experiment indicate coordinated invasion of cancer associated fibroblasts with T47D cells through extracellular matrix composed of collagen, matrigel and crosslinking ribose (70000 CAFs, green; 70000 T47D, red; CAFs + T47D, 35000 each, blue; 35000 CAFs, purple). Western blot analysis shows downregulation of asporin by shRNA **B.** and its upregulation upon 3D cultivation of CAFs in Alvetex polystyrene scaffold **C. D.** Downregulation of asporin attenuated invasion of CAFs and T47D cells (70000 CAFs shRNA ctrl, green; 70000 CAFs shRNA18, red; CAFs shRNA ctrl + T47D, 35000 each, purple; CAFs shRNA 18 + T47D, 35000 each, blue). All experiments were performed three times and representative blots/charts are shown.

166

D14 genotype which has recently been associated with poor metastasis-free survival of prostate cancer patients (Supplementary Figure S6) [32].

DISCUSSION

The need for complex 3D culture models to unravel novel pathways and identify accurate biomarkers in breast cancer has been recently reviewed [33]. We have observed asporin upregulation during 3D culture of Hs578T cells in polystyrene scaffold Alvetex, but not in collagen or Geltrex matrices. Similarly, expression of tropomyosin kinase 2 was increased by 3D cultivation on collagen-coated polystyrene scaffold Alvetex [34] in comparison to 2D culture. Stiffness of the matrix is a major factor which could affect tumor progression [1, 35]. Polystyrene scaffold is stiffer than collagen or Geltrex matrices and may better simulate solid stress in tumors *in vivo*.

In contrast to chondrocytes or peridontal ligament cells [28, 29] we found no asporin modulation by TGF-beta or BMP2 in Hs578T cells. Nevertheless, asporin was slightly downregulated by BMP4. Similarly to BMP7, expression of BMP4 is low in invasive breast cancer and its signalling has been implicated in the maintenance of tumor dormancy [36–38]. Unexpectedly, we observed upregulation of asporin in serum-free cultivation. We speculate that this induction could lead to enhanced motility towards better sources of nutrients. Upregulation by starvation has also been reported for other proteins in cancer, e.g. glucose regulated chaperones [39, 40].

MDA-MB-231 and BT-549 cells invaded faster through collagen matrix which was prepared with the recombinant asporin in comparison to the matrix without asporin. This may be related to less dense matrix due to inhibition of collagen fibrillogenesis by asporin [10]. Invasion of Hs578T cells, which is accompanied by endogenous asporin production, was not further enhanced in the matrix with recombinant asporin. On the contrary, asporin downregulation decreased invasion of Hs578T cells through the collagen matrix. As shown by Satoyoshi et al. [24] asporin itself can promote invasion of both cancer cells and cancer associated fibroblast by interaction with CD44 and activation of Rac1. We have not observed any phenotypic change of BT-549 and MDA-MB-231 cells after stable transfection, however, asporin from ORF sequence may not reflect its native form.

We found asporin dot-like positivity mostly in cancer associated fibroblasts both in ductal and lobular invasive carcinomas which is in concordance with studies both from breast or other solid tumors such as prostate or gastric cancer [23, 24]. Cancer associated fibroblasts resemble myofibroblasts with smooth muscle actin which is also expressed in so called myoepithelial Hs578T [41]. Interestingly, Hs578T and SNB-75 cancer cells with asporin expression have the highest expression of smooth muscle actin (Supplementary Figure S1B). Asporin has also been reported to be expressed by gastric cancer cell lines and to enhance their oncogenic properties [42]. As asporin belongs to secreted proteins, regulation of its expression both in CAFs and cancer cell is relevant for alteration of the tumor microenvironment.

Maris et al. [25] have recently described the tumor suppressive potential of asporin in several breast cancer models. This may be true for a subset of breast cancer patients. However, *in silico* data mining on a large cohort of breast cancer patients showed that high asporin expression is associated with significantly worse RFS both in estrogen receptor positive and negative grade 3 tumors, even with metastasis to lymph nodes. Asporin was also associated with worse RFS in the chemotherapy treated basal subtype and in the untreated luminal B subtype. Van't Veer et al. [43] have recently assembled 72 breast cancer-related gene expression datasets, containing approximately 5700 samples altogether, where asporin is included in an extracellular matrix module comprising 58 genes. Importantly, asporin was found as one of the most upregulated genes in bone metastases of breast and prostate cancer [17, 44]. This is in accord with its primary function in skeletal development (mainly bones, cartilage and teeth), regulation of collagen fibrillogenesis and induction of mineralization [6, 7, 10, 45]. Hurley et al. [32] have recently described different metastatic potential of prostate cancer related to the polymorphism of aspartic acid (D)-repeat of asporin, which may also be important in breast cancer. D14 allele (heterozygous with D13 as well as in combination with any other allele) was associated with poor patient outcome, which is in line with decreased invasion of our breast CAFs upon asporin downregulation (genotype D13/D14). The unique asporin polyaspartate repeat binds calcium [10] and the D14 allele was also linked to predisposition for knee joint osteoarthritis whereas the D13 allele was more prevalent in healthy individuals [46]. The role of other alleles, such as D15 (homozygous in breast cancer cell line Hs578T) or D16 (heterozygous with D13 in bacteria used for production of recombinant asporin) awaits elucidation. Maris et al. [25] used transcript variant 1 with 13 aspartic acids (NM_017680.4) for both *in vitro* and *in vivo* experiments and therefore their results are in good concordance with the study of Hurley et al. [32] which decribed tumor suppressive effects of the D13 allele in prostate cancer.

In conclusion, we found that asporin can be downregulated by bone morphogenetic protein 4 in Hs578T cells and its upregulation may be facilited by serum-free cultivation or by three dimensional growth. Downregulation by shRNA inhibited invasion of Hs578T as well as of CAFs and T47D cells while invasion of asporin-negative MDA-MB-231 and BT549 breast cancer cells through collagen type I was enhanced by recombinant asporin. In contrast to these observations, tumor suppressive effects of asporin were published by others. Importantly, opposing effects of different asporin alleles

have recently been reported in prostate cancer. Its dual role is also documented by our metaanalysis of publicly available data which found asporin to be associated with better prognosis in low-grade tumors but not in high-grade breast cancer. Besides other investigations, large scale analysis of aspartic acid (D)-repeat polymorphism will be needed for clarification of asporin dual role in progression of breast cancer.

MATERIALS AND METHODS

Cell culture

Authenticated and mycoplasma-free breast cancer cell lines were used for all experiments. Hs578t cells, obtained from ECACC (Salisbury, UK), were grown in Dulbecco's Modified Eagle's Medium (DMEM; Life Technologies, Carlsbad, CA) supplemented with 10 % foetal bovine serum (FBS, ThermoFisher, MA, USA), 1% penicillin/streptomycin (Life Technologies) and insulin (10 μg/ml) (Life Technologies). MDA-MB-231, BT-549 and T47D, obtained from American Type Culture Collection (Rockville, MD, USA), were cultivated in DMEM supplemented with 10% FBS and 1% penicillin/streptomycin.

Normal human gingival fibroblasts (kindly provided by prof. Jitka Ulrichova from Department of Medical Chemistry and Biochemistry at Palacky University Olomouc) were obtained from medically healthy donors who were clinically free of periodontal disease. Samples of gingiva were obtained from patients undergoing surgical removal of third molars at the Department of Oral and Maxillofacial Surgery (University Hospital, Olomouc). The tissue acquisition protocol adhered to the requirements of the local Ethics Committee. All patients had signed written informed consent.

The gingival tissues were washed three times in phosphate buffered saline (PBS) containing antibiotics (pamycon and colinomycin). The excised gingiva were cut into 1 mm pieces, plated in Petri dishes (10 cm diameter) and cultured in DMEM with 10% FBS and 1% penicillin/streptomycin. Explants were incubated in humidified atmosphere with 5 % CO_2 at 37°C. Cells were fed weekly until the fibroblasts reached confluence. After 4-6 weeks cell cultures were trypsinized and transferred into 75 cm^2 cultivation flasks. Cells were used between the 3 and 10 passages for experiments.

Breast cancer associated fibroblasts

Isolation of primary breast cancer associated fibroblasts was performed according to modified protocol from Giannoni et al. [47, 48]. Tissue samples were obtained from patients undergoing surgical removal of breast cancer at Masaryk Memorial Institute (Brno, Czech Republic). The tissue acquisition protocol adhered to the requirements of the local Ethics Committee and all patients had signed written informed consent. Pieces of breast cancer tissue were minced and put on a 100 mm tissue culture dish. Sterile microscopic slide was put over the tissue pieces, slightly pushed, and overlaid with DMEM medium supplemented with 20% FBS, 2% penicillin/streptomycin, 100 μg/ml kanamycin (Serva, Germany), 2.5 μg/ml amphotericin B (ThermoFisher Scientific). Approximately 3 weeks after initial seeding of tissue pieces, the microscopic slide was removed and fibroblasts were further maintained as a regular cell culture. After approximately two weeks in culture, the cultivation medium was changed to DMEM with 10% FBS and 1% penicillin/streptomycin. To confirm homogeneity of isolated fibroblasts (denoted as BCAF3, BCAF4 and BCAF5), we performed FACS analysis of three selected surface fibroblast markers using BD FACSVerse flow cytometer. Cell suspension obtained by trypsinisation was stained with primary antibodies against FAP (human fibroblast activation protein alpha, MAB3715, R&D Systems), FSP (fibroblast surface protein, F4771, Sigma-Aldrich), followed by staining with secondary antibody Alexa Fluor 488 conjugated donkey anti-mouse IgG (H+L) (A21202, ThermoFisher Scientific); and primary antibody anti-fibroblast FITC (130-100-134, Miltenyi Biotec). Dead cells were excluded from the analysis using LIVE/DEAD Fixable Far Red Dead Cell Stain Kit (L10120, ThermoFisher Scientific). Analysis showed uniform positivity for all three markers and simultaneously, no positivity for non-fibroblast markers EpCAM (324220, BioLegend), CD31 (303104, BioLegend), and CD45 (304054, BioLegend) was detected.

We have isolated fibroblasts from three patients with invasive breast cancer, no special type (samples were denoted as BCAF3, BCAF4 and BCAF5) with the following clinicopatological parameters: BCAF3 - grade 3, pT2 (25 mm), pN0, M0, Her2 positive, ER 0%, PR 0%, AR 100%, Ki67 52%; BCAF4 - grade 3, pT2 (35mm), pN1a(3/15), M0, Her2 negative, all ER, PR and AR 100%, Ki67 30%; BCAF5 - grade 3, pT3 (55mm), pN0, Her2 negative, ER 100%, PR variable (0-3-80%), AR 80%, Ki67 55%. BCAF4 grew fastest and were more invasive than BCAF5 (data not shown; BCAF3 were not tested for invasion due to their slow proliferation). BCAF4 (passages from 9 to 11) were used as a model of cancer associated fibroblasts in all experiments.

Transient asporin knockdown in BCAF4 (specific sh18 and control scrambled shRNA; please see the sequences below) was performed with Viromer® transfection reagent (Lipocalyx, Hall, Germany) according to manufacturer protocol. Standard medium was changed after 12 hrs and one day later the cells were starved in serum free medium for 6 hrs before invasion experiments (see below).

Human dental pulp stem cells and odontogenic differentiation

The isolation and culture of dental pulp stem cells (hDPSC) were performed according to protocol described previously [49]. Impacted wisdom teeth were obtained from healthy donors with informed consent at the Department of Dentistry under approved guidelines of the Ethical Committee of the Faculty Hospital in Hradec Kralove. For odontogenic differentiation, hDPSCs were cultured until confluence and then incubated with alpha-MEM supplemented with 10 mM beta-glycerolphosphate, 50 μM ascorbic acid and 0.1 μM dexamethasone for one week [26].

Experimental treatments of Hs578T cells

Multiple treatments which could modify expression of asporin in Hs578T were tested. TGF β (transforming growth factor β) in concentration of 10 ng/ml, BMP2 and BMP4 (bone morphogenetic protein 2 and 4) in concentration 100 ng/ml were added to 24 hrs starved Hs578t cells in serum deprived media for 48 hrs. Additionally, each cytokine was combined with cycloheximide treatment (Sigma-Aldrich, Germany) in final concentration 10 μg/ml for the same period of time. Long term cultivation of Hs578t cells in presence of 1 μM dexamethasone (Sigma-Aldrich) was for three weeks. Expression of asporin was also monitored after 4 day cultivation of Hs578T cell under hypoxic conditions with 3% O_2.

Cultivation of Hs578t in 3D was performed with polystyrene scaffold Alvetex® (Reinervate, UK), Geltrex® (Life Technologies) or collagen type I (Purecol®, Advanced BioMatrix, Carlsbad, CA) and lasted for 12 days with medium exchange every 4 days. Geltrex thin gel method was performed by mixing 500 μl of Geltrex with 1000 μl cultivating media followed by 1.5 hour incubation in 37°C. Collagen in a concentration of 800 μg/ml was diluted to 150 mM NaCl buffered with 20 mM Hepes (pH 7.4), and left to solidify in 37 °C in cell culture incubator.

Generation of stably transfected cell lines

For asporin overexpression MDA-MB-231 and BT-549 were transfected either with ASPN full length sequence (TrueClone, pCMV6-AC, Origene) or with open reading frame (TrueORF, pCMV6, Origene) sequence using Neon Transfection System (Life Technologies). Stable cell lines were selected with 0.5 mg/ml geneticin for 2 weeks and then kept under low selection pressure at 0.1 mg/ml geneticin.

For asporin knockdown Hs578t cells were similarly transfected with short hairpin RNAs (sh17 AGT TGA AAT ACC TCC AGA TAA TCT TCC TT, sh18 TCC AAC AGT GCC AAA GAT GAA GAA ATC TT or sh19 CGA GTA TGT GCT CCT ATT ATT CCT GGC TT in pRS plasmid, OriGene) and selected with 1 μg/ml puromycin. Both overexpression and knockdown of ASPN were checked by RT-PCR and western blotting. Control cells were transfected with empty pRS plasmid or scrambled shRNA (GCA CTA CCA GAC CTA ACT CAG ATA GTA CT; Origene #TR30012).

Adhesion, proliferation and migration assays

Experiments were carried out using xCELLigence Real-Time Cell Analyzer (RTCA) DP instrument (Roche diagnostic, GmbH, Germany). For adhesion assay 40 000 cells were seeded on E-plate 16 surface and increasing cell index was monitored every 30 seconds for 4 hrs, followed by recording cell proliferation for additional 64 hrs.

For migration monitoring, fully confluent cell population in a standard cultivation medium supplemented with 10 μg/ml mitomycin C (Roche) was scratched by pipette tip. Migrating cell populations were monitored every second hour for 24 hrs. Experiments were repeated at least three times.

Invasion assays

Invasion assays were carried out using xCELLigence Real-Time Cell Analyzer (RTCA) DP instrument (Roche). Parental breast cancer cell lines (MDA-MB-231, BT-549 and Hs578t) were tested with collagen I matrix prepared with or without recombinant asporin (kind gift from Professor Åke Oldberg, University of Lund). Bottom side of CIM plate inserts (Roche/Acea, CA, USA) was pre-coated with 30 μl of cooled 200 μg/ml collagen I (Purecol®, Advanced BioMatrix, Carlsbad, CA) prepared as described by Kalamajski et al. [10]. Resulting aqueous phase was pipetted off after 30 minutes leaving a thin collagen film. The upper side of the inserts was coated with 30 μl of cooled collagen I with or without 10 nM recombinant asporin and the solution was incubated at 37°C for 6 hrs. The lower chamber of the CIM plate was filled with 170 μl 10% FCS DMEM as a chemoattractant. Background signal was measured after addition of 30 μl of serum-free media per well followed by 1 hour of incubation at 37°C. Cells which were starved in medium without serum for 6 hrs were harvested and seeded in amount of 75000 cells per insert well. Cell index was monitored every 10 minutes for 42 hrs.

For stably transfected cells (MDA-MB-231, BT-549 and Hs578t) the invasion barrier was prepared from 200 μg/ml pure collagen I (collagen G1, Matrix BioScience GmbH Germany) and CIM plate was coated as described above. Collagen solution in the upper part of wells was incubated for 1 hour at 37°C and then 75000 starved cells were added. Increasing cell index was recorded every 10 minutes for 24 hrs. All experiments were repeated at least three times.

169

Modified protocol was used for CAFs and T47D cells. The upper side of the inserts was coated with 30 µl of mixture of 200 µg/ml neutralized collagen I (Advanced BioMatrix, Carlsbad, CA), Matrigel® (200 µg/ml; Corning, MA, USA) and 15 mM D-ribose (Sigma-Aldrich) and then incubated at 37°C for 1 hr [50]. Lower chamber of the CIM plate was filled with 170 µl 10% FCS DMEM as a chemoattractant. Collagen-Matrigel-ribose clot was overlaid by 30 µl of serum-free media. Assembled coated CIM plates were left at 37°C in the incubator for 1 hr followed by reading background signal. Starved cells were harvested and seeded in amount of 70000 cells per insert well or 35000 per cell line in co-culture experiments. Cell index was monitored every 10 minutes for 42 hrs. Experiments were repeated three times.

Spheroid assay

Spheroid cell culture was performed by the liquid overlay method [51]. Briefly, all cell models were seeded at 10 000 cells on convex face of 1.5% agarose LE (Promega, WI, USA) in DMEM. The plate was then spun at 1000 rpm for 15 minutes to induce cell aggregation and subsequently each well was overlayed by 100 µl of standard cultivation media [52]. Formation of mammospheres was recorded after 4 days.

In situ hybridization by RNAscope and immunohistochemistry

Archival tissue samples of 50 invasive breast cancer, among them 43 invasive lobular (ILC) and 7 invasive ductal carcinomas (IDC) were obtained from patients between years 1999 and 2006. We have previously found high levels of asporin in ILC than IDC [12], therefore, ILC cases were specifically selected to reach sufficient sample size. The study was approved by the Ethics Committee of the Faculty of Medicine and Dentistry. Control cells from *in vitro* culture were homogeneously distributed in an agarose gel matrix, creating an artificial tissue, and then formalin-fixed and paraffin-embedded [53]. These controls were used also for immunohistochemistry validation (Supplementary Table S2), however, the results didn´t reflect the mRNA levels of asporin in any method tested (antigen retrieval by boiling either in a microwave for 20 min in citrate buffer pH6 or in a water bath in EDTA pH9; visualization by DAB chromogen and Dual Link secondary antibody, Dako) [13].

In situ hybridization for asporin RNA was performed using the RNAscope® 2.0 kit according to the manufacturer´s manual (Advanced Cell Diagnostics, USA; [54]. Briefly, 5 µm formalin-fixed, paraffin-embedded tissue sections were deparaffinized in xylene and pre-treated with heat (15 minutes, boiling in water bath) and protease (undiluted protease solution, 30 minutes, 40°C, DAKO hybridizer). Length of the main hybridization

step was increased from 2 to 3 hrs (40°C). Then signal amplification system which included hybridization with preamplifier, amplifier and label probes, and subsequently chromogenic detection with DAB (3,3'-diaminobenzidine) was used. Positive staining was identified as brown punctate dots in the cell. Negative control probes (DapB) and positive control probes (POLR2A) were also included for each examined slide.

With respect to the RNA scope technology, several aspects should be commented upon. The frequency of dot-like positivity was low, while other types of positivities occurred, for example in areas with inflammation or rare homogenous nuclear positivity in negative control. The results may also be affected by variability of patient samples. Although all samples were fixed and embedded in our pathology laboratory according to standard protocols, we observed differences in signals with RNA scope positive control. Dreyer et al. [55] compared RNA scope with DNA *in situ* hybridization and PCR detection of HPV infections. Surprisingly, RNA scope signals occurred even after digestion with RNase. In our opinion, RNA scope technology may support results obtained by other methods but is not a robust enough method for routine use as yet.

Mass spectrometry

Cells were lysed and 200 µg of proteins was separated on SDS-PAGE as described below. Gel was coomassie stained and relevant molecular weight bands were excised, destained, reduced by 50 mM Tris(2-carboxyethyl) phosphine hydrochloride (Sigma-Aldrich), alkylated by 50 mM Iodoacetamide (Sigma-Aldrich) and digested with proteomic grade trypsin (Promega) at 37°C overnight. Mass spectrometric analysis was performed on an Orbitrap Fusion (ThermoFisher Scientific) instrument fitted with a Proxeon Easy-Spray ionization source, coupled to an Ultimate 3000 RSLCnano chromatograph. Ten microliter of sample was loaded on a µ-Precolumn C18 PepMap 100 (300 µm x 5 mm, 5 µm, 100 Å pore size) desalting column (ThermoFisher Scientific) "in-line" with a PepMap RSLC (75 µm x 50 cm, 3 µm, 100 Å pore size) analytical column (ThermoFisher Scientific) heated at 35°C. The peptides were subsequently separated on the analytical column by ramping the organic phase from 5% to 35% during a total run time of 165 minutes. The aqueous and organic mobile phases were 0.1% formic acid diluted in water or acetonitrile, respectively. There were two parallel experiments running simultaneously on mass spectrometer. The first one was single FTMS scan with resolution was set to 120,000 and precursor ions were scanned across an m/z range of 400- 1600. The second experiment was targeted MS2 with HCD collision and Orbitrap detector. Resolution was set to 30,000 and HCD collision energy to 35%. The transitions of most suitable peptides were designed and resulting spectra were evaluated in Skyline-daily 2.6.1.6899 software [56].

170

Production of recombinant asporin

One milliliter aliquot of frozen bacterial culture containing the plasmid carrying the gene for asporin was thawed and suspended in 100 ml of LB medium (10 g NaCl, 10 g tryptone, 5 g yeast extract per liter) with kanamycin (75 mg/L) and incubated overnight on a heated rotation shaker (37°C) [10]. The next day this starting culture was added to 1 liter of fresh LB medium containing kanamycin. The culture was again incubated on rotation shaker until reaching the sufficient density for the induction (O.D.$_{600}$=0.6). Expression of the recombinant asporin was induced by Isopropyl-β-D-thiogalacto-pyranoside (IPTG, final concentration 1 mM). Bacteria were incubated for another 5 hours and then harvested by the centrifugation. Bacterial pellet was resuspended in denaturing lysis buffer (8 M urea buffered by 50mM Tris, pH 8), sonicated (10x10 s, 2 min intervals) and centrifugated (10000g, 10 min, room temperature). The obtained supernatant was used for protein purification using the metaloaffinity chromatografy with Ni-NTA agarose. Washing and elution was performed using buffers with decreasing pH (8 M urea, pH 6.3; 5.9 and 4.5). Glutathion-S-transferase recombinant asporin (H00054829-P01, Abnova, Taiwan) was purchased for quantification of asporin by mass spectrometry (Supplementary Figure S3).

RNA isolation, reverse transcription and quantitative PCR

Total RNA isolation was performed by the RNeasy Mini Kit (Qiagen), quantified by Nanodrop, pretreated with Dnase I (Invitrogen, Carlsbad, CA, USA) and reverse transcribed with random hexamers and SuperScript III Reverse Transcriptase (Invitrogen). The real-time polymerase chain reaction reactions were performed using specific primers and probes (ASPN forward 5´GGG TGA CGG TGT TCC ATA TC 3´, reverse 5´TTG GTG GTA AGC CTT TAG GAA 3´, probe 5´ BHQ1-TTG CAG AAG CAA AAC TGA CC-FAM 3´; TBP forward 5´CAC GAA CCA CGG CAC TGA TT 3´, reverse 5´TTT TCT TGC TGC CAG TCT GGA C 3´, probe 5' BHQ1-TCT TCA CTC TTG GCT CCT GTG CAC A-HEX 3') and Probes Master Mix for 50 cycles of denaturation, annealing and extension (95–60–72°C each for 20 s) at the LightCycler 480 instrument (Roche). Relative quantification was carried out according to the ΔΔCt method using TBP as a reference gene [57].

Protein extraction and western blot analysis

Cells were harvested into RIPA buffer (50 mM Tris HCl pH 8.0; 150 mM NaCl; 1% NP-40 0.5% sodium deoxycholate; 0.1% SDS) supplemented with protease/phosphatase inhibitor cocktail (Roche). Twenty micrograms of whole cell lysate were separated by electrophoresis in 10% Bis-Tris polyacrylamide gel followed by blotting to nitrocellulose membrane. Non-specific binding sites were blocked by incubating the blots for 2 hrs at room temperature with 5% (w/v) non-fat dry milk in PBS. Blots were incubated overnight with primary antibodies at the following concentrations: anti-ASPN (1:1000; # HPA008435; Sigma-Aldrich, Protein Atlas; plus other antibodies specified in Supplementary Table S2), anti-FAK pY397 (1:500, Life Technologies), and anti-GAPDH (1:25000; Sigma-Aldrich) was used as a loading control. Secondary antibodies were as follows anti-rabbit IgG, HRP-linked Antibody (#7074) and Anti-mouse IgG, HRP-linked Antibody (#7076), both purchased from Cell Signaling Technology, MA, USA. Signal detection was performed with Dura/ Femto ECL western blotting substrate (ThermoFisher Scientific). Analysis of optical density was performed using ImageJ software.

In silico data mining

Expression of asporin in cancer cell lines was checked either in Gene Expression Omnibus database at the National Center for Biotechnology Information or in ArrayExpress database at the European Bioinformatics Institute. The prognostic and predictive value of asporin (ASPN; 219087_at) was further evaluated by Kaplan-Meier Plotter (KMPLOT) [58], where all settings were left at default values (database version 2014; breast cancer analysis was performed from 8th to 10th September 2015; other cancer types were analysed on 23rd February 2016).

Genotyping of the asporin D-repeat polymorphism

The aspartic acid (D) - repeat polymorphism located in the N-terminal region of the ASPN gene was evaluated according to Hurley et al. [32]. Briefly, PCR products (fluorescent forward primer 5´ 6-FAM–ATT CCT GGC TTT GTG CTC TG and non-labeled reverse primer 5´ TGG CTT CTT GGC TCT CTT GT) were mixed with GeneScan™ 500 TAMRA™ Size Standard and Hi-Di™ Formamide (Applied Biosystems, Foster City, CA, USA) and denatured 3 min at 95°C. Then, it was separated on ABI PRISM® 3130 Genetic Analyzer and obtained data were analysed and visualised using GeneScan® Analysis Software (Applied Biosystems). Genotypes of BCAF3, 4 and 5 were D13/D14, D13/D14 and D13/D15, respectively.

Genotype of three samples (BCAF4, Hs578T and plasmid DNA from *E.coli* producing recombinant asporin; Supplementary Figure S6) was confirmed by Sanger sequencing. PCR was performed with the same primers (without FAM tag in the forward primer) using HotStarTaq Master Mix (Quiagen, Hilden, Germany). Products were purified by QIAquick PCR Purification kit (Qiagen) and used as a template for sequencing PCR which was performed using BigDye Terminator

171

v1.1 Cycle Sequencing kit (Applied Biosystems). PCR products of this second PCR were purified by BigDye XTerminator™ Purification Kit (Applied Biosystems) and then separated by capillary electrophoresis on ABI PRISM® 3100 Genetic Analyzer (Applied Biosystems). Obtained data were analysed using Sequencing Analysis Software™ and visualised with GeneMapper® Software (Applied Biosystems). Whole procedure was performed according to the manufacturer protocol.

CONFLICTS OF INTEREST

The authors have no conflicts of interest to declare.

REFERENCES

1. Kharaishvili G, Simkova D, Bouchalova K, Gachechiladze M, Narsia N, Bouchal J. The role of cancer-associated fibroblasts, solid stress and other microenvironmental factors in tumor progression and therapy resistance. Cancer Cell Int. 2014; 14: 41. doi: 10.1186/1475-2867-14-41.

2. Bosman FT, Stamenkovic I. Functional structure and composition of the extracellular matrix. J Pathol. 2003; 200: 423-8. doi: 10.1002/path.1437.

3. Jean C, Gravelle P, Fournie JJ, Laurent G. Influence of stress on extracellular matrix and integrin biology. Oncogene. 2011; 30: 2697-706. doi: 10.1038/onc.2011.27.

4. Mecham RP. Overview of extracellular matrix. Curr Protoc Cell Biol. 2001; Chapter 10: Unit 10.1. doi: 10.1002/0471143030.cb1001s00.

5. Schaefer L, Iozzo RV. Biological functions of the small leucine-rich proteoglycans: from genetics to signal transduction. J Biol Chem. 2008; 283: 21305-9. doi: 10.1074/jbc.R800020200.

6. Henry SP, Takanosu M, Boyd TC, Mayne PM, Eberspaecher H, Zhou W, de Crombrugghe B, Hook M, Mayne R. Expression pattern and gene characterization of asporin, a newly discovered member of the leucine-rich repeat protein family. J Biol Chem. 2001; 276: 12212-21. doi: 10.1074/jbc.M011290200.

7. Lorenzo P, Aspberg A, Onnerfjord P, Bayliss MT, Neame PJ, Heinegard D. Identification and characterization of asporin, a novel member of the leucine-rich repeat protein family closely related to decorin and biglycan. J Biol Chem. 2001; 276: 12201-11. doi: 10.1074/jbc.M010932200.

8. Yamada S, Murakami S, Matoba R, Ozawa Y, Yokokoji T, Nakahira Y, Ikezawa K, Takayama S, Matsubara K, Okada H. Expression profile of active genes in human periodontal ligament and isolation of PLAP-1, a novel SLRP family gene. Gene. 2001; 275: 279-86.

9. Kou I, Nakajima M, Ikegawa S. Binding characteristics of the osteoarthritis-associated protein asporin. J Bone Miner Metab. 2010; 28: 395-402. doi: 10.1007/s00774-009-0145-8.

10. Kalamajski S, Aspberg A, Lindblom K, Heinegård D, Oldberg A. Asporin competes with decorin for collagen binding, binds calcium and promotes osteoblast collagen mineralization. Biochem J. 2009; 423: 53-9. doi: 10.1042/BJ20090542.

11. Ikegawa S. The genetics of common degenerative skeletal disorders: osteoarthritis and degenerative disc disease. Annu Rev Genomics Hum Genet. 2013; 14: 245-56. doi: 10.1146/annurev-genom-091212-153427.

12. Turashvili G, Bouchal J, Baumforth K, Wei W, Dziechciarkova M, Ehrmann J, Klein J, Fridman E, Skarda J, Srovnal J, Hajduch M, Murray P, Kolar Z. Novel markers for differentiation of lobular and ductal invasive breast carcinomas by laser microdissection and microarray analysis. BMC Cancer. 2007; 7: 55. doi: 10.1186/1471-2407-7-55.

13. Kharaishvili G, Cizkova M, Bouchalova K, Mgebrishvili G, Kolar Z, Bouchal J. Collagen triple helix repeat containing 1 protein, periostin and versican in primary and metastatic breast cancer: an immunohistochemical study. J Clin Pathol. 2011; 64: 977-82. doi: 10.1136/jclinpath-2011-200106.

14. Malanchi I, Santamaria-Martínez A, Susanto E, Peng H, Lehr HA, Delaloye JF, Huelsken J. Interactions between cancer stem cells and their niche govern metastatic colonization. Nature. 2012; 481: 85-9. doi: 10.1038/nature10694.

15. Ma XJ, Wang Z, Ryan PD, Isakoff SJ, Barmettler A, Fuller A, Muir B, Mohapatra G, Salunga R, Tuggle JT, Tran Y,

Tran D, Tassin A, et al. A two-gene expression ratio predicts clinical outcome in breast cancer patients treated with tamoxifen. Cancer Cell. 2004; 5: 607-16. doi: 10.1016/j.ccr.2004.05.015.

16. Best CJ, Gillespie JW, Yi Y, Chandramouli GV, Perlmutter MA, Gathright Y, Erickson HS, Georgevich L, Tangrea MA, Duray PH, González S, Velasco A, Linehan WM, et al. Molecular alterations in primary prostate cancer after androgen ablation therapy. Clin Cancer Res. 2005; 11: 6823-34. doi: 10.1158/1078-0432.CCR-05-0585.

17. Stanbrough M, Bubley GJ, Ross K, Golub TR, Rubin MA, Penning TM, Febbo PG, Balk SP. Increased expression of genes converting adrenal androgens to testosterone in androgen-independent prostate cancer. Cancer Res. 2006; 66: 2815-25. doi: 10.1158/0008-5472.CAN-05-4000.

18. Mackay A, Urruticoechea A, Dixon JM, Dexter T, Fenwick K, Ashworth A, Drury S, Larionov A, Young O, White S, Miller WR, Evans DB, Dowsett M. Molecular response to aromatase inhibitor treatment in primary breast cancer. Breast Cancer Res. 2007; 9: R37. doi: 10.1186/bcr1732.

19. Rajkumar T, Vijayalakshmi N, Gopal G, Sabitha K, Shirley S, Raja UM, Ramakrishnan SA. Identification and validation of genes involved in gastric tumorigenesis. Cancer Cell Int. 2010; 10: 45. doi: 10.1186/1475-2867-10-45.

20. Cima I, Schiess R, Wild P, Kaelin M, Schüffler P, Lange V, Picotti P, Ossola R, Templeton A, Schubert O, Fuchs T, Leippold T, Wyler S, et al. Cancer genetics-guided discovery of serum biomarker signatures for diagnosis and prognosis of prostate cancer. Proc Natl Acad Sci U S A. 2011; 108: 3342-7. doi: 10.1073/pnas.1013699108.

21. Turtoi A, Musmeci D, Wang Y, Dumont B, Somja J, Bevilacqua G, De Pauw E, Delvenne P, Castronovo V. Identification of novel accessible proteins bearing diagnostic and therapeutic potential in human pancreatic ductal adenocarcinoma. J Proteome Res. 2011; 10: 4302-13. doi: 10.1021/pr200527z.

22. Klee EW, Bondar OP, Goodmanson MK, Dyer RB, Erdogan S, Bergstralh EJ, Bergen HR, Sebo TJ, Klee GG. Candidate serum biomarkers for prostate adenocarcinoma identified by mRNA differences in prostate tissue and verified with protein measurements in tissue and blood. Clin Chem. 2012; 58: 599-609. doi: 10.1373/clinchem.2011.171637.

23. Orr B, Riddick AC, Stewart GD, Anderson RA, Franco OE, Hayward SW, Thomson AA. Identification of stromally expressed molecules in the prostate by tag-profiling of cancer-associated fibroblasts, normal fibroblasts and fetal prostate. Oncogene. 2012; 31: 1130-42. doi: 10.1038/onc.2011.312.

24. Satoyoshi R, Kuriyama S, Aiba N, Yashiro M, Tanaka M. Asporin activates coordinated invasion of scirrhous gastric cancer and cancer-associated fibroblasts. Oncogene 2014; 34: 650-60. doi: 10.1038/onc.2013.584.

25. Maris P, Blomme A, Palacios AP, Costanza B, Bellahcène A, Bianchi E, Gofflot S, Drion P, Trombino GE, Di Valentin E, Cusumano PG, Maweja S, Jerusalem G, et al. Asporin Is a Fibroblast-Derived TGF-β1 Inhibitor and a Tumor Suppressor Associated with Good Prognosis in Breast Cancer. PLoS Med. 2015; 12: e1001871. doi: 10.1371/journal.pmed.1001871.

26. Lee EH, Park HJ, Jeong JH, Kim YJ, Cha DW, Kwon DK, Lee SH, Cho JY. The role of asporin in mineralization of human dental pulp stem cells J Cell Physiol. 2011; 226: 1676-82. doi: 10.1002/jcp.22498.

27. Kou I, Nakajima M, Ikegawa S. Expression and regulation of the osteoarthritis-associated protein asporin. J Biol Chem. 2007; 282: 32193-9. doi: 10.1074/jbc.M706262200.

28. Duval E, Bigot N, Hervieu M, Kou I, Leclercq S, Galéra P, Boumediene K, Baugé C. Asporin expression is highly regulated in human chondrocytes. Mol Med. 2011; 17: 816-23. doi: 10.2119/molmed.2011.00052.

29. Yamada S, Ozawa Y, Tomoeda M, Matoba R, Matsubara K, Murakami S. Regulation of PLAP-1 expression in periodontal ligament cells. J Dent Res. 2006; 85: 447-51.

30. Sodek KL, Brown TJ, Ringuette MJ. Collagen I but not Matrigel matrices provide an MMP-dependent barrier to ovarian cancer cell penetration. BMC Cancer. 2008; 8: 223. doi: 10.1186/1471-2407-8-223.

31. Prat A, Karginova O, Parker JS, Fan C, He X, Bixby L, Harrell JC, Roman E, Adamo B, Troester M, Perou CM. Characterization of cell lines derived from breast cancers and normal mammary tissues for the study of the intrinsic molecular subtypes. Breast Cancer Res Treat. 2013; 142: 237-55. doi: 10.1007/s10549-013-2743-3.

32. Hurley PJ, Sundi D, Shinder B, Simons BW, Hughes RM, Miller RM, Benzon B, Faraj SF, Netto GJ, Vergara IA, Erho N, Davicioni E, Karnes RJ, et al. Germline Variants in Asporin Vary by Race, Modulate the Tumor Microenvironment, and Are Differentially Associated with Metastatic Prostate Cancer. Clin Cancer Res. 2016; 22: 448-58. doi: 10.1158/1078-0432.CCR-15-0256.

33. Correia AL, Bissell MJ. The tumor microenvironment is a dominant force in multidrug resistance. Drug Resist Updat. 2012; 15: 39-49. doi: 10.1016/j.drup.2012.01.006.

34. Rajan N, Elliott R, Clewes O, Mackay A, Reis-Filho JS, Burn J, Langtry J, Sieber-Blum M, Lord CJ, Ashworth A. Dysregulated TRK signalling is a therapeutic target in CYLD defective tumours. Oncogene. 2011; 30: 4243-60. doi: 10.1038/onc.2011.133.

35. Stylianopoulos T, Martin JD, Chauhan VP, Jain SR, Diop-Frimpong B, Bardeesy N, Smith BL, Ferrone CR, Hornicek FJ, Boucher Y, Munn LL, Jain RK. Causes, consequences, and remedies for growth-induced solid stress in murine and human tumors. Proc Natl Acad Sci U S A. 2012; 109: 15101-8. doi: 10.1073/pnas.1213353109.

36. Kretschmer C, Sterner-Kock A, Siedentopf F, Schoenegg W, Schlag PM, Kemmner W. Identification of early molecular markers for breast cancer. Mol Cancer. 2011; 10: 15. doi: 10.1186/1476-4598-10-15.

173

37. Gao H, Chakraborty G, Lee-Lim AP, Mo Q, Decker M, Vonica A, Shen R, Brogi E, Brivanlou AH, Giancotti FG. The BMP inhibitor Coco reactivates breast cancer cells at lung metastatic sites. Cell. 2012; 150: 764-79. doi: 10.1016/j.cell.2012.06.035.

38. Wan L, Pantel K, Kang Y. Tumor metastasis: moving new biological insights into the clinic. Nat Med. 2013; 19: 1450-64. doi: 10.1038/nm.3391.

39. Levin VA, Panchabhai SC, Shen L, Kornblau SM, Qiu Y, Baggerly KA. Different changes in protein and phosphoprotein levels result from serum starvation of high-grade glioma and adenocarcinoma cell lines. J Proteome Res. 2010; 9: 179-91. doi: 10.1021/pr900392b.

40. Lee AS. Glucose-regulated proteins in cancer: molecular mechanisms and therapeutic potential. Nat Rev Cancer. 2014; 14: 263-76. doi: 10.1038/nrc3701.

41. Leccia F, Nardone A, Corvigno S, Vecchio LD, De Placido S, Salvatore F, Veneziani BM. Cytometric and biochemical characterization of human breast cancer cells reveals heterogeneous myoepithelial phenotypes. Cytometry A. 2012; 81: 960-72. doi: 10.1002/cyto.a.22095.

42. Ding Q, Zhang M, Liu C. Asporin participates in gastric cancer cell growth and migration by influencing EGF receptor signaling. Oncol Rep. 2015; 33: 1783-90. doi: 10.3892/or.2015.3791.

43. Wolf DM, Lenburg ME, Yau C, Boudreau A, van 't Veer LJ. Gene co-expression modules as clinically relevant hallmarks of breast cancer diversity. PLoS One. 2014; 9: e88309. doi: 10.1371/journal.pone.0088309.

44. Klein A, Olendrowitz C, Schmutzler R, Hampl J, Schlag PM, Maass N, Arnold N, Wessel R, Ramser J, Meindl A, Scherneck S, Seitz S. Identification of brain- and bone-specific breast cancer metastasis genes. Cancer Lett 2009; 276: 212-20. doi: 10.1016/j.canlet.2008.11.017.

45. Dangaria SJ, Ito Y, Luan X, Diekwisch TG. Differentiation of neural-crest-derived intermediate pluripotent progenitors into committed periodontal populations involves unique molecular signature changes, cohort shifts, and epigenetic modifications. Stem Cells Dev. 2011; 20: 39-52. doi: 10.1089/scd.2010.0180.

46. Kizawa H, Kou I, Iida A, Sudo A, Miyamoto Y, Fukuda A, Mabuchi A, Kotani A, Kawakami A, Yamamoto S, Uchida A, Nakamura K, Notoya K, et al. An aspartic acid repeat polymorphism in asporin inhibits chondrogenesis and increases susceptibility to osteoarthritis. Nat Genet. 2005; 37: 138-44. doi: 10.1038/ng1496.

47. Giannoni E, Bianchini F, Masieri L, Serni S, Torre E, Calorini L, Chiarugi P. Reciprocal activation of prostate cancer cells and cancer-associated fibroblasts stimulates epithelial-mesenchymal transition and cancer stemness. Cancer Res. 2010; 70: 6945-56. doi: 10.1158/0008-5472.

48. Giannoni E, Bianchini F, Calorini L, Chiarugi P. Cancer associated fibroblasts exploit reactive oxygen species through a proinflammatory signature leading to epithelial mesenchymal transition and stemness. Antioxid Redox Signal. 2011; 14: 2361-71. doi: 10.1089/ars.2010.3727.

49. Mokry J, Soukup T, Micuda S, Karbanova J, Visek B, Brcakova E, Suchanek J, Bouchal J, Vokurkova D, Ivancakova R. Telomere attrition occurs during ex vivo expansion of human dental pulp stem cells. J Biomed Biotechnol. 2010; 2010: 673513. doi: 10.1155/2010/673513.

50. Levental KR, Yu H, Kass L, Lakins JN, Egeblad M, Erler JT, Fong SF, Csiszar K, Giaccia A, Weninger W, Yamauchi M, Gasser DL, Weaver VM. Matrix crosslinking forces tumor progression by enhancing integrin signaling. Cell. 2009; 139: 891-906. doi: 10.1016/j.cell.2009.10.027.

51. Friedrich J, Seidel C, Ebner R, Kunz-Schughart LA. Spheroid-based drug screen: considerations and practical approach. Nat Protoc. 2009; 4: 309-24. doi: 10.1038/nprot.2008.226.

52. Li Q, Chen C, Kapadia A, Zhou Q, Harper MK, Schaack J, LaBarbera DV. 3D models of epithelial-mesenchymal transition in breast cancer metastasis: high-throughput screening assay development, validation, and pilot screen. J Biomol Screen. 2011; 16: 141-54. doi: 10.1177/1087057110392995.

53. Andersson AC, Strömberg S, Bäckvall H, Kampf C, Uhlen M, Wester K, Pontén F. Analysis of protein expression in cell microarrays: a tool for antibody-based proteomics. J Histochem Cytochem. 2006; 54: 1413-23. doi: 10.1369/jhc.6A7001.2006.

54. Wang F, Flanagan J, Su N, Wang LC, Bui S, Nielson A, Wu X, Vo HT, Ma XJ, Luo Y. RNAscope: a novel in situ RNA analysis platform for formalin-fixed, paraffin-embedded tissues. J Mol Diagn. 2012; 14: 22-9. doi: 10.1016/j.jmoldx.2011.08.002.

55. Dreyer JH, Hauck F, Oliveira-Silva M, Barros MH, Niedobitek G. Detection of HPV infection in head and neck squamous cell carcinoma: a practical proposal. Virchows Arch. 2013; 462: 381-9. doi: 10.1007/s00428-013-1393-5.

56. MacLean B, Tomazela DM, Shulman N, Chambers M, Finney GL, Frewen B, Kern R, Tabb DL, Liebler DC, MacCoss MJ. Skyline: an open source document editor for creating and analyzing targeted proteomics experiments. Bioinformatics. 2010; 26: 966-8. doi: 10.1093/bioinformatics/btq054.

57. Pfaffl MW. A new mathematical model for relative quantification in real-time RT-PCR. Nucleic Acids Res. 2001; 29: e45. doi: 10.1093/nar/29.9.e45

58. Györffy B, Lanczky A, Eklund AC, Denkert C, Budczies J, Li Q, Szallasi Z. An online survival analysis tool to rapidly assess the effect of 22,277 genes on breast cancer prognosis using microarray data of 1,809 patients. Breast Cancer Res Treat. 2010; 123: 725-31. doi: 10.1007/s10549-009-0674-9.

174

The dual role of asporin in breast cancer progression
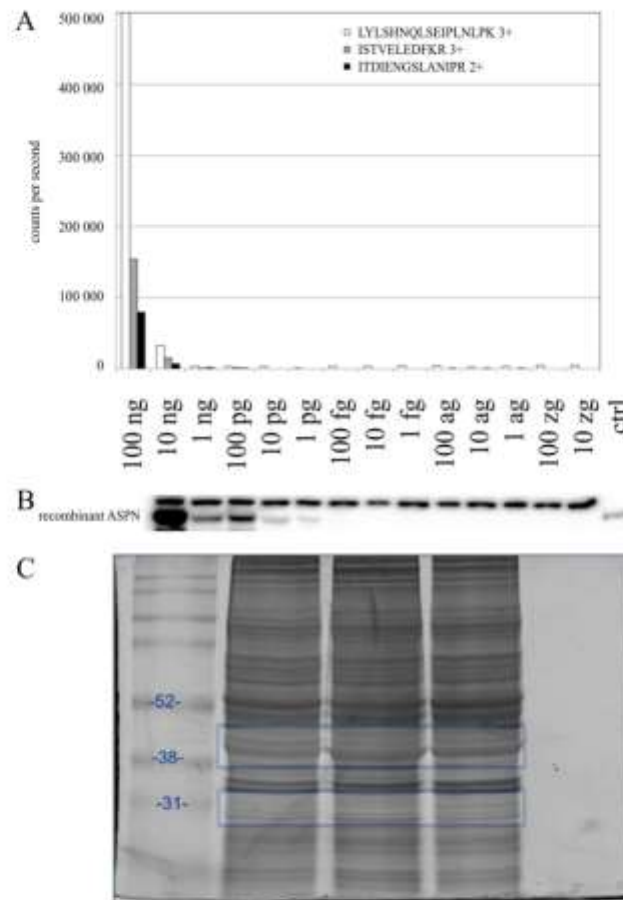
SUPPLEMENTARY FIGURES AND TABLES
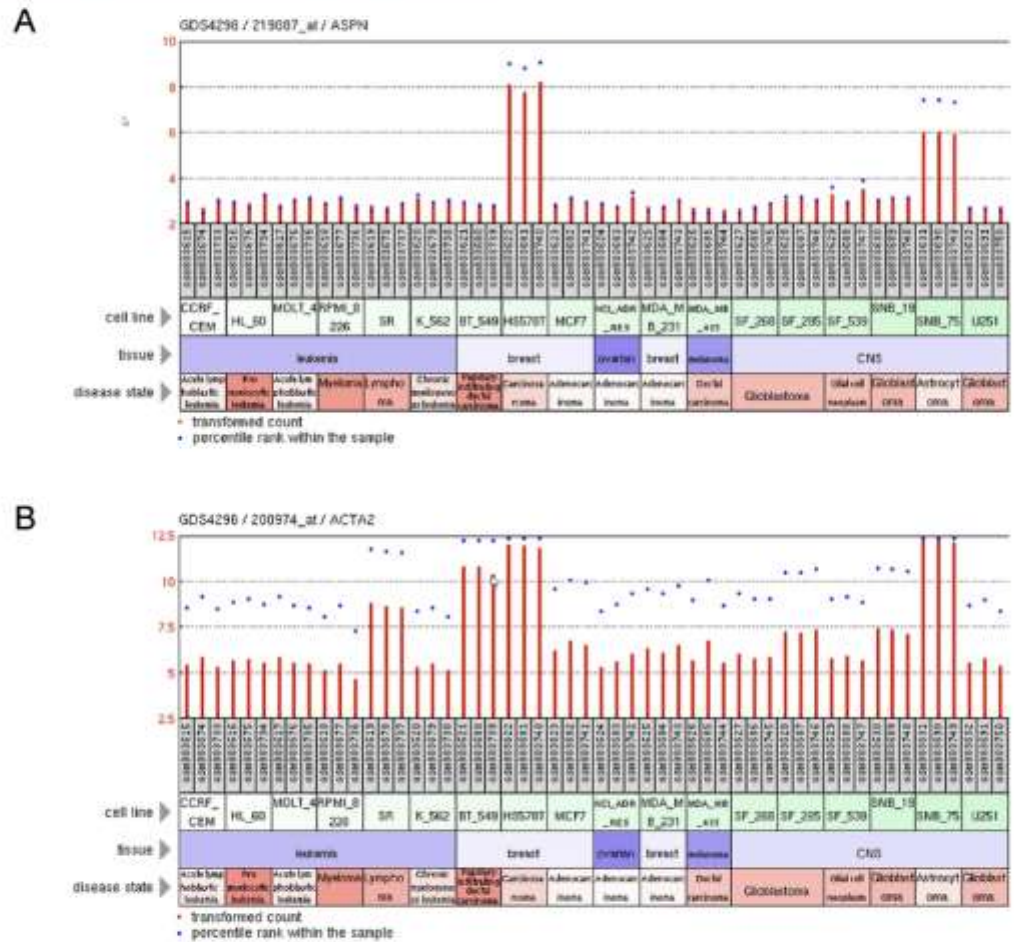
Supplementary Figure S1: Microarray analysis in NCI-60 cancer cell line panel (GDS4296 in GEO Database at the National Center for Biotechnology Information). A. Histogram with two positive cell lines is shown while all other cell lines were negative (full histogram can be easily found in the GEO Database). Hs578T breast cancer cell line has higher signal also in GSE5720, GSE15026 and GSE24717. Other cell lines with asporin mRNA expression can be found in Cancer Cell Line Encyclopedia (Supplementary Table S4). **B.** Expression of smooth muscle actin (gene name ACTA2) is highest in the asporin positive cell lines Hs578T and SNB-75.
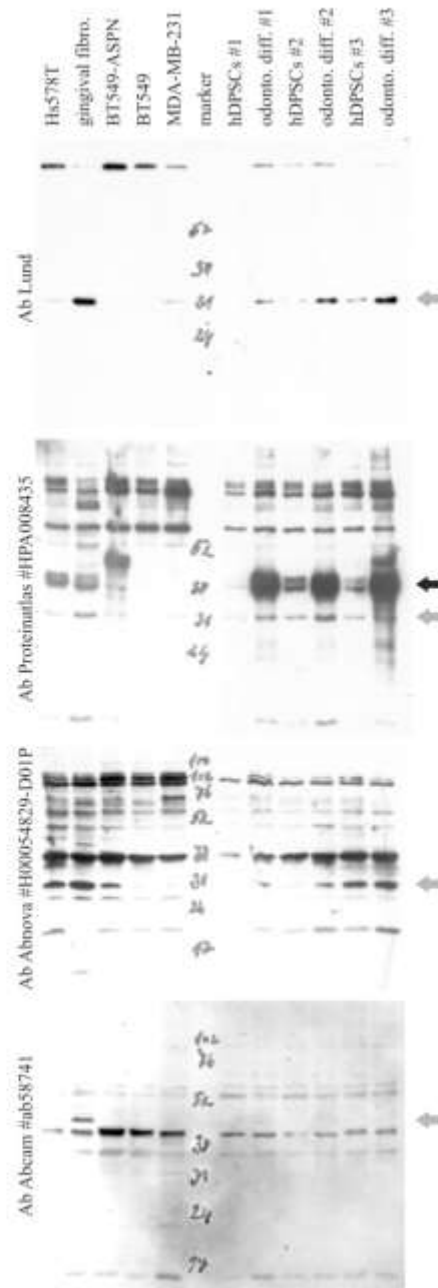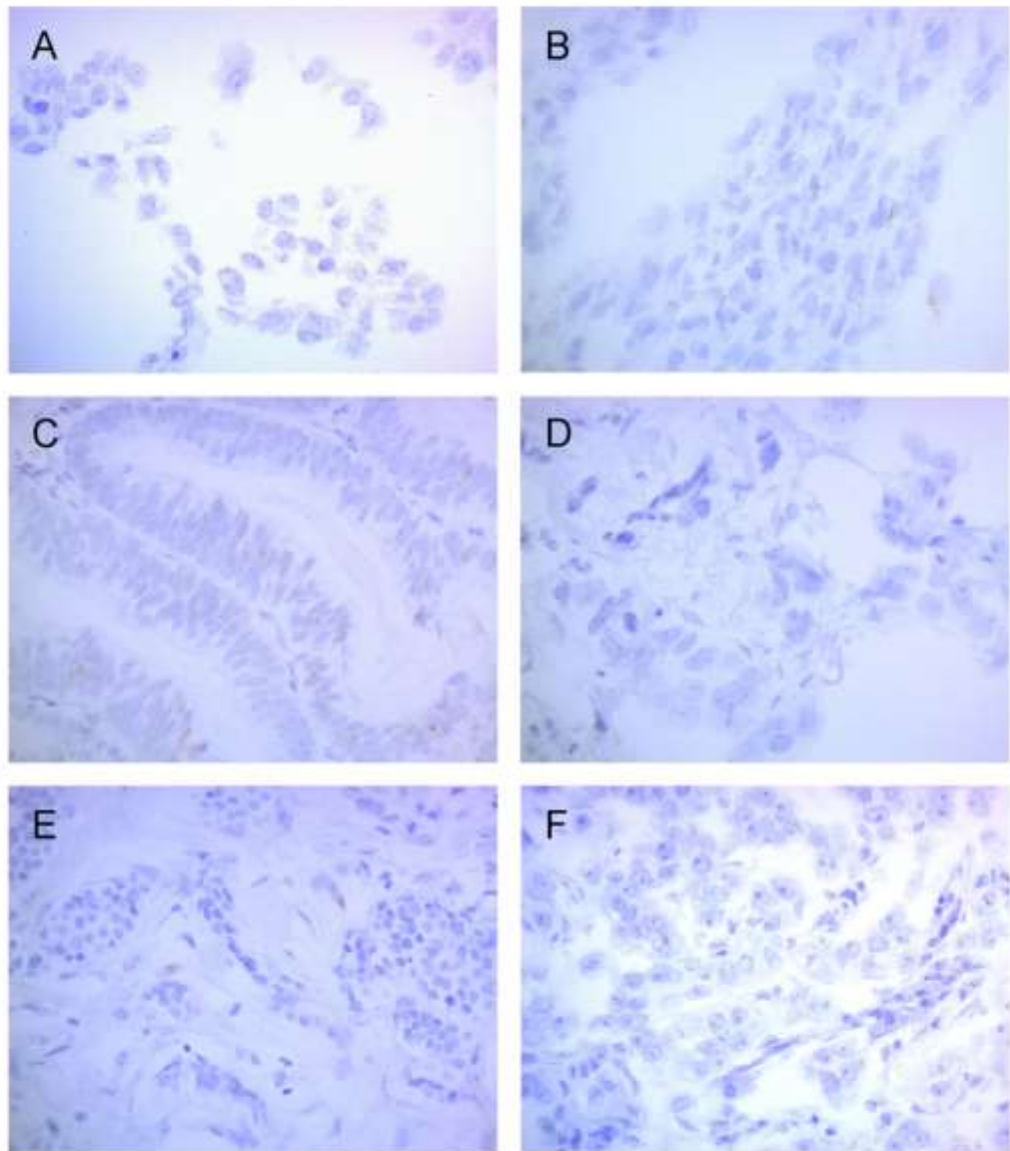
Supplementary Figure S2: Validation of the indicated antibodies (antibody Lund was kindly provided by prof. Oldberg; see also Supplementary Table S2). Grey arrows indicate bands which partially correspond with RT-PCR results. Black arrow is commented in the main Results section.

Supplementary Figure S3: Mass spectrometric analysis of asporin. Dilution series was prepared by spiking recombinant asporin in indicated amount into 10 ul lysate of HeLa cells. Mass spectrometric analysis detected a sample with 10 ng recombinant asporin **A.,** while western blotting with the Proteinatlas antibody #HPA008435 detected a sample with 1 pg of asporin **B.** The sample with 100 ng was not analysed by western blotting due to expected very high signal. One ng of recombinant asporin (without HeLa lysate) was loaded as a control directly for electrophoresis (ctrl; molecular weight of the recombinant asporin is 70 kDa due to the fusion with glutathion S-transferase). Experiment was repeated three times, representative image is shown. Supplementary Figure S3 **C.** displays polyacrylamide gel with loading of 200 μg protein of hDPSCs after odontogenic differentiation (three independent samples). The gel was cut at the indicated molecular weights and analysed by mass spectrometry (see Supplementary Table S3). A similar gel was prepared for BT549-ASPN.

Supplementary Figure S4: Expression of asporin by RNA sequencing in different human tissues (Expression Atlas of ArrayExpress database at the European Bioinformatics Institute). The highest expression was detected in endometrium, gall bladder, heart **A.** and breast (**B.,** Illumina Body Map).

Supplementary Figure S5: Expression of asporin by RNA scope *in situ* hybridization. The same samples as in the Figure 4 are shown at lower magnification (400x).
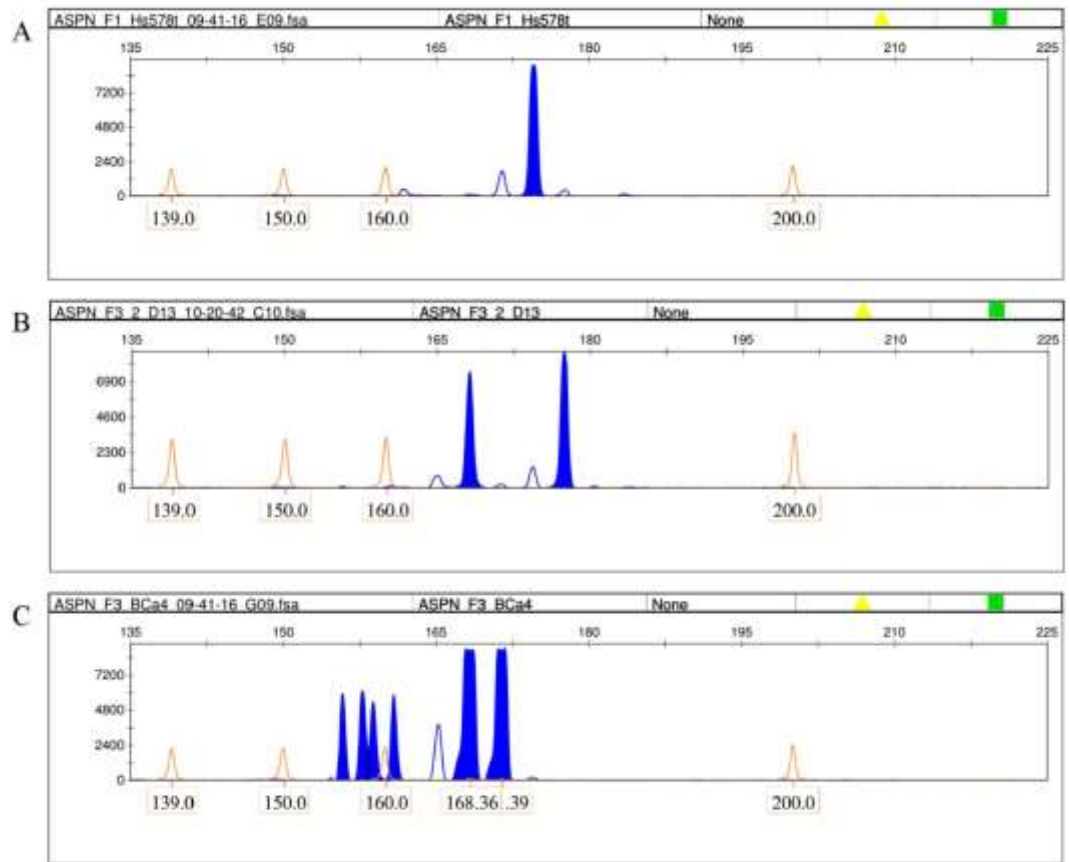
Supplementary Figure S6: Genotyping of asporin D-repeat polymorphism. Hs578T cells contain D15 allele **A.**, while breast cancer associated fibroblasts have D13/D14 genotype **C.** Besides expected D13 allele, the asporin producing *E.coli* contained also D16 allele **B.**, which was confirmed by Sanger sequencing (along with Hs578T cells and CAFs). Axes Y and X display fluorescent intensity and capillary electrophoresis mobility, respectively (please see Material and Methods for further details).

Supplementary Table S2: Overview of asporin antibodies

Supplier	producer	host	cat.no.	immunogen	tested in this study	published article [1]
Sigma-Aldrich	proteinatlas.org	rabbit	HPA008435	IPLNLPKSLAELRIHENKVKKIQKDTFKGMNALHV LEMSANPLDNNGIEPGAFEGVTVFHIRIAEAKLTS VPKGLPPTLLELHLDYNKISTVELEDFKRYKELQRL GLGNNKITDIENGSLAN	Figure 1, Suppl. Figure 2, IHC not shown	Satoyoshi et al. 2014, Maris et al. 2015, Hurley et al. 2016
Abcam/ GenWay	not specified	rabbit	ab58741/ GWB-6BA1C3	ISTVELEDFKRYKELQRLGLGNNKITDIENGSLANIP RVREIHLENNKL	Suppl. Figure 2, IHC not shown	Gruber et al. 2009 (GenWay), Lee et al. 2011 (Abcam catalog # not specified)
Prof. Oldberg	Lund group	rabbit	in-house	expressed from cDNA	Suppl. Figure 2, IHC not shown	
Abcam/ Everest-Biotech	not specified	goat	ab31303/ EB06670	IHENKVKKIQKDT	data not shown	Ho et al. 2010, Hurng et al. 2011
Abnova	Abnova	rabbit	H00054829-D01P	full recombinant protein	Suppl. Figure 2	
our antibody	Dr. Vojtesek	rabbit	in-house	TVELEDFKRYKELQR	data not shown	
our antibody	Dr. Vojtesek	rabbit	in-house	HLENNKLKKIPSGLPE	data not shown	
not available	Yamada group	rabbit	in-house	EPRSHFFPFD	not tested	Yamada et al. 2006
Abcam	not specified	rabbit	ab154404 and ab75317	YGLILNNNKLTKIHPKAFLTTKKLRRLYLSHNQLSEIP LNLPKSLAELRIHENKVKKIQKDTFKGMNALHVLEM SANPLDNNGIEPGAFEGVT and synthetic peptide derived from the C-terminal domain, respectively	not tested	
Sigma-Aldrich	proteinatlas.org	rabbit	HPA024230	EYVLLLFLALCSAKPFFSPSHIALKNMMLKDMEDTDDD DDDDDDDDDDDEDNSLFPTREPRSHFFPFDLFPMCPF GCQCYSRVVHCSDLGLTSVPTNIPFDTRMLDLQNNKI KEIKENDFKGLTSLYGLILNNNKLTKIHPKA	not tested	

[1] References not included in the main article: Gruber et al. Asporin, a susceptibility gene in osteoarthritis, is expressed at higher levels in the more degenerate human intervertebral disc. Arthritis Research & Therapy 2009, 11: R47; Ho et al. The biomechanical characteristics of the bone-periodontal ligament-cementum complex. Biomaterials 2010; 31:6635-46; Hurng et al. Discontinuities in the human bone-PDL-cementum complex. Biomaterials 2011; 32:7106-17. IHC, immunohistochemistry.

Supplementary Table S3: Mass-spectometric intensities (counts per second) of asporin peptides in samples cut from a gel with loading of 200 µg protein (Figure S3C)

sample	Pep1	Pep2	Pep3
Odonto.diff#1 40 kDa	3000	1500	3000
Odonto.diff#1 30 kDa	0	0	0
Odonto.diff#2 40 kDa	2000	1500	2000
Odonto.diff#2 30 kDa	0	0	0
Odonto.diff#3 40 kDa	3000	1000	2000
Odonto.diff#3 30 kDa	0	0	0
BT549-ASPN 50 kDa	2000	2000	2500
BT549-ASPN 40 kDa	0	0	0

Pep1, thrice charged LYLSHNQLSEIPLNLPK; Pep2, twice charged ISTVELEDFKR; Pep3, thrice charged ISTVELEDFKR. Specific transitions were validated by the analysis of the recombinant asporin.