Novel (phospho-)proteomics approaches to study aging and senescence

Erik L. de Graaf

The research in this thesis was performed in the Biomolecular Mass Spectrometry & Proteomics Group, Utrecht University, Utrecht, The Netherlands

The research was financially supported by the Centre for Biomedical Genetics and the Netherlands Proteomics Centre.

Printed by: Proefschriftmaken.nl Uitgeverij BOXPress

ISBN: 978-90-8891-804-9

Novel (phospho-)proteomics approaches to study aging and senescence

Vernieuwende (fosfo-)proteomics methoden voor de studie van veroudering en senescence.

(met een samenvatting in het Nederlands)

Proefschrift

ter verkrijging van de graad van doctor aan de Universiteit Utrecht op gezag van de rector magnificus, prof.dr. G.J. van der Zwaan, ingevolge het besluit van het college voor promoties in het openbaar te verdedigen op woensdag 12 februari 2014 des middags te 2.30 uur

door

Erik Leonardus de Graaf

geboren op 3 december 1984, te Bernheze

Promotor	Prof. dr. A.J.R. Heck
Co-promotor	Dr. A.F.M Altelaar

Table of Contents

Chapter 1

General Introduction

1. Introduction to Proteomics

A protein is a large biological molecule that consists of a chain of amino acid building blocks that is folded into a specific three-dimensional structure. The amino acid sequence of each protein is encoded in the DNA and its' coding region is called a gene. Genes are transcribed into mRNA molecules which in turn are translated into the final product; a protein. Proteins are very important as they are the key effector directly involved in virtually all cellular processes, ranging from cell growth and homeostasis to cell death. As a result, perturbations in protein levels, sequence and modification states are often linked to many types of diseases.

The large-scale study of genes (i.e. genomics) is mainly a qualitative study used to identify genome structures of different organisms focussed at coding and regulatory DNA sequences. Over the past decades genomics formed the fundamental basis for protein research, studying differences in protein blueprints between species, individuals and disease phenotypes.

Transcriptomics is a more quantitative field of research in which the amount of gene transcription (in mRNA levels) is determined, yielding important biological information about what genes are turned on, and when. The applied techniques are relatively straightforward and a single experiment is often enough to generate a high-throughput and comprehensive dataset. (1) The disadvantage however is that, although the measurements reflect the genome's plans for protein synthesis, it does not yet represent the realization of those plans. The correlation between mRNA and protein levels is poor and generally below 0.5 (2). This limited predictive value is a result of additional biological layers of regulation including mRNA splicing and editing, mRNA translation regulation and variable mRNA transcript or protein half-lives (3). Moreover, the activity of a protein is not simply defined by the amount of protein but more often by its' localization or the multiple types of modifications that occur after production, termed posttranslational modifications (PTMs). PTMs are particularly important in the propagation of cellular signals, where the attachment of a phosphate group to a protein triggers either activation or inactivation of a signalling pathway. For instance, when studying the origin of uncontrolled cell growth in cancer, it can be very useful to know what kind of PTMs occur on growth associated signalling proteins.

The direct analysis of protein expression and modification state was traditionally performed on a small scale, using immunoassays, two-hybrid assays or gel-based methods, limiting the analysis throughput to one or a few proteins. The large-scale study of (thousands) of proteins in a single experiment is referred to as proteomics, in analogy to genomics (4, 5). Proteomics techniques are needed as more often than not multiple proteins are involved in a single cellular process. The study of multiple proteins simultaneously allows for a more "systems biology" type of approach mapping protein content, crosstalk and activation to a certain point of time for a certain type of condition. As an example, proteomics can be used to screen for proteins that are differentially expressed in healthy or diseased tissues for better diagnosis and therapy monitoring (6) or to identify drug targets (7) to allow for more effective and selective drug design, minimizing side-effects.

Proteome analysis can be quite challenging as the ~20 000 genes studied in genomics, on average result in double the amount of transcript variants (8) that can be transcribed into proteins that become posttranslational modified in different ways, resulting in a total of at least 100 000 protein isoforms. On top of the vast amount of different protein isoforms, protein localization, interaction and abundance (that can vary over six orders of magnitude) are crucial parameters that further increase the complexity of protein analysis. The increasing sample complexity poses serious challenges on analytical methods to accurately determine the cellular protein contents, expression levels and PTM states. As a result of this there is no single one-method-fits-all workflow and therefore a myriad of different, mainly mass spectrometry-based, approaches has been developed over the past decades. (9-11) The analytical methods can be focussed on directly measuring intact proteins or protein-derived peptides referred to as top-down and bottom-up proteomics, respectively. In the following sections the fundamentals of the most frequently used bottom-up mass spectrometry-based proteomic analyses are described.

2. Mass spectrometry-based proteomics

A typical proteomics workflow consists of three essential parts that can be extended with additional methods to fine-tune the experiment to the investigators' need. The basic proteomics workflow consists of sample preparation, LC-MS and data analysis as illustrated in Figure 1 and each part will be explained in more detail in the following paragraphs.

2.1 Sample preparation

Different types of samples can be analysed ranging from synthetic protein mixtures, biochemically enriched organelle fractions or whole cell extracts from in vitro cultures, to heterogeneous in vivo tissue samples from varying organisms. Generally for bottom-up proteomics, the first steps involve the solubilisation and denaturation of proteins in a lysis buffer followed by a digestion of all proteins into peptides, using proteolytic enzymes such as Trypsin and Lys-C. Trypsin is most often used due to its specific cleavage at the C-terminus of arginine and lysine (12), generating peptides of similar length (mass) and charge that are readily identifiable by the mass spectrometer. Another reason why peptide analysis is preferred over proteins is the higher resolution of liquid chromatography (LC)-based separation of peptides when compared to proteins. The application of LC in proteomics is described in the next paragraph.

2.2 Sample complexity reduction: separation, fractionation or enrichment

Depending on the biological system, proteolytic digestion of full proteomes can generate hundreds of thousands of peptides and such complexity is not directly amenable for MS analysis. The mass spectrometer can only sequence each peptide one by one and therefore

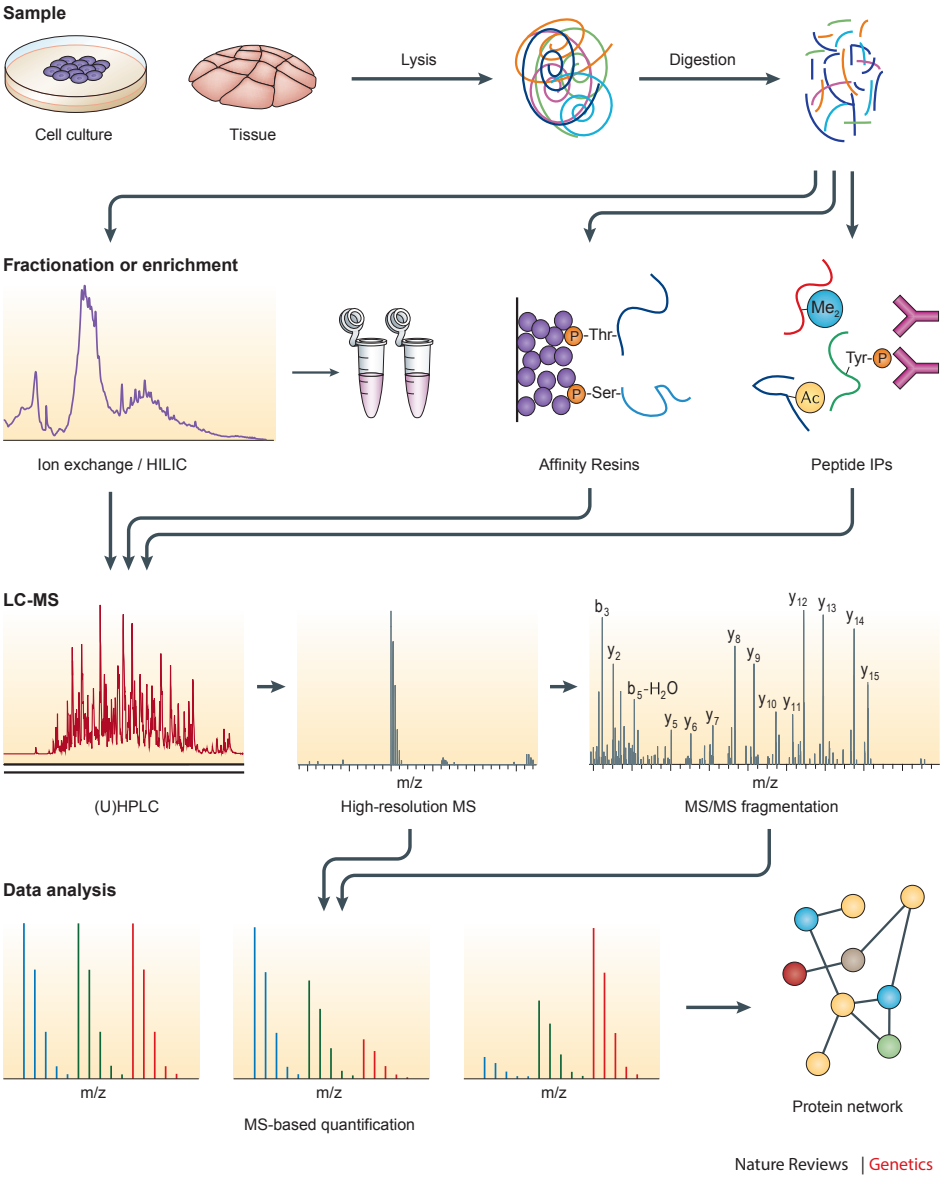

***Figure 1**: A typical proteomics workflow is flexible and needs to be adjusted according to the research question. (Adopted from Altelaar et al.; Ref. 11)*

they need to be separated using LC in order to be presented to the MS one after the other. Liquid chromatography is used to separate peptides based on their physicochemical properties, which in turn is determined by the peptide amino acid sequence. The most often used stationary phases include reversed phase (RP), and strong cation exchange (SCX), that separate peptides based on hydrophobicity and the number of positive charges, respectively. Due to its' low resolving power and high orthogonality to RP, SCX is often used

as a first step to prefractionate the sample. When used 'offline', this fractionation typically results in the conversion of a single highly complex sample into multiple less complex peptide mixtures, each ready to be injected into a second online RP-LC-MS setup. The latter setup takes advantage of the superior chromatographic resolving power of reverse phase material. Furthermore, the eluent used in RP is compatibility with electrospray peptide ionisation (ESI), allowing it to be directly interfaced to the MS for instant peptide mass measurements (hence 'online'). However, quite often posttranslational modified peptides are low abundant and therefore masked by the ubiquitous non-modified peptides hampering PTM analysis. Therefore, for the study of PTMs or low abundant peptides, an alternative or additional enrichment step is often included into the proteomics workflow.

Enrichment is usually performed before, after or instead of prefractionation methods. Enrichment strategies are mainly focussed on affinity chromatography or antibody-based immuniprecipitation methods. For phosphopeptides, the most popular enrichment strategies are based on ionic interactions of the phosphate moiety with a metal ion such as Fe^{3+} (13), Ga^{3+} (14) and Ti^{4+} (15, 16). The general principle is to immobilize positively charged metal ions to form a stationary phase that specifically coordinate to negatively charged phosphopeptides under acidic conditions. In immobilized metal ion affinity chromatography (IMAC) the metal ion is chelated by an immobilized polydentate ligand to form multiple coordinated bonds. (17) In metal oxide affinity chromatography (MOAC) porous metal dioxide microspheres are used as a stationary phase (i.e. TiO_2 (15), ZrO_2 (18)). Recently, an optimized protocol for high specificity and reproducible phosphopeptide enrichment was introduced using Ti^{4+}-IMAC material (16, 19), that was applied and further characterized in chapter 4.

The analysis of tyrosine phosphorylation poses a greater challenge than serine and threonine phosphorylation due to its' even lower endogenous frequency of occurrence. Therefore, to selectively probe tyrosine phosphorylation levels, immunoprecipitation (IP) is the most frequently used method of choice. (19-21) Typically, antibodies are raised against a mixture of synthetic phosphotyrosine containing peptides flanked by 6 to 9 "random" amino acids. There are also some downsides to antibody enrichment. Due to the intrinsic specificity of an antibody to a stretch of a minimum of 6 amino acids, the antibody enrichment can be biased towards certain amino acid sequences. Another side effect of a smaller epitope than antibody binding pocket size is poor enrichment reproducibility. Furthermore, the sample amount needed for phosphotyrosine IPs is typically one order of magnitude higher than phosphoserine/-threonine enrichments due to the low sensitivity of IP enrichment methods and the low natural frequency of phosphotyrosine signalling in mammals.

2.3 Mass spectrometry & Peptide sequencing

The mass spectrometer (MS) detects the mass-to-charge ratio (m/z) of an analyte ion. The MS analysis consists of roughly three parts. Firstly, the analyte needs to be ionized in the ion source from where it is directed into the vacuum of the mass spectrometer. Secondly, the ions are guided, manipulated and separated by electric fields in the mass analysers.

Lastly, the discrete ion packages created in the analyser are detected by the mass spectrometer. The following paragraphs describe the different processes involved in more detail.

Ionization sources

In order to measure the weight and identity of peptides they need to be analysed as ions in a vacuum. Therefore, non-volatile hydrophobic peptides need to be converted into gas phase ions without degradation. The most frequently used soft ionization techniques in proteomics are matrix assisted laser desorption ionization (MALDI) (22) and electrospray ionization (ESI) (23). (9)

In MALDI a matrix (i.e. alpha-cyano and dihydroxybenzoic acid) is mixed with the analyte (peptides) to form analyte crystals upon solvent evaporation. To generate bare ions, laser pulses are used to transfer energy to the solid phase matrix crystals resulting in desorption of matrix and analyte ions (Figure 2A). After laser desorption, analytes become predominantly singly charged [M+H]+ species. (22, 24) The main disadvantages of MALDI are low laser shot-to-shot reproducibility, the restriction to offline seperation/fractionation methods only and sample preparation methods depending on pure crystal formation. Advantages of MALDI are low complexity spectra due to predominantly singly charged ions, a higher tolerance to detergents and salts and the possibility to reanalyse samples.

Contrary to MALDI, in ESI ions are produced from a liquid phase. ESI is driven by a high voltage (1-6kV) applied between an emitter and the mass spectrometer inlet. The charged spray forms a Taylor cone that forms small droplets that burst into smaller droplets due to coulombic interactions and further desolvation is aided by a heated transfer capillary (Figure 2B). (23, 25, 26) When all solvent is evaporated the analyte obtains the droplet charges and this typically leads to multiply charged ions depending on the length and type of amino acid sequence (for tryptic peptides generally $z \geq 2$). The main reason why ESI is used in the majority of proteomics experiments is its ability to directly interface the LC separation with MS for high-throughput analysis.(27) Additionally, the continuous flow of ions generated by ESI maximizes MS instrument analysis time when compared to the laser pulsed ion packets formed in MALDI. The development of micro- and nano-ESI (referring to flow rate regimes per minute) have greatly increased sensitivity.

As most often acidic sample/chromatography conditions are used, peptides gain a net positive charge and the resulting gas phase ions travel through the vacuum of the mass spectrometer guided by negative electrostatic potentials.

Mass analysers

To measure the mass of ions, different mass analysers have been developed in the past decades. Mass analysers separate ions based on their mass-to-charge ratio using different principles. Separation of ions in Quadrupole (Q) mass analysers is based on their stable trajectory in combined static and alternating electric fields; time-of-flight (TOF) analysers make use of ion velocities or flight times and ion trap (IT), orbitrap and ion cyclotron resonance (ICR)

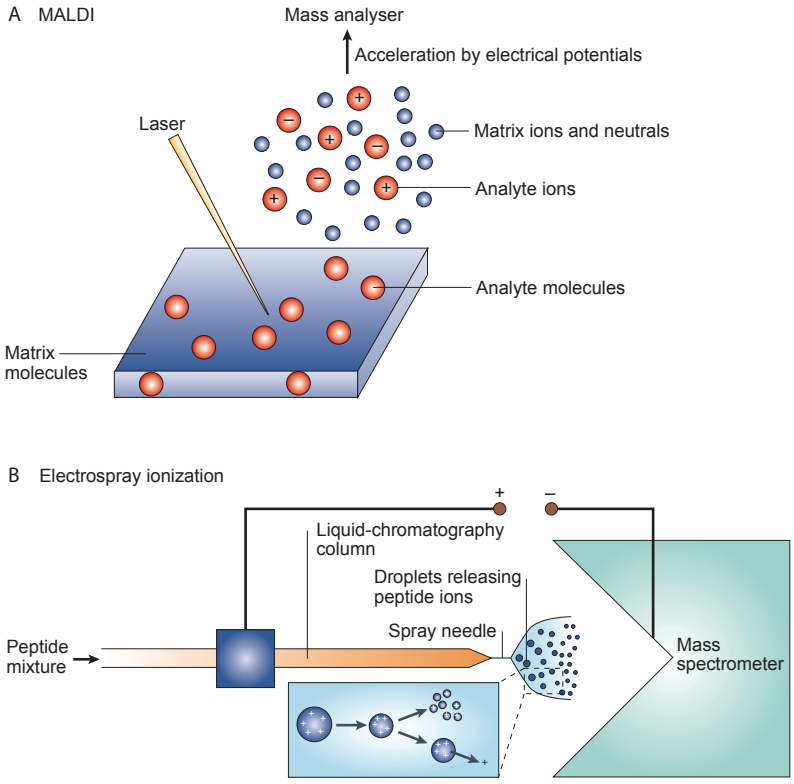

***Figure 2**: Soft ionization techniques most frequently used in proteomics experiments. A) Matrix Assisted Laser Desorption Ionization (MALDI) is based on the desorption of matrix and analyte by intense UV laser pulses, resulting in a hot plume and matrix-to-analyte proton transfer. B) Electrospray Ionization (ESI) induces the formation of small droplets that are further desolvated due to increasing temperature and increasing electrostatic repulsions. (Adopted from Ref. 9)*

analysers are based on m/z resonance frequency.(28) The choice of mass analyser depends on the type of proteomics experiment. In the discovery proteomics approach detection speed and resolution is of high importance, whereas in targeted experiments selectivity, sensitivity and dynamic range have priority. Mass analysers can also be used sequentially to address different needs and this is referred to as tandem MS and MS^n or hybrid MS when different types of analysers are coupled together.(10) The most frequently used mass analysers are discussed below.

Quadrupole (Q) – As the name suggests this mass analyser consists of a square array of four precision-aligned cylindrical or ideally hyperbolic shaped metal rods (Figure 3A). (28, 29) While ions are traveling through the assembly, variable voltages can be applied to the rods. The potentials are the same for opposing and 180° out of phase for adjacent rods and these sets are frequently switched in polarity, giving rise to a dynamic electric field, manipulating the trajectory of ions traveling through. When a specific set of RF and DC voltages is applied, this trajectory is circular in the xy-plane for a specific m/z value and therefore

analytes with the same m/z are able to pass through in a three dimensional corkscrew trajectory (Figure 3B). Ions with other m/z values will follow a trajectory that is instable resulting in the elimination of ions by ejection into the vacuum or collisions into the rods thereby losing its' charge and ability to reach the detector. When ramping specific sets of RF and DC voltages, different m/z regimes are allowed to pass through, resulting in the filtering of different ion packages that travel to the next mass analyser or detector. However, when only RF voltages are used the trajectory for a broad spectrum of m/z values are stable allowing it to capture and focus a whole range of charged ions. Different RF only configurations are used to maximize the transmission efficiency of a broad m/z range and can include six or eight rods or square shaped quadrupoles resulting in hexapole, octopole and flatapoles, respectively. RF only multipole configurations are mainly used as ion guides to transport ions between different sections in the MS or in collision cells to focus the paths of fragment ions formed after precursor ion dissociation.

Quadrupoles are often used in tandem and hybrid MS configurations to select for a precursor m/z in the first quadrupole (Q1), followed by precursor ion fragmentation in a second inert gas filled RF only quadrupole (Q2), and delivery of the formed fragments to a second mass analyser to scan for resulting m/z species. General applications are triple quadrupole

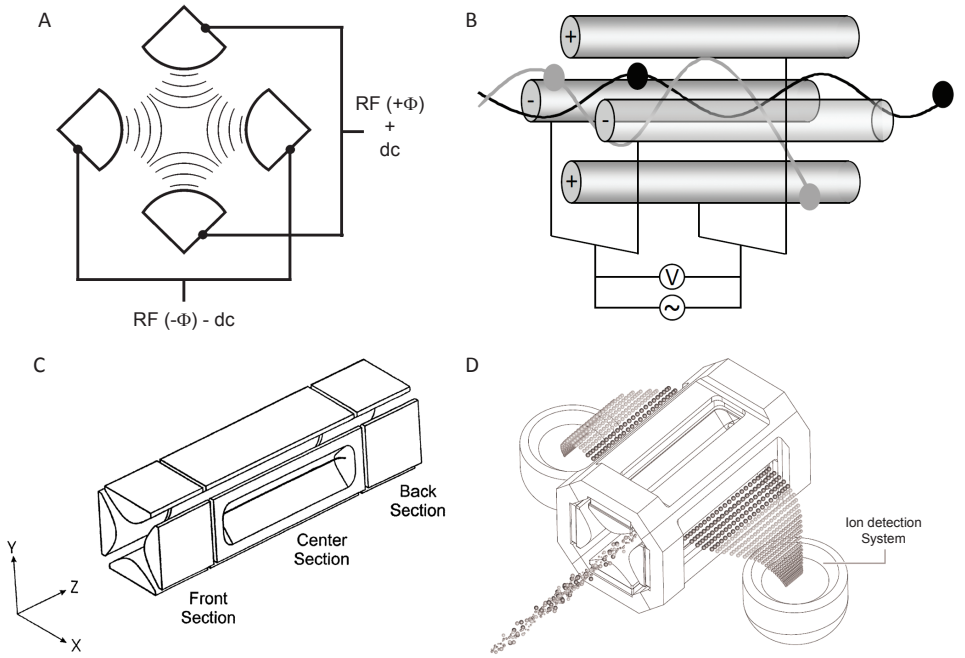

Figure 3: *Quadrupole-based mass analysers. (A) The Electric field generated by the differential voltages applied on the quadrupole array results in the repulsion/attraction of positively charged peptides to the metal rods. As a result, only those ions with a certain m/z remain in a stable trajectory over the quadrupole distance thereby removing all other ions with different m/z values (B). When the quadrupole is flanked by front and back lenses or quadrupoles, ions can be trapped inside the assembly (C). When the ions are selectively ejected from the trap (based on voltages destabilizing certain m/z values) through a slit in the rods, the signal of each ion species can be measured by a detector.*

(QqQ) MS and quadrupole time-of-flight (Q-TOF) MS. This type of mass spectrometers are referred to as tandem MS in space, as the ion manipulations are performed in different places.

Linear ion trap (LIT) – The LIT is essentially a quadrupole analyser with two electrodes on each side that repel or constrain the ions inside the rods (Figure 3C).(30, 31) The lineair ion trap is referred to as a 2D trap as ions are confined in the radial dimension by the quadrupolar field and in axial dimensions by the electric field of the closing electrodes. The entry and exit electrodes can reside in the form of a metal plate with bores or short quadrupoles. Contrary to quadrupoles, all ions of different masses within a specified m/z, are stored together and are ejected one by one according to their m/z to obtain a spectrum. When ions are ejected from a slit in two of the rods, they are collected at detectors for ion counting (Figure 3D). Due to charge-spacing effects ions repel each other when in close proximity thereby destabilizing their trajectories. Therefore, the quadrupole trap is soaked in a bath of inert gas (usually Helium) to collisionally cool down the excess kinetic energy formed by charge spacing effects, focussing all ions towards the centre of the trap. Due to the 2D trapping along an axis there is less space-charge repulsion when compared to the 3D trapping, which traps at a focal point. As a result of this LITs are preferred over the older 3D ion trap designs as their trapping efficiency is 10-fold higher and their capacity is at least 400-fold higher, allowing for better sensitivity. (28) Another advantage of the IT is the possibility to fragment trapped ions. By ramping up auxiliary voltages the amplitude of ion oscillation can be increased resulting in additional gas collisions that transfer the gained kinetic energy into internal vibrational energy, eventually leading to ion fragmentation. The trapping of ions together with the possibility of fragmenting them allows for tandem MS in time. In contrast to tandem MS in space, the precursor mass analysis, fragmentation and fragment ion mass analysis can be performed within the same analyser. Using a single MS analyser one could also analyse fragments of fragments creating MS^3 or higher experiments. Recently LIT design optimizations have been made by the introduction of a dual pressure ion trap that decouples the pressure regimes for ion trapping and fragmentation in one cell and mass analysis in a second cell for more efficient trapping/fragmentation and faster scan rates, respectively.(32)

Orbitrap – The orbitrap is the most recently developed mass analyser that employs orbital trapping as a novel method for m/z analysis. The orbitrap was designed by Makarov in 2000 (33) and is based on the Kingdon trap (34) that was first described in 1923. The electrostatic trap consists of a barrel shaped outer electrode that is split in two by an insulating ring and an inner spindle-shaped electrode (Figure 4A). (35) The ions injected into the orbitrap are trapped due to the balance between centrifugal forces and electrical attractions to the inner electrode. The quadro-logarithmic potential distribution that is achieved by the DC voltage between the inner and grounded outer electrode and the geometry of the orbitrap, results in a harmonic oscillation of the ions along the spindle axis. The ion oscillations in the orbitrap induce an image current on the outer electrodes that is recorded over time by a differential amplifier connected to the two halves. The resulting time-domain signal is converted using fast Fourier transform into the frequency domain. (36) Due to the harmonic properties of the axial oscillations the angular ion frequency is independent of initial parameters, de-

pending only on the mass and charge of the ion. Therefore, the mass-to-charge spectrum is easily calculated from the recorded angular frequency.

The orbitrap is a popular mass analyser in proteomics experiments because of its' high resolution (routinely 60 000, going up to 450 000) and high mass accuracy (<2ppm) at LC separation time scales. (37, 38) Together with the speed of linear ion traps and tandem in time MS functionality, the hybrid LTQ-Orbitrap (Thermo Fisher) (Figure 4B), is one of the

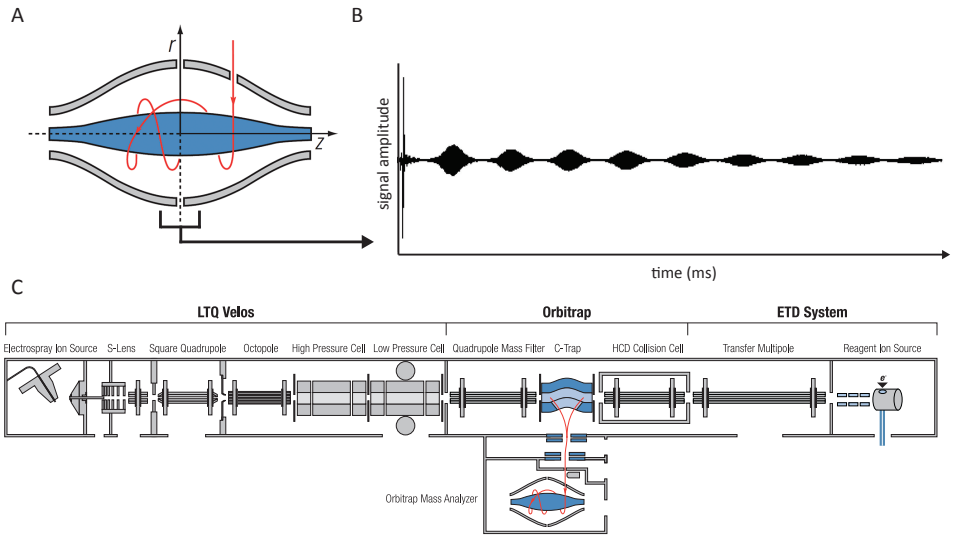

Figure 4: *Orbitrap mass analyser. A) Schematic of the Orbitrap geometry. Ions move in spiral orbits around the inner electrode while their image current is recorded at the outer electrodes resulting in a time domain spectrum (B) that is converted using fourier transform into a frequency domain that in turn is calculated to a mass-to-charge spectrum (using $m/z=k/\omega^2$). C) A hybrid LIT-Orbitrap mass spectrometer (Thermo LTQ-Orbitrap Velos), including a dual pressure ion trap, dedicated HCD collision cell and reagent anion source for ETD fragmentation.*

most frequently used configurations in many proteomics labs nowadays. (39) Later additions of multiple fragmentation techniques, improvements in speed and sensitivity and the development of a new Q-Orbitrap hybrid MS resulted in the Orbitrap as the central mass analyser in most proteomic analyses. (32, 37, 40, 41)

Detectors

When ions are separated/detected by mass analysers other than orbitraps, the signal of each m/z value needs to be determined in an additional step. This task is preferably performed by electron multipliers (EM) because of their high signal amplification and fast response times. A typical EM consists of a continuous dynode with a concave metal surface that catches ions flying in originating from the Q, LIT or TOF mass analyser (Figure 5). When an ion hits the surface, one or multiple secondary electrons are formed that collide again resulting in an cascade of electrons and an exponential signal increase (up to 10^7) delivering a final measurable current at the end of the dynode.(28)

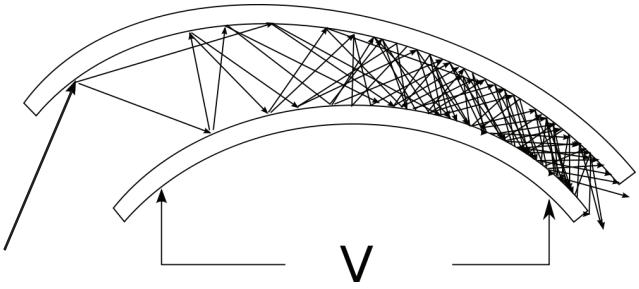

Figure 5: *The electron multiplier detector principle is based on the generation of an electron multiplication cascade, amplifying the initial ion signal up to 7 orders of magnitude.*

Peptide fragmentation

The m/z of a precursor ion alone is in most cases not unique and therefore the identity of a peptide and thus protein remains ambiguous. Therefore, in order to uncover the identity of a peptide, its' precursor ion is fragmented into its' building blocks to determine the peptide sequence and consequential protein ID. There

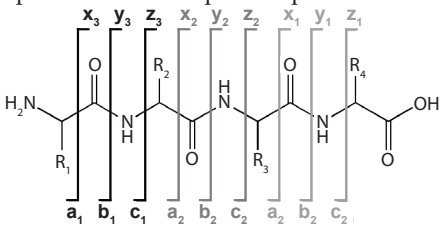

Figure 6: *Roepstorff-Fohlman nomenclature of peptide backbone fragments observed in peptide fragmentation spectra. Fragment ions created by collision-induced fragmentation (CID) are predominantly b- and y-ions. Peptide fragments generated by ETD are typically c- and z-ions.*

are several low-energy fragmentation techniques available that dissociate a peptide at its' backbone. The nomenclature of different peptide fragments is depicted in Figure 6 and was first described by Roepstorf and Fohlman. (42)

The most frequently used peptide fragmentation method is collision-induced dissociation (CID). (43) In CID peptides are excited by an electric field resulting in an increase in translational/kinetic energy that is converted into internal, mostly vibrational, energy upon collision with inert diatomic gas molecules such as Helium, Nitrogen, Argon or Xenon. The site of peptide fragmentation is depicted in the mobile proton model (44) and is primarily dictated by the proton affinity of the fragment. Upon excitation protons localised on the basic side chain can migrate along the backbone with a preference for amines. The proximity of a proton to the backbone amino-groups weakens the C-N bonds that together with the increase of internal energy will result in peptide fragmentation, predominantly generating b- and y-ions. This "slow-heating" process can be further subcategorized into resonant-excitation and beam-type CID. The former fragmentation type is performed in an ion trap and is also referred to as IT-CID or tandem in time. Due to the restricted space, only a small amount of kinetic energy can be deposited onto the peptide using resonating frequencies (typically a few eV). The low amount of energy in concert with a higher pressure of small

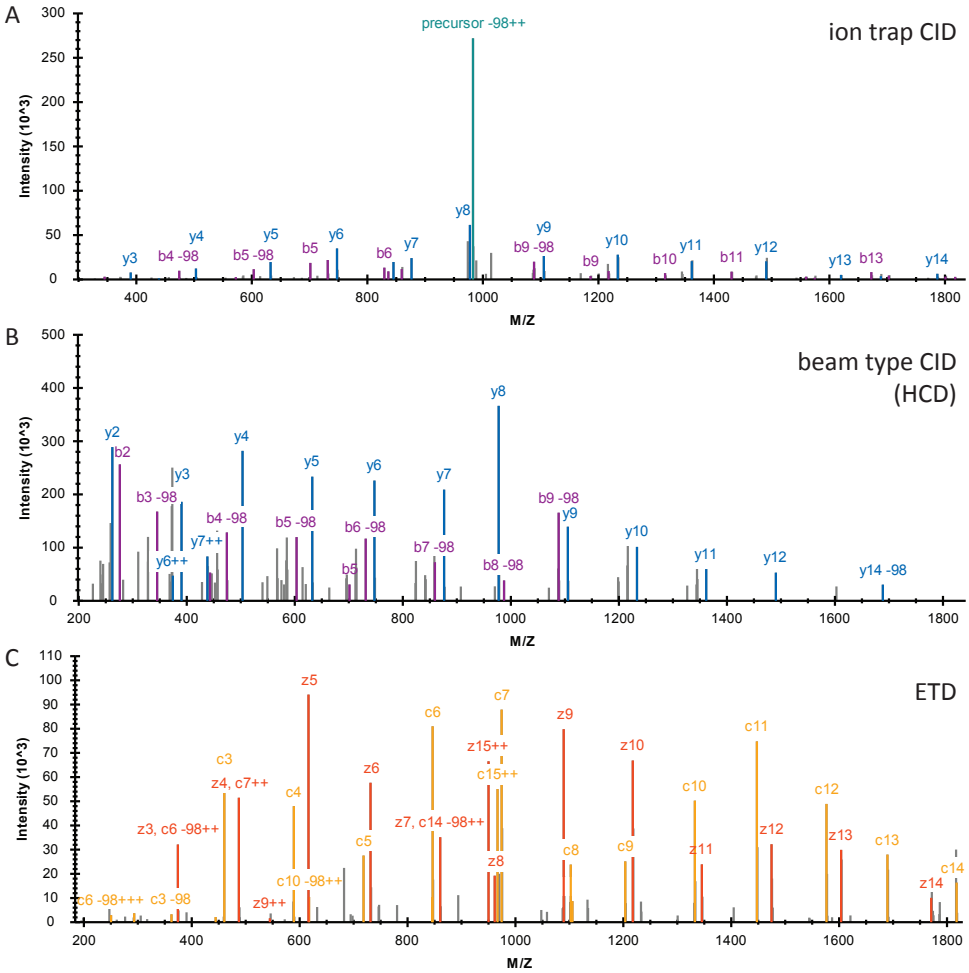

Figure 7: *MS/MS spectra of a phosphorylated peptide (FQSEEQQQTEDELQDK). The doubly charged precursor (1031.42 m/z) was fragmented using IT-CID (A) and beam type CID (HCD) (B). The triply charged precursor (687.95 m/z) was best fragmented using ETD (C).*

Helium gas molecules results in the accumulation of energy via multiple gas collisions in the millisecond time scale, before fragmentation occurs. Beam type CID is applied in dedicated collision cells in tandem quadrupoles (QqQ, Q-TOF) or multipoles (HCD) and thus is sometimes referred to as tandem in space. On the contrary to IT-CID, beam type CID allows for a higher amount of energy transfer on a faster time scale. In beam type CID the kinetic energy is applied before entering a gas chamber via the acceleration of the ions in an electric field. This allows for higher energy transfer rates, typically several tens to hundreds of eV, that are mainly converted into internal energy by a first collision with a larger gas molecule such as Nitrogen or Argon. As a result of the fewer collisions a higher energy is converted in a smaller timeframe (microseconds range). The faster energy deposition in beam-type CID has a benefit over the slower IT-CID in that fewer chemical reactions/rearrangements can

occur before fragmentation, resulting in less (unexplainable) fragments or neutral losses. This is especially important in the case of PTM analysis such as phosphorylation where a weak phosphate bond results in predominant precursor phosphate losses that dominate the MS/MS spectrum (Figure 7AB). (45) However, a downside of beam type CID is that fragment ions retain part of the kinetic energy allowing for further fragment decomposition of more instable b-ions. (46) This loss of b-ion fragments is less pronounced in IT-CID where only precursor ions are excited by resonance frequencies. In addition, fragment ion scattering is restricted in the ion trap by multiple lower energy gas collisions resulting in "collisional cooling" leading to better fragment ion capture and thus better fragmentation efficiency.

A more recently introduced fragmentation method is electron transfer dissociation (ETD). (47) In ETD, electrons are transferred from a radical anion (e.g. fluoranthene) to a peptide cation brought in close proximity to each other in the ion trap. When the positively charged multi-protonated peptides capture an electron, the backbone fragments preferentially at N-Cα bonds generating c- and z-type ions (Figure 6 and 7C). Due to low charge densities of doubly charged cation peptides, ETD fragmentation efficiency is very low resulting in charge reduced and non-fragmented precursor ions. (48) However, the fragmentation efficiency of lower charged peptides can in part be overcome by a supplemental (resonance) excitation. (49, 50) Since ion trap CID is less efficient on peptides with more than 2 positive charges, ETD can be used as a complementary fragmentation technique to CID. Recently this concept was developed into an hybrid CID and ETD enabled mass spectrometer (Figure 4C), allowing for an automated on-the-fly decision tree that picks the most optimal fragmentation method based on the precursor charge and length. (51, 52) Next to the ability of ETD to more efficiently fragment higher charged peptides, acid labile modifications such as phosphorylation and glycosylation are left intact during fragmentation. Therefore ETD fragmentation is more frequently applied in experiments studying PTMs to more reliably identify and localize PTM sites. (53, 54)

2.4 Data analysis and statistics

Generated MS/MS spectra contain information concerning the peptide sequence and consequently protein identity. However, as single LC-MS/MS runs typically generate tens of thousands of spectra, manual data interpretation is not feasible. In order to assign spectra to peptides in a more automated manner several database search algorithms have been developed, such as Sequest (55) and Mascot (56). The general first step in these automated methods involves the generation of a list of theoretical peptides and their corresponding masses based on an *in silico* proteome digest. The size of this list is mainly determined by the amount of proteins in the database, the type of cleavage enzyme used and the number of peptide modifications. Subsequently, the masses of all the in silico derived peptides are compared to the acquired precursor masses and only peptide candidates matching within a given mass tolerance are kept for further MS/MS comparison. In the next step, in silico fragment masses are generated for each in silico peptide sequence matching the empirical precursor m/z. Finally, a score is calculated for similarity between the experimental and in

silico-derived spectrum, based on the number and type of fragment ion masses matched. The sequence with the highest score is considered the observed peptide sequence resulting in a peptide spectrum match (PSM). However, when matching tens of thousands spectra to millions of possible candidates the chance of obtaining random incorrect hits is likely. The amount of random hits can be reduced by lowering the amount of possible sequence candidates by decreasing the mass tolerance for precursor and fragment matching and can be achieved by employing high accuracy mass analysers. Even so, the sheer amount of data will always result in false-positives. To estimate the rate of false discovery (FDR), several statistical methods have been developed. The most straightforward approach is to perform a second search of the experimental data against a nonsense database (derived from reversed or scrambled protein sequences). (57) Any match to this so-called decoy database is considered a false positive and the number of decoy PSMs divided by the total number of PSMs gives the FDR ratio. Typically an FDR of 1 % is accepted for large-scale proteomics experiments. (58) The FDR can be lowered by applying additional filters to PSM parameters, such as a minimum peptide score and/or a minimum sequence length. The application of these filters often result in a loss of true positives, therefore more sophisticated machine learning algorithms have been developed that calculate the weight of multiple discriminating parameters taking into account other parameters such as e.g. peptide hydrophobicity, charge and number of peptides that match to a protein. (59)

2.5 Quantitative proteomics

Mass spectrometry cannot directly be used to quantify peptides or proteins. In bottom-up proteomics relative protein abundances are derived from peptide signal intensities. However, the signal of peptides in the MS are not solely dependent on their concentration. On the contrary, the measured peptide intensity is strongly influenced by other parameters such as peptide digestion efficiency, solubility and ionizability. Therefore, the MS responses obtained for different peptide sequences, hence different physicochemical properties, are not lineair to their concentration and cannot be compared. Different strategies have been developed to enable quantitative proteomics experiments and they are always based on the comparison of identical sequences. The quantification methods can be divided into two groups namely label-free and stable isotope labelling. (60) Both strategies are based on the comparison of identical peptide sequences between different origin (Figure 8).

Label-free quantification

This quantification method is based on the comparison of peptides from different samples in seperate LC-MS/MS runs. The comparison between different runs can be based on either peptide MS signal (61) or number of spectral matches (62). (63) It is the most straightforward quantification method suitable for all types of samples requiring no additional sample preparation steps. Each sample is analysed separately and therefore the number of samples that can be compared is infinite, albeit limited by the instrument time available. Quantification precision of label-free methods can be compromised due to variations in

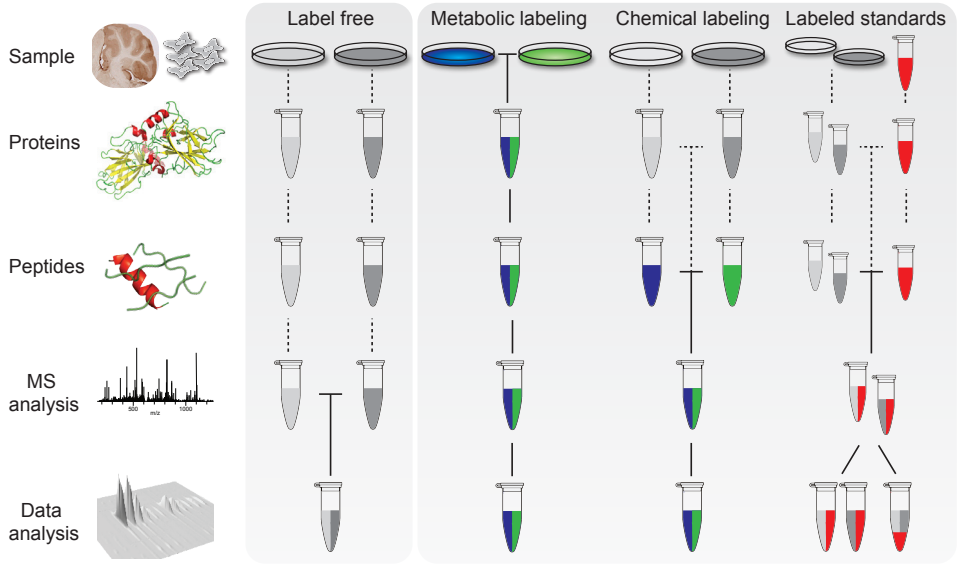

Figure 8: *Quantitative mass spectrometry-based proteomics workflows.*

sample preparation (e.g. sample lysis and digestion variability) and LC-MS conditions (e.g. peptide retention or ionization variability) that can be different for each sample/injection. Therefore, strong requirements are needed for a minimal number of reproducible sample preparation steps. As a result of these criteria proteome coverage is generally lower due to lack of additional fractionation steps. Furthermore, stable reproducible LC peak shape and retention time are required, which, to some extent, can be alleviated by normalization between runs. Normalization can be achieved with spiked-in calibrants and abundant non-changing peptides as landmarks between runs or by using software for peak matching between runs. (64-66)

Isotope labelling quantification

Labelling in quantitative MS is based on the incorporation of different stable isotopes to distinguish between samples. When comparing identical peptide sequences from differently labeled samples their chemical properties remain the same however their mass will be distinguishable in the mass spectrometer. Therefore, when combining the samples before or halfway the proteomics pipeline, variation in sample preparation and LC-MS are equal for both samples (Figure 8). As a result experimental fluctuations are mitigated resulting in a more precise protein regulation determination. The introduction of the heavy isotope (e.g. ^{15}N, ^{13}C) can be achieved in different ways and the most frequently used methods are described below.

Metabolic labelling – is achieved by growing cells, plants or even animals on a heavy isotope enriched food source or by the substitution of specific amino acids that contain ^{15}N and/or ^{13}C instead of the naturally abundant ^{14}N and ^{12}C isotopes. (67-69) The proteins of cells or

organisms grown in these conditions will contain exclusively (~98%) heavy peptides. The early implementation of labels in the proteomics workflow allows for minimal variations and the highest theoretical accuracy of protein ratio determination. The most popular method is stable isotope labelling by amino acids in cell culture (SILAC). (69) In this method cells are grown in culture medium that lack essential amino acids that can be supplemented with heavy versions of the same amino acid. When switching to the modified medium the cells need to undergo several cell doublings in order to contain a complete heavy proteome. In proteomics experiments heavy arginine and lysine are most often used since every tryptic peptide will contain a label (except for c-terminal peptides). Serveral factors are limiting its use including the cost and time for maintaining the model system on expensive media as well as the inability of certain cell lines and organisms to implement the labels. Furthermore, the number of samples that can be compared is classically restricted to two or three.

Chemical labelling – forms a good alternative for biological systems such as human tissue or fluid and other post-biosynthetic samples that are not suited for metabolic labelling. There is a big variety in isotopic labels that can be introduced at different stages of the proteomics workflow. The addition of a chemical label at the protein level can be performed using an isotope-coded affinity tag (ICAT). (70) In ICAT experiments, protein reduction is followed by labelling of free reactive cysteines. The implementation of the label on the protein level allows for sample pooling before digestion removing the variability in protein digestion. However, due to the low number of peptides that can be used for quantification (only those containing cysteines), its' application is not ideal.

The majority of chemical labelling strategies target the primary amine groups present on lysines and free amino termini of peptides or proteins. One example of a fast and largely complete labelling strategy is stable isotope dimethyl labelling. (71, 72) In this protocol formaldehyde and cyanoborohydride with a different number of ^{13}C and deuterium atoms are used to create dimethylated peptides with a minimum of 4 Da difference between each sample. The reagents used are very cost-effective and this type of labelling is applicable to any type of biological sample. A possible downside of using this type of label is the slightly different chemical properties of deuterium when compared to hydrogen that can result in subtle retention time shifts potentially compromising quantification. (73) Furthermore, as the label is introduced at a later stage in the sample workup more experimental variability can be expected. However, a recent study comparing different labelling strategies concluded this type of labelling has a similar accuracy to metabolic labelling. (74)

Other popular chemical labels targeting primary amines are the isobaric tags termed tandem mass tag (TMT) (75) and isobaric tag for relative and absolute quantification (iTRAQ) (76). Both reagents consist of a reporter moiety with a differential mass and a balance group that both make up a total mass that is identical for each label. As a result the peptides originating from different samples remain identical in precursor mass leading to less complex MS1 spectra with higher precursor signals. Upon precursor fragmentation the isobaric tags fall apart in the different reporter ions and neutral balancing groups. In contrast to the labels described before that use MS1 spectra for quantification, isobaric tags use the reporter ion signals from MS/MS spectra. The benefit of this approach is its' superior multiplexing

capability as neither LC separation nor MS1 spectra suffer from increased complexity. Currently, reagents are available allowing for the analysis of up to 10 samples in a single run, greatly increasing the analytical throughput. However, one need to take into consideration that other peptides with similar precursor masses are co-isolated and fragmented, negatively influencing quantification accuracy. (77) Moreover, mass spectrometers capable of measuring low m/z values are needed. Particularly, ion traps are not suitable for these types of analysis as they inherently miss the low m/z range in their MS/MS spectrum due to low trapping efficiencies for low m/z species.

Stable isotope labelled standards (SIS) – are used for both relative and absolute quantification of peptides/proteins. The absolute concentration of proteins can be measured when the concentration of the SIS peptide is known and is often expressed as copies per cell. (78, 79) In contrast to the labelling strategies described before, this strategy is not based on labelling of the sample but on the introduction of a labelled standard in each sample. Basically, stable isotope dilution is an improved label-free quantification method in which isotopically labelled peptide analogues are spiked into the sample to determine the ratio between natural peptide and standard peptide levels (NAT/SIS). By comparing the NAT/SIS ratio between different samples experimental deviations are normalized and more precise quantifications can be achieved when compared to label-free analysis. The main benefits of this method are its' improved accuracy, applicability to any type of sample and the absence of additional sample labelling steps. The heavy standard can be introduced just before LC-MS analyses using heavy synthetic peptides (often referred to as AQUA) (79) or before sample digestion using either biosynthesized heavy proteotypic peptide concatemers (QconCAT) (80) or synthetic heavy full length proteins (PSAQ) (81) (Figure 8). Obviously, the latter methods are preferred due to the additional normalization for digestion and protein enrichment/solubility variability. However, synthesis of fully labelled proteins is still expensive, therefore AQUA peptides are more frequently used. Due to the lack of multiplexing capabilities the analytical throughput of SIS-based methods is lower compared to metabolic and chemical labelling strategies.

3. Aging, Cancer and Senescence

Aging can be arguably defined as the progressive loss of bodily functions, reduced fertility and increased mortality over an organisms' lifespan. It has been well established that aging is an unpredictable, complex and stochastic process. (82) Damaging factors appear randomly during the organism lifespan that affect all biological molecules. Exogenous factors include various toxic chemicals and UV radiation exposure whereas endogenous stressors can include spontaneous chemical reactions (e.g. hydrolysis) and radical enzymatic side products such as reactive oxygen and nitrogen species. (83, 84) Under these stressful conditions proteins, lipids and DNA can be modified with different intermediate metabolites (e.g. oxidation), rendering these biomolecules unfunctional or even cytotoxic when accumulated. (85, 86) Misfolded proteins are associated with age-related pathologies including Alzheimer's and Parkinsons' disease. As a defensive mechanism, most altered protein forms are not repaired but selectively degraded and replaced by new copies.

In contrary to protein, metabolites and lipids, there is no template for de novo DNA biosynthesis that allows for an easy replacement of damaged DNA strands. The DNA is known to be increasingly oxidized as age progresses, where mitochondrial DNA is more strongly affected than nuclear DNA. (87) Mitochondrial aging has indeed been heavily studied for its' contributions to organismal aging. (88) However, nuclear DNA forms the main stage for aging research. Nuclear DNA is generally present at 2-4 copies per cell whereas mitochondrial DNA is present at several thousand copies per cell. Not surprisingly, different nuclear DNA-damage repair mechanisms have evolved over time. (89) Due to the magnitude of damages that occur (estimated to include up to 10^5 DNA lesions per cell per day (83)) and the complexity of these repair mechanisms(90), defence mechanisms do not remain error free. As a consequence, unrepaired DNA is propagated into daughter cells accumulating damage over time. Indeed, many premature aging syndromes have aberrant nuclear DNA-repair mechanisms, pinpointing DNA as a critical target of aging. (89, 91, 92)

Eventually, the accumulation of DNA perturbations can alter protein function and activity compromising cellular functions and thus organismal integrity.(93) Different types of DNA damage such as base conversion, single and double strand DNA breakage can give rise to different cellular fates including DNA mutations, chromosome aberration, blocked transcription/replication or programmed cell death (apoptosis). Mutations and chromosomal rearrangement can result in perturbed regulating or coding sequences, leading to increased or decreased gene transcription, constitutive active or unactive proteins or protein misfolding. The increased activity of growth promoting oncogenes or the loss of tumor suppressor genes can result in uncontrolled growth, eventually leading to cancer. Well known tumor suppressor pathways perturbed in most cancers include p16(INK4A), Retinoblastoma(94) and cellular tumor antigen p53(95). (96) Proliferative signalling mediators such as RAS(97), BRAF(98) and PI3K-Akt(99) are among the most frequently observed cancer oncogenes.

Another type of DNA damage can occur due to the intrinsic shortening of DNA after each cell division. Normal cells face a limited number of cell divisions before the DNA has reached a critically short length. When this limit is reached the DNA becomes instable, leading to the activation of different DNA repair mechanisms instigating cell cycle arrest. (100, 101) This process, called cellular or replicative senescence(102), is characterized by a state of physiologically irreversible growth arrest, increased cell size, the expression of tumor suppressor protein p16(INK4A), increased senescence-associated β-galactosidase activity (103) and increased secretion of growth factors, proteases and cytokines(104-106). After the first description of cellular senescence by Hayflick and Moorhead in 1961, several other factors leading to senescence-like phenotypes have been described. It has been shown that senescence can be initiated by severe genomic damage(107), chromatin relaxation(108) oncogene activation(109-112) and hyperphysiological stress induced by artificial cell culture (plastic dish substrate, highly mitogenic serum or atmospheric oxygen levels)(113, 114). In contrary, to the definite and deleterious tumor suppressive mechanism apoptosis, senescence allows damaged cells to survive and exert their (sometimes compromised) function, prolonging tissue homeostasis. Therefore, the induction of senescence in tumor tissues by external intervention could pose a promising anti-cancer therapy.

Senescence; connecting cancer protection to aging

In an evolutionary context tumor suppressive mechanisms, like senescence, are beneficial for species survival. In a hazardous milieu, tissue renewal and growth is essential for the viability of the organism allowing it to produce offspring, thereby increasing the chances of species survival. However, the uncontrolled expansion of tissue (resulting in cancer) is often fatal to the organism and therefore this growth needs to be dampened. In order to outlive their reproductive lifespan (which is a few decades for humans), organisms have acquired mechanisms to effectively reduce the negative effects of tumor growth. However, these mechanisms have grown to be beneficial for only the first reproductive years in their lifespan. Therefore, should such mechanism be harmful at a later age, there would be little selective pressure to eliminate such harmful mechanism. As a result, organisms that live longer, due to (external) stress relieve, suffer from senescent tissues that cannot be renewed. Consequently, tissue repair is hampered leading to a slow deterioration of bodily functions. This evolutionary theory, also called antagonistic pleiotropy, is one of the main theories of how aging becomes apparent.(115) Indeed increasing number of senescent cells correlate positively with age and age-related pathology, such as decreased neurogenesis and type II diabetes (116-118). Extensive research efforts have been made to understand aging in order to increase longevity. Similarly anti-tumor mechanisms, such as senescence, are heavily studied to find better medicines for cancer. Interestingly, these seemingly different research fields show a great overlap in terms of underlying biological mechanisms. Therefore, acquiring fundamental knowledge about the mechanisms that induce senescence can benefit both anti-aging and anti-cancer therapies.

4. Thesis outline

This thesis describes proteomics technology development and implementations to increase the knowledge about aging and cancer suppression mechanisms. The core tool in all performed experiments is mass spectrometry based proteomics. In addition to the application of these techniques, newly developed proteomics techniques are being introduced that help to increase the analytical power of proteomics experiments.

In chapter two, the influence of DNA-damage on tissue aging was studied for a specific part of the mouse brain over time. Quantitative discovery proteomics experiments were performed in order to find biomarkers and mechanisms of aging in the cerebellum. After the identification of several interesting differentially regulated proteins associated with age and neuron-deterioration, a more targeted validation was performed using immunohistochemistry. By using this complementary method a spatio-temporal analysis of differentially regulated proteins could be monitored. The main observations of this study are the DNA-damage induced down regulation of a tightly interconnected network of proteins responsible for synaptic signaling associated with morphological transformation of Purkinje cells and a molecular link between DNA-damage and motoric diseases such as spinocerebellar ataxia.

Chapter three describes a technical method development to optimize the transition from a discovery orientated proteomics approach to a targeted proteomics phase. Different fragmentation techniques used in discovery experiments were compared to the fragmentation technique used in single reaction monitoring (SRM, also referred to as MRM) measurements. Furthermore, the feasibility of using these, discovery phase acquired, spectra for SRM assay development was tested. The work demonstrates that CID spectra acquired on SRM instruments are more similar to spectra acquired in the HCD cell than those acquired in the ion trap of hybrid LTQ-Orbitrap instruments. Concomitantly, SRM assays generated using HCD spectra showed a higher sensitivity when compared to ion trap spectra-generated SRM assays. Therefore, when planning a targeted MS experiment, choosing for HCD fragmentation in the discovery phase can help facilitate SRM assay development later on.

In chapter four, the phosphopeptide enrichment robustness of a new Ti^{4+}-IMAC method was assessed. First it was established that Ti^{4+}-IMAC enrichment resulted in a highly reproducible quantification of phosphorylation sites in HeLa cells. Subsequently, this strategy was applied to monitor the phosphoproteome of Jurkat T-cells upon Prostaglandin E_2 stimulation over a time scale of 60 minutes. It was demonstrated that using this enrichment strategy and label-free quantification a comprehensive temporal phosphoproteome of Jurkat T-cells could be constructed, indicating differential regulation of different kinases over time. The proved straight-forward yet comprehensive phosphoproteomics workflow and its' applicability to every sample type could form a good alternative method for the phosphoproteomics community.

Chapter five describes a study performed to increase knowledge about the mechanisms involved in oncogene-induced senescence (OIS). Using both whole proteome and phosphoproteome analysis, differential protein expression and phosphorylation were compared between cycling, senescent and tumorigenic human skin cells. Proteome analysis revealed a strong resemblance between senescent and tumorigenic cells. In both cell lines an up regulation of the retinoblastoma tumor suppressor mechanism was observed that was more pronounced in OIS. Proteins specifically regulated in OIS included senescence markers, inflammatory proteins and extracellular processing factors amongst other previously senescence-unassociated proteins. Phosphorylation screening identified strong reduction in cyclin-dependent kinase activity and numerous sites to be differentially regulated in OIS. This work forms an extensive resource of protein and phosphorylation regulations associated with OIS, benefitting researchers in the field of tumor suppressor research.

In chapter six an outlook is described regarding the validation of proteomics-derived target lists with the emphasis on cell signaling analysis. Currently, antibody-based methods are used predominantly to validate and investigate the involvement of proteins and post translation modifications (PTMs) in different biological processes. The limitation of antibody-based methods are described and targeted mass spectrometry-based methods are proposed as an alternative approach to overcome these problems. Especially for PTM analysis, targeted mass spectrometry has a high potential that will likely become a standard methodology for dynamic cell signaling analysis.

References

1. Rockman, M. V., and Kruglyak, L. (2006) Genetics of global gene expression. Nat Rev Genet 7, 862-872.

2. Gygi, S. P., Rochon, Y., Franza, B. R., and Aebersold, R. (1999) Correlation between protein and mRNA abundance in yeast. Mol Cell Biol 19, 1720-1730.

3. Hegde, P. S., White, I. R., and Debouck, C. (2003) Interplay of transcriptomics and proteomics. Current opinion in biotechnology 14, 647-651.

4. Abbott, A. (1999) A post-genomic challenge: learning to read patterns of protein synthesis. Nature 402, 715-720.

5. Wilkins, M. R., Pasquali, C., Appel, R. D., Ou, K., Golaz, O., Sanchez, J. C., Yan, J. X., Gooley, A. A., Hughes, G., Humphery-Smith, I., Williams, K. L., and Hochstrasser, D. F. (1996) From proteins to proteomes: large scale protein identification by two-dimensional electrophoresis and amino acid analysis. Biotechnology (N Y) 14, 61-65.

6. Umar, A., Kang, H., Timmermans, A. M., Look, M. P., Meijer-van Gelder, M. E., den Bakker, M. A., Jaitly, N., Martens, J. W., Luider, T. M., Foekens, J. A., and Pasa-Tolic, L. (2009) Identification of a putative protein profile associated with tamoxifen therapy resistance in breast cancer. Molecular & cellular proteomics : MCP 8, 1278-1294.

7. Ong, S. E., Schenone, M., Margolin, A. A., Li, X., Do, K., Doud, M. K., Mani, D. R., Kuai, L., Wang, X., Wood, J. L., Tolliday, N. J., Koehler, A. N., Marcaurelle, L. A., Golub, T. R., Gould, R. J., Schreiber, S. L., and Carr, S. A. (2009) Identifying the proteins to which small-molecule probes and drugs bind in cells. Proceedings of the National Academy of Sciences of the United States of America 106, 4617-4622.

8. Lander, E. S., et al. (2001) Initial sequencing and analysis of the human genome. Nature 409, 860-921.

9. Steen, H., and Mann, M. (2004) The ABC's (and XYZ's) of peptide sequencing. Nat Rev Mol Cell Biol 5, 699-711.

10. Yates, J. R., Ruse, C. I., and Nakorchevsky, A. (2009) Proteomics by mass spectrometry: approaches, advances, and applications. Annu Rev Biomed Eng 11, 49-79.

11. Altelaar, A. F., Munoz, J., and Heck, A. J. (2013) Next-generation proteomics: towards an integrative view of proteome dynamics. Nat Rev Genet 14, 35-48.

12. Olsen, J. V., Ong, S. E., and Mann, M. (2004) Trypsin cleaves exclusively C-terminal to arginine and lysine residues. Molecular & cellular proteomics : MCP 3, 608-614.

13. Li, X., Gerber, S. A., Rudner, A. D., Beausoleil, S. A., Haas, W., Villen, J., Elias, J. E., and Gygi, S. P. (2007) Large-scale phosphorylation analysis of alpha-factor-arrested Saccharomyces cerevisiae. Journal of proteome research 6, 1190-1197.

14. Posewitz, M. C., and Tempst, P. (1999) Immobilized gallium(III) affinity chromatography of phosphopeptides. Analytical chemistry 71, 2883-2892.

15. Larsen, M. R., Thingholm, T. E., Jensen, O. N., Roepstorff, P., and Jorgensen, T. J. (2005) Highly selective enrichment of phosphorylated peptides from peptide mixtures using titanium dioxide microcolumns. Molecular & cellular proteomics : MCP 4, 873-886.

16. Zhou, H., Ye, M., Dong, J., Han, G., Jiang, X., Wu, R., and Zou, H. (2008) Specific phosphopeptide enrichment with immobilized titanium ion affinity chromatography adsorbent for phosphoproteome analysis. Journal of proteome research 7, 3957-3967.

17. Ficarro, S. B., McCleland, M. L., Stukenberg, P. T., Burke, D. J., Ross, M. M., Shabanowitz, J., Hunt,

D. F., and White, F. M. (2002) Phosphoproteome analysis by mass spectrometry and its application to Saccharomyces cerevisiae. Nature biotechnology 20, 301-305.

18. Kweon, H. K., and Hakansson, K. (2006) Selective zirconium dioxide-based enrichment of phosphorylated peptides for mass spectrometric analysis. Analytical chemistry 78, 1743-1749.

19. Zhou, H., Ye, M., Dong, J., Corradini, E., Cristobal, A., Heck, A. J., Zou, H., and Mohammed, S. (2013) Robust phosphoproteome enrichment using monodisperse microsphere-based immobilized titanium (IV) ion affinity chromatography. Nat Protoc 8, 461-480.

20. Pandey, A., Podtelejnikov, A. V., Blagoev, B., Bustelo, X. R., Mann, M., and Lodish, H. F. (2000) Analysis of receptor signaling pathways by mass spectrometry: identification of vav-2 as a substrate of the epidermal and platelet-derived growth factor receptors. Proceedings of the National Academy of Sciences of the United States of America 97, 179-184.

21. Steen, H., Kuster, B., Fernandez, M., Pandey, A., and Mann, M. (2002) Tyrosine phosphorylation mapping of the epidermal growth factor receptor signaling pathway. J Biol Chem 277, 1031-1039.

22. Tanaka, K. (2003) The origin of macromolecule ionization by laser irradiation (Nobel lecture). Angew Chem Int Ed Engl 42, 3860-3870.

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

24. Karas, M., and Hillenkamp, F. (1988) Laser desorption ionization of proteins with molecular masses exceeding 10,000 daltons. Analytical chemistry 60, 2299-2301.

25. Taylor, F. B. (1964) The Application of Carbon Chloroform Extraction Technic to Several Municipal Water Supplies. Health Lab Sci 1, 33-43.

26. Iribarne, J. V., and Thomson, B. A. (1976) On the evaporation of small ions from charged droplets. Journal of Chemical Physics 64, 2287-2295.

27. Whitehouse, C. M., Dreyer, R. N., Yamashita, M., and Fenn, J. B. (1985) Electrospray interface for liquid chromatographs and mass spectrometers. Analytical chemistry 57, 675-679.

28. Hoffmann, E. d., and Stroobant, V. (2007) Mass spectrometry : principles and applications, 3rd ed. Ed., Wiley ; Chichester : John Wiley [distributor], Hoboken, N.J.

29. Paul, W., and Steinwedel, H. (1953) Ein neues Massenspektrometer ohne Magnetfeld. Zeitschrift Naturforschung Teil A 8, 448.

30. Hager, J. W., and Le Blanc, J. C. (2003) High-performance liquid chromatography-tandem mass spectrometry with a new quadrupole/linear ion trap instrument. J Chromatogr A 1020, 3-9.

31. Douglas, D. J., Frank, A. J., and Mao, D. (2005) Linear ion traps in mass spectrometry. Mass spectrometry reviews 24, 1-29.

32. Olsen, J. V., Schwartz, J. C., Griep-Raming, J., Nielsen, M. L., Damoc, E., Denisov, E., Lange, O., Remes, P., Taylor, D., Splendore, M., Wouters, E. R., Senko, M., Makarov, A., Mann, M., and Horning, S. (2009) A dual pressure linear ion trap Orbitrap instrument with very high sequencing speed. Mol Cell Proteomics 8, 2759-2769.

33. Makarov, A. (2000) Electrostatic axially harmonic orbital trapping: a high-performance technique of mass analysis. Analytical chemistry 72, 1156-1162.

34. Kingdon, K. H. (1923) A Method for the Neutralization of Electron Space Charge by Positive Ionization at Very Low Gas Pressures. Physical Review 21, 408-418.

35. Makarov, A., Hardman, M. E., Schwartz, J. C., and Senko, M. (2004) US Patent 2004108450.

36. Senko, M. W., Canterbury, J. D., Guan, S., and Marshall, A. G. (1996) A high-performance modular data system for Fourier transform ion cyclotron resonance mass spectrometry. Rapid communications in mass spectrometry : RCM 10, 1839-1844.

37. Michalski, A., Damoc, E., Lange, O., Denisov, E., Nolting, D., Muller, M., Viner, R., Schwartz, J., Remes, P., Belford, M., Dunyach, J. J., Cox, J., Horning, S., Mann, M., and Makarov, A. (2012) Ultra high resolution linear ion trap Orbitrap mass spectrometer (Orbitrap Elite) facilitates top down LC MS/MS and versatile peptide fragmentation modes. Molecular & cellular proteomics : MCP 11, O111 013698.

38. Olsen, J. V., de Godoy, L. M., Li, G., Macek, B., Mortensen, P., Pesch, R., Makarov, A., Lange, O., Horning, S., and Mann, M. (2005) Parts per million mass accuracy on an Orbitrap mass spectrometer via lock mass injection into a C-trap. Molecular & cellular proteomics : MCP 4, 2010-2021.

39. Makarov, A., Denisov, E., Kholomeev, A., Balschun, W., Lange, O., Strupat, K., and Horning, S. (2006) Performance evaluation of a hybrid linear ion trap/orbitrap mass spectrometer. Analytical chemistry 78, 2113-2120.

40. McAlister, G. C., Berggren, W. T., Griep-Raming, J., Horning, S., Makarov, A., Phanstiel, D., Stafford, G., Swaney, D. L., Syka, J. E., Zabrouskov, V., and Coon, J. J. (2008) A proteomics grade electron transfer dissociation-enabled hybrid linear ion trap-orbitrap mass spectrometer. Journal of proteome research 7, 3127-3136.

41. Michalski, A., Damoc, E., Hauschild, J. P., Lange, O., Wieghaus, A., Makarov, A., Nagaraj, N., Cox, J., Mann, M., and Horning, S. (2011) Mass spectrometry-based proteomics using Q Exactive, a high-performance benchtop quadrupole Orbitrap mass spectrometer. Molecular & cellular proteomics : MCP 10, M111 011015.

42. Roepstorff, P., and Fohlman, J. (1984) Proposal for a common nomenclature for sequence ions in mass spectra of peptides. Biomed Mass Spectrom 11, 601.

43. McLuckey, S. A. (1992) Principles of Collisional Activation in Analytical Mass Spectrometry. Journal of the American Society for Mass Spectrometry 3, 599-614.

44. Paizs, B., and Suhai, S. (2005) Fragmentation pathways of protonated peptides. Mass Spectrom Rev 24, 508-548.

45. Boersema, P. J., Mohammed, S., and Heck, A. J. (2009) Phosphopeptide fragmentation and analysis by mass spectrometry. J Mass Spectrom 44, 861-878.

46. Lau, K. W., Hart, S. R., Lynch, J. A., Wong, S. C., Hubbard, S. J., and Gaskell, S. J. (2009) Observations on the detection of b- and y-type ions in the collisionally activated decomposition spectra of protonated peptides. Rapid Commun Mass Spectrom 23, 1508-1514.

47. Syka, J. E., Coon, J. J., Schroeder, M. J., Shabanowitz, J., and Hunt, D. F. (2004) Peptide and protein sequence analysis by electron transfer dissociation mass spectrometry. Proceedings of the National Academy of Sciences of the United States of America 101, 9528-9533.

48. Good, D. M., Wirtala, M., McAlister, G. C., and Coon, J. J. (2007) Performance characteristics of electron transfer dissociation mass spectrometry. Molecular & cellular proteomics : MCP 6, 1942-1951.

49. Pitteri, S. J., Chrisman, P. A., and McLuckey, S. A. (2005) Electron-transfer ion/ion reactions of doubly protonated peptides: effect of elevated bath gas temperature. Analytical chemistry 77, 5662-5669.

50. Horn, D. M., Ge, Y., and McLafferty, F. W. (2000) Activated ion electron capture dissociation for mass spectral sequencing of larger (42 kDa) proteins. Analytical chemistry 72, 4778-4784.

51. Swaney, D. L., McAlister, G. C., and Coon, J. J. (2008) Decision tree-driven tandem mass spectrom-

etry for shotgun proteomics. Nature methods 5, 959-964.

52. Frese, C. K., Altelaar, A. F. M., Hennrich, M. L., Nolting, D., Zeller, M., Griep-Raming, J., Heck, A. J. R., and Mohammed, S. (2011) Improved Peptide Identification by Targeted Fragmentation Using CID, HCD and ETD on an LTQ-Orbitrap Velos. J Proteome Res.

53. Wiesner, J., Premsler, T., and Sickmann, A. (2008) Application of electron transfer dissociation (ETD) for the analysis of posttranslational modifications. Proteomics 8, 4466-4483.

54. Mikesh, L. M., Ueberheide, B., Chi, A., Coon, J. J., Syka, J. E., Shabanowitz, J., and Hunt, D. F. (2006) The utility of ETD mass spectrometry in proteomic analysis. Biochim Biophys Acta 1764, 1811-1822.

55. Eng, J. K., McCormack, A. L., and Yates, J. R., 3rd (1994) An Approach to Correlate Tandem Mass Spectral Data of Peptides with Amino Acid Sequences in a Protein Database. J Am Soc Mass Spectrom 5, 976-989.

56. Perkins, D. N., Pappin, D. J., Creasy, D. M., and Cottrell, J. S. (1999) Probability-based protein identification by searching sequence databases using mass spectrometry data. Electrophoresis 20, 3551-3567.

57. Higgs, R. E., Knierman, M. D., Freeman, A. B., Gelbert, L. M., Patil, S. T., and Hale, J. E. (2007) Estimating the statistical significance of peptide identifications from shotgun proteomics experiments. Journal of proteome research 6, 1758-1767.

58. Weatherly, D. B., Atwood, J. A., 3rd, Minning, T. A., Cavola, C., Tarleton, R. L., and Orlando, R. (2005) A Heuristic method for assigning a false-discovery rate for protein identifications from Mascot database search results. Molecular & cellular proteomics : MCP 4, 762-772.

59. Kall, L., Canterbury, J. D., Weston, J., Noble, W. S., and MacCoss, M. J. (2007) Semi-supervised learning for peptide identification from shotgun proteomics datasets. Nature methods 4, 923-925.

60. Ong, S. E., and Mann, M. (2005) Mass spectrometry-based proteomics turns quantitative. Nat Chem Biol 1, 252-262.

61. Chelius, D., and Bondarenko, P. V. (2002) Quantitative profiling of proteins in complex mixtures using liquid chromatography and mass spectrometry. Journal of proteome research 1, 317-323.

62. Liu, H., Sadygov, R. G., and Yates, J. R., 3rd (2004) A model for random sampling and estimation of relative protein abundance in shotgun proteomics. Analytical chemistry 76, 4193-4201.

63. Old, W. M., Meyer-Arendt, K., Aveline-Wolf, L., Pierce, K. G., Mendoza, A., Sevinsky, J. R., Resing, K. A., and Ahn, N. G. (2005) Comparison of label-free methods for quantifying human proteins by shotgun proteomics. Molecular & cellular proteomics : MCP 4, 1487-1502.

64. Cox, J., and Mann, M. (2011) Quantitative, high-resolution proteomics for data-driven systems biology. Annu Rev Biochem 80, 273-299.

65. Christin, C., Hoefsloot, H. C., Smilde, A. K., Suits, F., Bischoff, R., and Horvatovich, P. L. (2010) Time alignment algorithms based on selected mass traces for complex LC-MS data. Journal of proteome research 9, 1483-1495.

66. Finney, G. L., Blackler, A. R., Hoopmann, M. R., Canterbury, J. D., Wu, C. C., and MacCoss, M. J. (2008) Label-free comparative analysis of proteomics mixtures using chromatographic alignment of high-resolution muLC-MS data. Analytical chemistry 80, 961-971.

67. de Leenheer, A. P., and Thienpont, L. M. (1992) Applications of isotope dilution-mass spectrometry in clinical chemistry, pharmcokinetics, and toxicology. Mass Spectrom Rev 11, 249-307.

68. Oda, Y., Huang, K., Cross, F. R., Cowburn, D., and Chait, B. T. (1999) Accurate quantitation of protein expression and site-specific phosphorylation. Proceedings of the National Academy of Sciences of

the United States of America 96, 6591-6596.

69. Ong, S. E., Blagoev, B., Kratchmarova, I., Kristensen, D. B., Steen, H., Pandey, A., and Mann, M. (2002) Stable isotope labeling by amino acids in cell culture, SILAC, as a simple and accurate approach to expression proteomics. Molecular & cellular proteomics : MCP 1, 376-386.

70. Gygi, S. P., Rist, B., Gerber, S. A., Turecek, F., Gelb, M. H., and Aebersold, R. (1999) Quantitative analysis of complex protein mixtures using isotope-coded affinity tags. Nature biotechnology 17, 994-999.

71. Hsu, J. L., Huang, S. Y., Chow, N. H., and Chen, S. H. (2003) Stable-isotope dimethyl labeling for quantitative proteomics. Analytical chemistry 75, 6843-6852.

72. Boersema, P. J., Raijmakers, R., Lemeer, S., Mohammed, S., and Heck, A. J. (2009) Multiplex peptide stable isotope dimethyl labeling for quantitative proteomics. Nat Protoc 4, 484-494.

73. Zhang, R., Sioma, C. S., Thompson, R. A., Xiong, L., and Regnier, F. E. (2002) Controlling deuterium isotope effects in comparative proteomics. Analytical chemistry 74, 3662-3669.

74. Altelaar, A. F., Frese, C. K., Preisinger, C., Hennrich, M. L., Schram, A. W., Timmers, H. T., Heck, A. J., and Mohammed, S. (2013) Benchmarking stable isotope labeling based quantitative proteomics. J Proteomics 88, 14-26.

75. Thompson, A., Schafer, J., Kuhn, K., Kienle, S., Schwarz, J., Schmidt, G., Neumann, T., Johnstone, R., Mohammed, A. K., and Hamon, C. (2003) Tandem mass tags: a novel quantification strategy for comparative analysis of complex protein mixtures by MS/MS. Analytical chemistry 75, 1895-1904.

76. Ross, P. L., Huang, Y. N., Marchese, J. N., Williamson, B., Parker, K., Hattan, S., Khainovski, N., Pillai, S., Dey, S., Daniels, S., Purkayastha, S., Juhasz, P., Martin, S., Bartlet-Jones, M., He, F., Jacobson, A., and Pappin, D. J. (2004) Multiplexed protein quantitation in Saccharomyces cerevisiae using amine-reactive isobaric tagging reagents. Molecular & cellular proteomics : MCP 3, 1154-1169.

77. Christoforou, A. L., and Lilley, K. S. (2012) Isobaric tagging approaches in quantitative proteomics: the ups and downs. Anal Bioanal Chem 404, 1029-1037.

78. Barr, J. R., Maggio, V. L., Patterson, D. G., Jr., Cooper, G. R., Henderson, L. O., Turner, W. E., Smith, S. J., Hannon, W. H., Needham, L. L., and Sampson, E. J. (1996) Isotope dilution--mass spectrometric quantification of specific proteins: model application with apolipoprotein A-I. Clin Chem 42, 1676-1682.

79. Gerber, S. A., Rush, J., Stemman, O., Kirschner, M. W., and Gygi, S. P. (2003) Absolute quantification of proteins and phosphoproteins from cell lysates by tandem MS. Proc Natl Acad Sci U S A 100, 6940-6945.

80. Pratt, J. M., Simpson, D. M., Doherty, M. K., Rivers, J., Gaskell, S. J., and Beynon, R. J. (2006) Multiplexed absolute quantification for proteomics using concatenated signature peptides encoded by QconCAT genes. Nat Protoc 1, 1029-1043.

81. Brun, V., Dupuis, A., Adrait, A., Marcellin, M., Thomas, D., Court, M., Vandenesch, F., and Garin, J. (2007) Isotope-labeled protein standards: toward absolute quantitative proteomics. Molecular & cellular proteomics : MCP 6, 2139-2149.

82. Kirkwood, T. B. (2005) Understanding the odd science of aging. Cell 120, 437-447.

83. Lindahl, T. (1993) Instability and decay of the primary structure of DNA. Nature 362, 709-715.

84. Sander, M., Cadet, J., Casciano, D. A., Galloway, S. M., Marnett, L. J., Novak, R. F., Pettit, S. D., Preston, R. J., Skare, J. A., Williams, G. M., Van Houten, B., and Gollapudi, B. B. (2005) Proceedings of a workshop on DNA adducts: biological significance and applications to risk assessment Washington, DC, April 13-14, 2004. Toxicol Appl Pharmacol 208, 1-20.

85. Hipkiss, A. R. (2006) Accumulation of altered proteins and ageing: causes and effects. Exp Gerontol 41, 464-473.

86. Ross, C. A., and Poirier, M. A. (2004) Protein aggregation and neurodegenerative disease. Nat Med 10 Suppl, S10-17.

87. Richter, C., Park, J. W., and Ames, B. N. (1988) Normal oxidative damage to mitochondrial and nuclear DNA is extensive. Proceedings of the National Academy of Sciences of the United States of America 85, 6465-6467.

88. Balaban, R. S., Nemoto, S., and Finkel, T. (2005) Mitochondria, oxidants, and aging. Cell 120, 483-495.

89. Lombard, D. B., Chua, K. F., Mostoslavsky, R., Franco, S., Gostissa, M., and Alt, F. W. (2005) DNA repair, genome stability, and aging. Cell 120, 497-512.

90. Friedberg, E. C. (2003) DNA damage and repair. Nature 421, 436-440.

91. Vijg, J. (2000) Somatic mutations and aging: a re-evaluation. Mutat Res 447, 117-135.

92. Garinis, G. A., van der Horst, G. T., Vijg, J., and Hoeijmakers, J. H. (2008) DNA damage and ageing: new-age ideas for an age-old problem. Nat Cell Biol 10, 1241-1247.

93. Hoeijmakers, J. H. (2009) DNA damage, aging, and cancer. N Engl J Med 361, 1475-1485.

94. Stein, G. H., Drullinger, L. F., Soulard, A., and Dulic, V. (1999) Differential roles for cyclin-dependent kinase inhibitors p21 and p16 in the mechanisms of senescence and differentiation in human fibroblasts. Mol Cell Biol 19, 2109-2117.

95. Vogelstein, B., Lane, D., and Levine, A. J. (2000) Surfing the p53 network. Nature 408, 307-310.

96. Sherr, C. J., and McCormick, F. (2002) The RB and p53 pathways in cancer. Cancer Cell 2, 103-112.

97. Barbacid, M. (1987) ras genes. Annu Rev Biochem 56, 779-827.

98. Davies, H., Bignell, G. R., Cox, C., Stephens, P., Edkins, S., Clegg, S., Teague, J., Woffendin, H., Garnett, M. J., Bottomley, W., Davis, N., Dicks, E., Ewing, R., Floyd, Y., Gray, K., Hall, S., Hawes, R., Hughes, J., Kosmidou, V., Menzies, A., Mould, C., Parker, A., Stevens, C., Watt, S., Hooper, S., Wilson, R., Jayatilake, H., Gusterson, B. A., Cooper, C., Shipley, J., Hargrave, D., Pritchard-Jones, K., Maitland, N., Chenevix-Trench, G., Riggins, G. J., Bigner, D. D., Palmieri, G., Cossu, A., Flanagan, A., Nicholson, A., Ho, J. W., Leung, S. Y., Yuen, S. T., Weber, B. L., Seigler, H. F., Darrow, T. L., Paterson, H., Marais, R., Marshall, C. J., Wooster, R., Stratton, M. R., and Futreal, P. A. (2002) Mutations of the BRAF gene in human cancer. Nature 417, 949-954.

99. Yuan, T. L., and Cantley, L. C. (2008) PI3K pathway alterations in cancer: variations on a theme. Oncogene 27, 5497-5510.

100. Harley, C. B., Futcher, A. B., and Greider, C. W. (1990) Telomeres shorten during ageing of human fibroblasts. Nature 345, 458-460.

101. von Zglinicki, T., Saretzki, G., Ladhoff, J., d'Adda di Fagagna, F., and Jackson, S. P. (2005) Human cell senescence as a DNA damage response. Mech Ageing Dev 126, 111-117.

102. Hayflick, L. (1965) The Limited in Vitro Lifetime of Human Diploid Cell Strains. Exp Cell Res 37, 614-636.

103. Dimri, G. P., Lee, X., Basile, G., Acosta, M., Scott, G., Roskelley, C., Medrano, E. E., Linskens, M., Rubelj, I., Pereira-Smith, O., and et al. (1995) A biomarker that identifies senescent human cells in culture and in aging skin in vivo. Proceedings of the National Academy of Sciences of the United States

of America 92, 9363-9367.

104. Acosta, J. C., O'Loghlen, A., Banito, A., Guijarro, M. V., Augert, A., Raguz, S., Fumagalli, M., Da Costa, M., Brown, C., Popov, N., Takatsu, Y., Melamed, J., d'Adda di Fagagna, F., Bernard, D., Hernando, E., and Gil, J. (2008) Chemokine signaling via the CXCR2 receptor reinforces senescence. Cell 133, 1006-1018.

105. Coppe, J. P., Patil, C. K., Rodier, F., Sun, Y., Munoz, D. P., Goldstein, J., Nelson, P. S., Desprez, P. Y., and Campisi, J. (2008) Senescence-associated secretory phenotypes reveal cell-nonautonomous functions of oncogenic RAS and the p53 tumor suppressor. PLoS Biol 6, 2853-2868.

106. Kuilman, T., Michaloglou, C., Vredeveld, L. C., Douma, S., van Doorn, R., Desmet, C. J., Aarden, L. A., Mooi, W. J., and Peeper, D. S. (2008) Oncogene-induced senescence relayed by an interleukin-dependent inflammatory network. Cell 133, 1019-1031.

107. Di Leonardo, A., Linke, S. P., Clarkin, K., and Wahl, G. M. (1994) DNA damage triggers a prolonged p53-dependent G1 arrest and long-term induction of Cip1 in normal human fibroblasts. Genes Dev 8, 2540-2551.

108. Ogryzko, V. V., Hirai, T. H., Russanova, V. R., Barbie, D. A., and Howard, B. H. (1996) Human fibroblast commitment to a senescence-like state in response to histone deacetylase inhibitors is cell cycle dependent. Mol Cell Biol 16, 5210-5218.

109. Serrano, M., Lin, A. W., McCurrach, M. E., Beach, D., and Lowe, S. W. (1997) Oncogenic ras provokes premature cell senescence associated with accumulation of p53 and p16INK4a. Cell 88, 593-602.

110. Zhu, J., Woods, D., McMahon, M., and Bishop, J. M. (1998) Senescence of human fibroblasts induced by oncogenic Raf. Genes Dev 12, 2997-3007.

111. Lin, A. W., Barradas, M., Stone, J. C., van Aelst, L., Serrano, M., and Lowe, S. W. (1998) Premature senescence involving p53 and p16 is activated in response to constitutive MEK/MAPK mitogenic signaling. Genes Dev 12, 3008-3019.

112. Michaloglou, C., Vredeveld, L. C., Soengas, M. S., Denoyelle, C., Kuilman, T., van der Horst, C. M., Majoor, D. M., Shay, J. W., Mooi, W. J., and Peeper, D. S. (2005) BRAFE600-associated senescence-like cell cycle arrest of human naevi. Nature 436, 720-724.

113. Yaswen, P., and Stampfer, M. R. (2002) Molecular changes accompanying senescence and immortalization of cultured human mammary epithelial cells. Int J Biochem Cell Biol 34, 1382-1394.

114. Parrinello, S., Samper, E., Krtolica, A., Goldstein, J., Melov, S., and Campisi, J. (2003) Oxygen sensitivity severely limits the replicative lifespan of murine fibroblasts. Nat Cell Biol 5, 741-747.

115. Kirkwood, T. B., and Austad, S. N. (2000) Why do we age? Nature 408, 233-238.

116. Zindy, F., Quelle, D. E., Roussel, M. F., and Sherr, C. J. (1997) Expression of the p16INK4a tumor suppressor versus other INK4 family members during mouse development and aging. Oncogene 15, 203-211.

117. Krishnamurthy, J., Ramsey, M. R., Ligon, K. L., Torrice, C., Koh, A., Bonner-Weir, S., and Sharpless, N. E. (2006) p16INK4a induces an age-dependent decline in islet regenerative potential. Nature 443, 453-457.

118. Molofsky, A. V., Slutsky, S. G., Joseph, N. M., He, S., Pardal, R., Krishnamurthy, J., Sharpless, N. E., and Morrison, S. J. (2006) Increasing p16INK4a expression decreases forebrain progenitors and neurogenesis during ageing. Nature 443, 448-452.

Spatio-temporal Analysis of Molecular Determinants of Neuronal Degeneration in the Aging Mouse Cerebellum

Erik L. de Graaf[1,2,3], Wilbert P. Vermeij[3,4,5], Monique C. de Waard[3,4,5], Yvonne Rijksen[3,4,5], Ingrid van der Pluijm[3,4,5], Casper C. Hoogenraad[6], Jan H. J. Hoeijmakers[3,4,5], A. F. Maarten Altelaar[1,2], and Albert J. R. Heck[1,2,3]

published in
Molecular & Cellular Proteomics 12, May 2013

[1] Biomolecular Mass Spectrometry and Proteomics, Bijvoet Center for Biomolecular Research and Utrecht Institute for Pharmaceutical Sciences, Utrecht University, Padualaan 8, 3584 CH Utrecht, The Netherlands

[2] Netherlands Proteomics Centre, Padualaan 8, 3584 CH Utrecht, The Netherlands

[3] Center for Biomedical Genetics

[4] Department of Genetics, Erasmus Medical Centre, 3015 GE Rotterdam, The Netherlands

[5] Cancer Genomics Centre

[6] Cell Biology, Faculty of Science, Utrecht University, 3584 CH Utrecht, The Netherlands

Abstract

The accumulation of cellular damage, including DNA damage, is hypothesized to contribute to ageing-related neurodegenerative changes. DNA excision repair cross-complementing group 1 (Ercc1) knockout mice represent an accepted model of neuronal ageing, showing gradual neurodegenerative changes, including loss of synaptic contacts and cell body shrinkage. Here, we used the Purkinje cell specific Ercc1 DNA-repair knockout mouse model to study ageing in the mouse Cerebellum. We performed an in-depth quantitative proteomics analysis, using stable isotope dimethyl labelling, to decipher changes in protein expression between the early (8 weeks), intermediate (16 weeks) and late (26 weeks) stages of the phenotypically ageing Ercc1 knock-out and healthy littermate control mice. The expression of over 5200 proteins from the cerebellum was compared quantitatively, whereby 79 proteins (i.e. 1.5%) were found to be substantially regulated during ageing. These molecular markers of the early ageing onset did nearly all belong to a strongly interconnected network involved in excitatory synaptic signalling. Using immunohistological staining we obtained temporal and spatial profiles of these markers confirming not only the proteomics data, but in addition revealed how the change in protein expression correlates to synaptic changes in the cerebellum. In summary, this study provides a highly comprehensive spatial- and temporal view of the dynamic changes in the cerebellum and Purkinje cell signalling in particular, indicating that synapse signalling is one of the first processes to be affected in this premature aging model, leading to neuron morphological changes, neuron degeneration, inflammation and ultimately behaviour disorders.

Introduction

A link between DNA damage and the process of ageing has been firmly established.(1, 2) The brain in particular is a vulnerable organ that is plagued by various neurodegenerative disorders that have been related to ageing, i.e. Alzheimer's and Parkinson's disease. The study of the early onset of age-related neurodegenerative diseases is challenging, since there are not many confident early molecular determinants that predict their development. Therefore, progeroid syndromes (showing premature ageing) are often used as a model for segmental ageing as they show consistent and predictive elements of the ageing phenotype (e.g.: cessation of growth and development, hearing loss, severe and progressive neuron dysfunction).(1, 3) These accelerated ageing syndromes have in common that they bear defects in one or multiple proteins involved in DNA damage repair mechanisms.

A well-established progeroid mouse model is the excision repair cross-complementing group 1 (Ercc1) gene knockout.(4, 5) The Ercc1/XPF complex acts as a nuclease in the nucleotide excision repair (NER) pathway and has an important function in both global genome and transcription coupled DNA damage repair. Besides its role in NER, Ercc1 is also involved in interstrand cross-link repair and oxidative damage repair. The accumulation of DNA damage following Ercc1 gene inactivation results in the progeroid phenotype.(2, 6) Recently, it was shown that global and neuron specific Ercc1 mutant mice showed age-related neuronal changes in the spinal cord and the hippocampus.(7, 8) Here we set out to monitor molecular changes in the cerebellum of Purkinje cell specific Ercc1 knockout mice, by quantitative proteomics. The advantage of a cell-type specific knockout, over a global knockout, is that the organism is not compromised by secondary side effects from other organs (i.e. liver, kidney) or general tissue stress responses. Therefore, the changes observed in animal behaviour and protein expression levels can be attributed to the specific cell type that is affected, and the surroundings it directly interacts with. Purkinje cells are neurons that act as signal converters in the cerebellar cortex and are essential for the proper functioning of the cerebellum. Their main role is to convert input signals from many different neurons into the sole output signal of the cerebellum, thereby controlling motor coordination, learning and cognitive functions.

Up till now, most neuroproteomics studies have been directed at general protein composition of the synapse, with the emphasis on the synaptosomes(9), synaptic membranes(10, 11) or post- and pre-synaptic densities(8, 12) by using subcellular fractionation strategies on the whole brain or sub compartments. Here, we compared quantitatively the proteome from intact cerebella from Purkinje neuron specific DNA-repair KO mice with cerebella from control mice at the age of 8, 16 and 26 weeks, to monitor ageing-related molecular changes in the cerebellum over time. Using this mouse model coupled with stable isotope dimethyl labelling and high-resolution nanoLC-MS, we were able to quantify the proteins present in ex vivo cerebellum Purkinje synapses without introducing any bias induced by subcellular fractionation methods or artificial cell culture systems. A small set of consistently down-regulated proteins originated from a strongly interconnected network known to be involved in synaptic signalling. Furthermore, immunohistochemistry experiments were performed to confirm mass spectrometry data and to assess protein localization as well as

Purkinje cell and cerebellum morphology.

Experimental Procedures

Mouse models - The generation and characterization of Ercc1$^{+/-}$ mice has been described previously(5). To achieve Purkinje cell-specific Ercc1 gene inactivation, a transgenic line with Cre recombinase under the control of the L7/pcp2 promoter was used(13). Female L7/pcp2-Cre^{+} mice were crossed with male Ercc1$^{+/-}$ mice (both in the C57BL6J background). Female Ercc1$^{+/-}$ L7/pcp2-Cre^{+} mice, obtained from these breedings, were crossed with male Ercc1f/f FVB mice, to yield hybrid Ercc1$^{f/-}$ L7/pcp2-Cre^{+} mice. The homozygous Ercc1$^{f/f}$ (floxed Ercc1; with loxP sites inserted into its Ercc1 gene) mice (14) were kindly provided by Dr. D.W. Melton (University of Edinburgh, UK) and backcrossed for 10 generations to obtain a pure FVB background. Ercc1$^{f/-}$ L7/pcp2-Cre^{+} mice (in the F1 C57BL6J/FVB hybrid background) are heterozygous for Ercc1 in their entire body, except for the Purkinje cells in the cerebellum, which are homozygous for Ercc1 after Cre excision of the floxed allele. These mice will be referred to as KO in the remainder of the study. As controls, we used Ercc1$^{f/+}$ L7/pcp2-Cre^{+} littermates (referred to as CON or control), which are wild type in their entire body, except for the Purkinje cells in the cerebellum, which are heterozygous. All animals used in the studies described in this paper were of the same F1 C57BL6J/FVB hybrid background and had ad libitum access to standard mouse food (CRM pellets, SDS BP Nutrition Ltd; gross energy content 18.36 kJ/g dry mass, digestible energy 13.4 kJ/g) and water. Mice were weighted and visually inspected weekly, and were scored for gross morphological and motor abnormalities. Experiments were performed in accordance with the Principles of Laboratory Animal Care (NIH publication no. 86-23) and with the guidelines approved by the Erasmus University Animal Care Committee.

Sample preparation - Frozen female cerebella were lysed by sonication in lysis buffer (8M Urea in 50 mM ammonium bicarbonate, 1 tablet Complete mini EDTA-free Cocktail (Roche) and 1 tablet PhosSTOP phosphatase inhibitor Cocktail (Roche)). After centrifugation (20 000x g 30min at 4 °C), the supernatant was assayed for protein content using the BCA kit standard procedure (Pierce) (~4 mg per condition). Protein reduction and alkylation were performed using final concentrations of 2 mM dithiothreitol and 4 mM iodoacetamide, respectively. A first enzymatic digestion step was performed in 8 M urea lysis buffer using Lys-C at 37 °C for 4 h (enzyme/substrate ratio 1:75). The second digestion was performed overnight (37 °C) with trypsin (enzyme/substrate ratio 1:100) in 2 M Urea. Resulting peptides were chemically labelled using stable isotope dimethyl labelling as described before. (15) In the first experiment CON tissue was labelled "Intermediate" and Ercc1 PkJ KO was labelled "Heavy". In the second replicate experiment the Intermediate and Heavy labels were swapped. In both experiments a 1:1:1 mixture of CON tissue lysates of 8, 16 & 26 weeks were labelled "Light" as an internal control. The labelling efficiency for all labels was higher than 96%. Next, an aliquot of each label was measured on a regular LC-MS/MS run and samples were mixed 1:1:1 (L:I:H) based on their peptide intensities. This was found to result in a more precise ratio than using the total protein amounts as determined by a BCA assay. After mixing, peptides were dried to completion and subsequently reconstituted in

10% formic acid prior to fractionation using strong-cation exchange (SCX) as described previously.(16) Briefly, SCX was performed using a Zorbax BioSCX-Series II column (0.8-mm inner diameter x 50-mm length, 3.5 μm). SCX Solvent A consisted of 0.05% formic acid in 20% acetonitrile and solvent B of 0.05% formic acid, 0.5 M NaCl in 20% acetonitrile. The SCX gradient was as followed: 0-0.01 min (0-2% B); 0.01-8.01 min (2-3% B); 8.01-14.01 min (3-8% B); 14.01-28 min (8-20% B); 28-38 min (20-40% B); 38-48 min (40-90% B); 48-54 min (90% B); 54-60 min (0% B). After injection of 200 μg labelled tissue lysate a total of 45 SCX fractions were collected per cell lysate and dried in a vacuum centrifuge.

Liquid Chromatography and Mass Spectrometry **-** After reconstituting the SCX fractions containing doubly and triply charged peptides (approx. 15 fractions each SCX) in 10% formic acid, peptides were analysed using nano flow reverse phase liquid chromatography coupled to a LTQ-Orbitrap Velos mass spectrometer (Thermo, San Jose, CA). Depending on the SCX UV-trace, 1-10% of each fraction was injected. Densely populated 2+ fractions were injected twice to minimize undersampling of the mass spectrometer. Peptides were trapped on a trap column (ReproSil-Pur C18-AQ, 3 μm, (Dr. Maisch GmbH, Ammerbuch, Germany); 20 mm x 100 μm ID, packed in house) at 5 μl/min in 100% solvent A (0.1 M acetic acid in water). Next, peptides were eluted from the trap column onto the analytical column (ReproSil-Pur C18-AQ, 3 μm (Dr. Maisch GmbH, Ammerbuch, Germany); 40 cm x 50-μm ID, packed in house) at ˜100 nl/min in 1h, 2h or 3 h gradients from 10 to 50% solvent B (0.1 M acetic acid in 8:2 (v/v) acetonitrile/water). Nanospray was achieved with an in-house pulled and gold-coated fused silica capillary (360 μm OD; 20 μm ID; 10 μm tip ID) and an applied voltage of 1.7 kV. The mass spectrometer was configured to perform a FT survey scan from 350 to 1500 m/z (resolution 30 000) followed by HCD fragmentation of the 10 most intense peaks (35% normalized CE at a target value of 50 000 ions, resolution 7500).

Data analysis **-** All MS data were processed with Proteome Discover (version 1.3, Thermo Scientific) using a standardized workflow. Peak lists, generated in Proteome Discover, were searched against a concatenated forward-decoy Swissprot (v56.2, taxonomy Mus Musculus, 32 402 protein entries) database, supplemented with frequently observed contaminants, using Mascot (version 2.3.02 Matrix Science, UK). The following search parameters were used: 50 ppm precursor mass tolerance, 0.05 Da fragment ion tolerance, trypsin cleavage with maximum of 2 miscleavages, cysteine carbamidomethyl static modification and methionine oxidation and dimethyl labelling (L,I,H) of lysine residues and the peptide N-termini as dynamic modification. Triplex dimethyl labelling was used as a quantification method, with a mass precision of 2 ppm for consecutive precursor mass scans and a normalisation on the mean peptide ratio. A RT tolerance of 0.5 min was used to account for the effect of deuterium on the retention time. To filter for high quality data and to control the false discovery rate on identifications, only the PSMs adhering to the following criteria were kept for analysis: minimal Mascot score of 20, minimal peptide length of 7, only unique rank 1 peptides and a mass deviance of 20 ppm and 10 ppm for experiment 1 and experiment 2 respectively. As a result we obtained a peptide FDR of 0.89% for experiment 1 and 0.79% for experiment 2. To distinguish between proteins with high similarities, only unique peptides were considered for protein identification and quantification. The estimation of the false discovery rate in quantification was calculated using Significance Analysis of Microarrays

(SAM), which is increasingly used by the proteomics community because conventional t-testing tends to result in an underestimation of false-positives.(17-19) The obtained SAM q-value is the empirically estimated false discovery rate (FDR) per individual protein and should therefore not be interpreted as a global FDR or p-value. The q-values were obtained using the multi experiment viewer (MeV v4.6.2) with the method proposed by Tusher et al. (20) using 1000 permutations. Proteins with a minimum rounded average 2.0-fold change and a q-value below 0.1 were considered significantly regulated.

Antibodies - Primary antibodies (supplier; dilutions) used in this study were as follows: rabbit anti-Calbindin (Calb1) (Swant; 1:10,000); rabbit anti-cleaved Caspase 3 (CASP-3) (Asp175; Cell Signalling Technology; 1:500); rabbit anti-Connexin43 (CX43) (Abcam; 1:1,000); rabbit anti-GFAP (DAKO; 1:8,000); rabbit anti-GluRδ1/2 (Chemicon; 1:1,000); rabbit anti-HtrA1 (Abcam; 1:100); rabbit anti-Itpr1 (Abcam; 1:2,500); rat anti-Mac2 (Cedarlane; 1:1,000); rabbit anti-PKCγ (Santa Cruz; 1:2,000); goat anti-Prkg1 (cGK1) (Santa Cruz; 1:500). For avidin–biotin–peroxidase immunocytochemistry, biotinylated secondary antibodies (Vector Laboratories; 1:200) were used.

Immunohistochemical procedures - Mice were anesthetized with pentobarbital and perfused transcardially with 4% paraformaldehyde. The brain was carefully dissected out and postfixed overnight in 4% paraformaldehyde. Routinely, brain tissue was embedded in 10% gelatine blocks, rapidly frozen, and sectioned at 40 μm with a freezing microtome or stored at -80°C until use. (21) Control and KO tissue was embedded together, in one gelatine block per time point, to avoid fluctuations in antibody affinity. In order to quantitatively asses protein regulation using IHC results, all immunohistochemical procedures were performed simultaneously for each antibody. Sections were processed, free-floating, using a standard avidin– biotin–immunoperoxidase complex method (ABC, Vector Laboratories, Burlingame, CA) with diaminobenzidine (0.05%; Sigma) as the chromogen as described previously (21). In addition, a selected number of frozen gelatin sections was collected in 4% paraformaldehyde, and processed with a modified Gallyas silver impregnation procedure of Nadler and Evenson (22) that selectively labels degenerating neurons and their processes. Immunoperoxidase-stained sections were analysed and photographed using an Olympus BX40 microscope and a ColorView digital camera. Quantification was performed using MetaMorph Image Analysis software (Molecular Devices, Sunnyvale, CA). Background intensity was subtracted per antibody and the average intensity of the 8 week control sample was set to 1. Statistical significance was calculated by Anova with a Bonferroni post-test to compare the replicate means by age using Prism (GraphPad Software, Inc., La Jolla, CA).

Results

To decipher the molecular changes that underlie the observed phenotype, an in-depth quantitative proteomics screen was performed with biological duplicates (and technical duplicates) on cerebella from control and Purkinje cell specific $Ercc1^{f/-}$ knockout mice. The knockout mice have the same lifespan as compared to wild type littermates and show normal growth development. At the age of 8 weeks, the Purkinje specific Ercc1 KO mice showed

no aberrant behaviour and were phenotypically undistinguishable from littermate controls. Around 16 weeks the normal weight gain of these animals starts to level off (Figure 1A). Furthermore, first signs of imbalance were observed at this stage and developed into severe motoric dysfunctioning at 26 weeks. The observed motor abnormalities are similar to the Ercc1$^{\Delta/-}$ full body mutant mice(23) and were not present in any of the control animals. Comprehensive behavioural studies of the Purkinje specific Ercc1 KO mice further showed a clear motoric function decline and lack of capacity in motoric learning at 26 weeks (van der Vaart et al., in preparation). Based on the phenotypical behaviour three time points for tissue analysis were chosen: an early time-point at 8 weeks to search for molecular markers before the onset of the observed phenotype; an intermediate time-point at 16 weeks; and a late time-point at 26 weeks, where major molecular changes are anticipated.

After dissection of the cerebella, the tissue was processed as described in the workflow in figure 1B. In order to compare protein levels from control and KO mice, the generated peptides were labelled with medium and heavy dimethyl stable isotope labels respectively.(15) As an internal control a mixture of all control samples was labelled with a light version of dimethyl. Next the internal control sample was mixed with equal protein amounts of control and KO samples per time point (1:1:1 ratio). To decrease sample complexity, per time point, each peptide mixture was fractionated using strong cati-

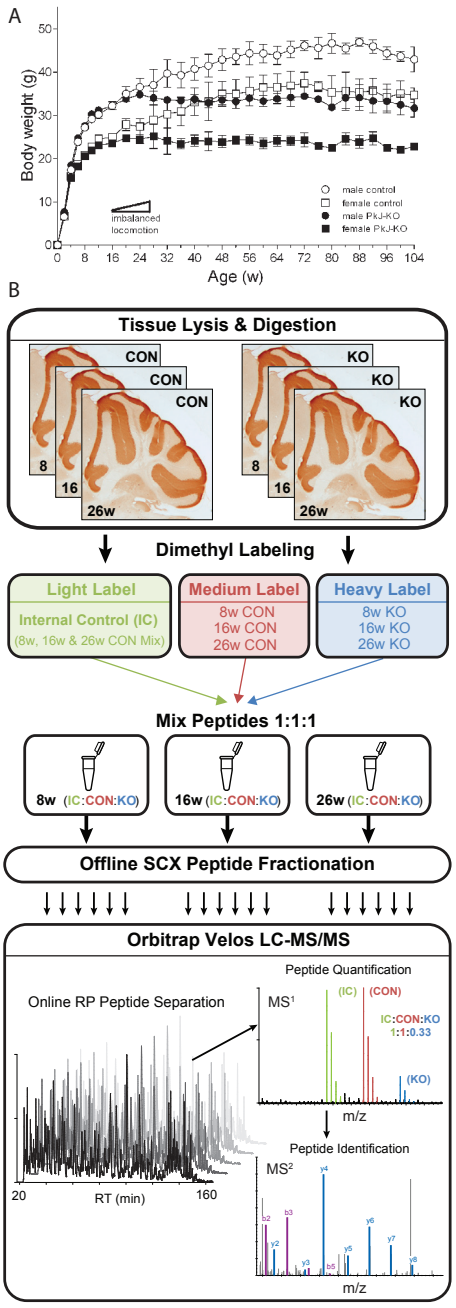

Figure 1: *Phenotypical weight and motoric function data of Purkinje cell specific Ercc1 KO mice (A). Proteomics workflow scheme (B). After cerebellum dissection from mice of different genetic background and age, tissues were lysed in 8 M Urea and digested with Lys-C and trypsin. Internal control, control and knockout peptides were labelled using stable isotope dimethyl labelling and mixed in three groups corresponding to the age of the sacrificed mice. In the replicate experiment heavy and medium labels were swapped. Sample complexity was reduced by generating 14 SCX fractions from each time point prior to 3h LC-MS/MS analysis.*

on exchange (SCX) in which the 14 main fractions were analysed by reverse phase LC-MS using an Orbitrap Velos with HCD fragmentation. Peptide identities and ratios between control (medium) and KO (heavy) were analysed using Proteome Discoverer software. In the biological replicate experiment the same procedure was followed, with the exception of a stable isotope label-swap between control and KO samples. In total 5254 proteins could be identified and quantified from all cerebella, over all time-points (Supplementary Tables S1 and S2).

To compare protein expression levels between control and KO replicates, their logarithmic ratios were plotted against each other (Figure 2). Only proteins with an average regulation from both replicates of at least 2-fold up or down, and a SAM q-value below 0.1, are considered to show a significant expression change. When comparing the KO/CON ratios of two replicates the left top and right bottom quadrant show the proteins that are consistently regulated in both replicates. Proteins located in the remaining two quadrants show an anticorrelation between both replicates and were found to be originating from contaminants and low confidence, single peptide quantifications.

In vivo proteomic screenings of progeroid cerebella

At 8 weeks of age (postnatal day (P) 56), the cerebella of control and KO mice are similar in morphology (Supplementary Figure S1A,B). When looking at the protein levels a similar trend is observed. The log ratio between Ercc1 Purkinje KO and control are close to zero for both replicates, as can be seen by the circular distribution of the replicate ratios (Figure 2A). The proteins that are consistently regulated in both replicates (Clic6, E-NPP2 and Lat2) were not significantly changed in the later time-points and did not show any functional relationship; therefore these proteins were discarded from further analysis.

At 16 weeks (P112) the KO cerebella size is still comparable to control however the individual protein ratios are already starting to show significant changes between control and KO tissue. While most of the protein levels remain unchanged the data distribution is starting to stretch out from the circular distribution into the two quadrants representing protein regulation (Figure 2B). This trend is not observed when comparing the protein expression of control tissue of 8 and 16 weeks to 26 week old control cerebella, proving the observed trend in KO is caused solely by the silencing of the Ercc1 gene and not for temporal/developmental reasons (Supplementary Figure S2A-B). Among the consistently regulated proteins at 16 weeks, 27 proteins were found to be up or down-regulated. The identity of the regulated proteins reveal a down-regulation of proteins involved in synaptic signalling (GluRδ2, Delphilin, IP3R1) and signal transduction (cGK1, PKCγ, Ahrgef33, RGS8, TN-C, Ppp1r16b). Furthermore the increase in astrocyte marker GFAP and complement factor C1qb indicates neuronal damage and an inflammatory response.

After 26 weeks (P182) a specific reduction in cerebella size is observed in KO mice as compared to control brains whereas all other parts of the brain were unaffected (Supplementary Figure S1C,D). Using the dimethyl labelling strategy equal amounts of proteins are compared, taking into account the variation of cerebellum size. This results in ~98% of all

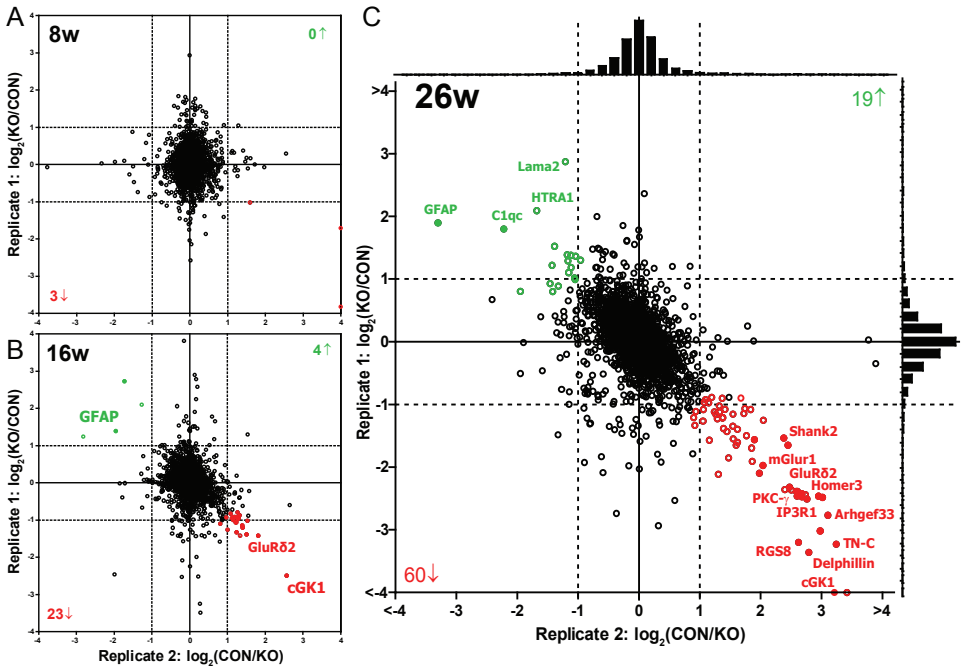

Figure 2: *Knockout vs. control protein ratio plotted for replicate 1 vs. replicate 2. Logarithmic protein ratios were plotted for 8 weeks (A), 16 weeks (B) and 26 weeks (C) old mice. Over 3000 proteins were quantified per time point. Proteins regulated after both 16 and 26 weeks are indicated by closed coloured circles. Proteins with an average 2-fold regulation and a SAM q-value below 0.1 were considered significant.*

proteins showing <2-fold expression changes, albeit with a clear difference in overall distribution of protein levels between control and knockout tissue (Figure 2C). Although such a small number of changes in protein expression (<~2%) is in line with a study by Walther et al.(24) in which no extensive proteome changes were observed in cerebella of healthy 5 and 26 months old mice, it shows a remarkable robustness in protein expression levels as our progeroid model displays aberrant neurological behaviour and visible brain morphological changes at 26 weeks. Interestingly, the proteins found to be down-regulated after 16 weeks were further down-regulated after 26 weeks, indicating a reproducible and predictable model system (Figure 2C). In addition to the proteins regulated after 16 weeks a new set of regulated proteins emerged in KO cerebella after 26 weeks. In total 60 proteins were down-regulated and 19 proteins were found to be up-regulated using our strict criteria.

Molecular function of regulated proteins

The proteins found to be up-regulated in KO tissues contained the astrocyte marker glial fibrillary acidic protein (GFAP), a gap junction glial marker protein Connexin43 (Cx43), macrophage marker Mac-2, complement factor C1qC as well as Metallothionein 1 (Mt-1). The up-regulation of these classes of proteins marks an increased recruitment of astrocytes and/or other glial cells indicating an increase in neuronal damage and inflammation in the

cerebellum of KO mice. This is in line with previous studies that have shown that ageing and the lack of efficient DNA repair can trigger neurodegenerative and inflammatory processes.(23, 25-27) Other proteins found to be up-regulated are extracellular proteins that are known to interact with or modify the extracellular environment (HTRA1, Lama2) as well as cell adhesion molecules (PECAM-1, MCAM). Caspase 3 was also found to be up-regulated, however the ratio between KO and control experiment could not be calculated in one replicate due to low protein levels in control tissue. The increase of Caspase enzymes could indicate an elevated level of apoptotic cells, due to excessive DNA damage, or locally activated apoptotic synaptic cascades influencing synaptic plasticity(28).

The group of down-regulated proteins contained synaptic scaffold proteins (i.e. Homer-3, Shank2) neurotransmitter receptors (i.e. mGluR1, GluRδ2, GABABR1 & GABABR2) ion channels/transporters (i.e. SERCA3, TrpC3, Kvβ1, Cavα2δ2) and signal transduction enzymes (i.e. cGK1, PKC-γ, CaMK-IIα, IP3KA, Pde5a). Many of these proteins are known to be present in synapses of different types of neurons throughout the brain indicating a strong effect of DNA-damage on synaptic maintenance. Interestingly, it has been described that reduced synaptic contact and activity is one of the early hallmarks of neurodegenerative diseases.(26, 29) Therefore, the down-regulated proteins found in this study can be considered as a set of proteins that are specifically expressed and active in Purkinje pre- and postsynaptic areas and potentially define the Purkinje synapses. A complete list of all regulated proteins including Purkinje cell markers L7 (Purkinje cell protein 2), Spot 35 (Calbindin) and PEP-19 (Purkinje cell protein 4) and their SAM q-value, can be found in Table 1.

In addition to proteins with known signalling transduction routes also signalling molecules without known interaction partners like Arhgef33, RGS8 and GNG13 were found to be regulated. These proteins are predicted to be involved in G-protein coupled signalling based on their homology to known protein domains. However their exact function and interaction partners are not known. In the proteomics data, these proteins are strongly co-regulated with other synaptic proteins, suggesting a functional role in the Purkinje synaptic signal transduction cascades. A possible role for these proteins could involve binding to similarly regulated G-protein coupled receptors (i.e. mGluR1) and their scaffold proteins (i.e. Homer-3) thereby activating other ion channels (i.e. IP3R1&2, TrpC3) or signal transduction cascades (i.e. cGK1, CaMK-IIα) that are part of the same scaffold complex or in the nearby vicinity.

Immunohistochemistry validation and follow-up

In order to verify our proteomics results, a set of proteins was analysed by immunohistochemistry (IHC) staining in sagittal sections of mice brains from the three different time points. The great advantage over commonly used verification methods by western blot is the ability to localize protein changes in different substructures of the cerebellum. An additional benefit is that cell type specificity and even subcellular (i.e. axon, cell body or dendrite) regulations can be observed. The spatial semi-quantitative IHC data of the validated proteins is plotted in supplementary figure S4, corroborating our proteomics based protein

quantification.

Table 1: Proteins differentially expressed in mouse cerebellum upon Purkinje cell DNA damage

Accession	Protein Name	Description	8w FC	16w FC	26w FC	q-value
Classical purkinje cell markers						
P12658	Spot 35 / Calbindin	Calbindin	-0.1	-0.9	-2.4	0.000
P12660	L7 / Pcp2	Purkinje cell protein 2	-0.1	-1.1	-2.6	0.000
P63054	PEP-19 / Pcp4	Purkinje cell protein 4	0.0	-0.8	-1.6	0.000
Transmembrane proteins						
P97772	mGluR1	Metabotropic glutamate receptor 1	-0.1	-1.0	-2.0	0.000
Q61625	GluRδ2	Glutamate receptor delta-2 subunit	-0.2	-1.5	-2.4	0.000
P23818	GluR-1	Glutamate receptor 1, AMPA 1	0.3	-0.4	-1.1	0.027
Q9WV18	GABABR1	Gamma-aminobutyric acid type B receptor subunit 1	-0.1	-0.6	-1.4	0.027
Q80T41	GABABR2	Gamma-aminobutyric acid type B receptor subunit 2	-0.1	-0.8	-1.3	0.017
P11881	IP3R1	Inositol 1,4,5-trisphosphate receptor type 1	-0.1	-1.1	-2.6	0.000
Q9Z329	IP3R2	Inositol 1,4,5-trisphosphate receptor type 2	0.0	-0.9	-2.6	0.000
O55143	SERCA2	Sarcoplasmic/endoplasmic reticulum calcium ATPase 2	0.0	-0.5	-1.0	0.016
Q64518	SERCA3	Sarcoplasmic/endoplasmic reticulum calcium ATPase 3	-0.1	-1.0	-2.8	0.000
Q9QZC1	TrpC3	Short transient receptor potential channel 3	0.1	-0.9	-4.5	0.019
P63143	K_vβ1	Voltage-gated potassium channel subunit beta-1	-0.1	-0.7	-1.2	0.016
Q6PHS9	Ca_vα2δ2	Voltage-gated calcium channel subunit alpha-2/delta-2	-0.3	-0.8	-1.2	0.017
O35544	Eaat4	Excitatory amino acid transporter 4	-0.2	-0.9	-1.9	0.000
Q5DTL9	Slc4a10	Sodium-driven chloride bicarbonate exchanger	-0.2	-0.6	-1.2	0.050
Q921R8	Slc41a3	Solute carrier family 41 member 3	0.2	-0.9	-1.4	0.084
Q6WQJ1	DGL-α	Sn1-specific diacylglycerol lipase alpha	-0.1	-1.1	-2.5	ND*
Q8JZM4	DNER	Delta and Notch-like epidermal growth factor-related receptor	-0.2	-0.9	-1.5	0.000
Q8C5W0	Calmin	Calmin / Calponin-like transmembrane domain protein	0.0	-1.0	-1.7	0.000
Q6NXK7	DPP X	Dipeptidyl peptidase X	-0.1	-0.6	-1.5	0.000
Q62092	Nsg1	Neuron-specific protein family member 1	-0.3	-0.5	-1.2	0.017
Q3UH99	Shisa6	UPF0626 protein B	-0.2	-1.0	-1.1	0.081
Q8K0T0	Rtn1	Reticulon-1	0.0	-0.5	-1.0	0.061
P41731	Cd63	CD63 antigen	0.1	0.4	1.1	0.082
P41233	ABC-1	ATP-binding cassette sub-family A member 1	ND*	0.5	1.2	0.082
Q08481	PECAM-1	Platelet endothelial cell adhesion molecule (CD31)	ND*	ND*	1.1	0.082
Q8R2Y2	MCAM	Cell surface glycoprotein MUC18 / Melanoma cell adhesion molecule / CD146	0.4	0.8	1.4	0.082
P23242	Cx43	Gap junction alpha-1 protein / Connexin-43	0.4	0.5	1.3	0.082
Extracellular proteins						
Q80YX1	TN-C	Tenascin-C / Hexabrachion	0.0	-1.6	-3.2	0.000
P02802	MT-1	Metallothionein-1	-0.1	0.5	1.1	0.082
P03995	GFAP	Glial fibrillary acidic protein	0.0	1.7	2.6	0.082
Q9R118	HTRA1	Serine protease HTRA1	-0.1	0.3	1.9	0.082
Q60675	Lama2	Laminin subunit alpha-2 / Laminin M	0.7	0.6	2.0	0.082

Table 1 Continued

Accession	Protein Name	Description	8w FC	16w FC	26w FC	q-value
Scaffold proteins						
Q0QWG9	Delphilin	Delphilin / Glutamate receptor, ionotropic, delta 2-interacting protein 1	-0.1	-1.3	-3.1	0.000
Q6WVG3	Kctd12	BTB/POZ domain-containing protein KCTD12	0.0	-0.6	-1.1	0.017
Q80Z38	Shank2	SH3 and multiple ankyrin repeat domains protein 2	0.0	-1.1	-2.0	0.029
Q99JP6	Homer-3	Homer protein homolog 3	-0.1	-1.1	-2.5	0.000
O54931	Akap2	A-kinase anchor protein 2	-0.2	-0.7	-1.4	0.000
Signaling proteins						
P11798	CaMK-IIα	Calcium/calmodulin-dependent protein kinase type II alpha chain	0.0	-0.9	-1.7	0.071
P63318	PKCγ	Protein kinase C gamma type	-0.1	-1.3	-2.5	0.000
P0C605	cGK1	cGMP-dependent protein kinase 1, alpha isozyme	-0.4	-2.5	-3.3	ND*
Q8CG03	Pde5a	cGMP-specific 3',5'-cyclic phosphodiesterase	-0.4	-1.0	-2.7	0.000
Q8R071	IP3KA	Inositol-trisphosphate 3-kinase A	0.0	-1.1	-3.0	0.000
Q9EPW0	Inpp4a	Type I inositol-3,4-bisphosphate 4-phosphatase	0.0	-0.6	-1.3	0.026
Q91WG7	DGK-γ	Diacylglycerol kinase gamma	0.0	-0.6	-1.0	2.957
Q80UP3	DGK-ζ	Diacylglycerol kinase zeta	0.1	-0.9	-1.7	0.079
Q8BW86	Arhgef33	Rho guanine nucleotide exchange factor 33	0.0	-1.3	-2.9	0.000
Q8BXT1	RGS8	Regulator of G-protein signaling 8	-0.1	-1.3	-2.9	0.000
Q9JMF3	GNG13	Guanine nucleotide-binding protein G(I)/G(S)/G(O) subunit gamma-13	-0.1	-1.1	-2.1	0.016
P57759	Erp29	Endoplasmic reticulum protein ERp29	0.0	-0.6	-1.4	0.000
P68510	14-3-3η / Ywhah	14-3-3 protein eta	-0.1	-0.4	-1.0	0.050
Cytoskeletal proteins						
Q9JJZ2	Tuba8	Tubulin alpha-8 chain	-0.2	-0.8	-1.8	0.000
P97434	RIP3	Myosin phosphatase Rho-interacting protein	-0.3	-0.7	-1.6	0.000
Q05BC3	EMAP-1	Echinoderm microtubule-associated protein-like 1	0.0	-0.4	-1.3	0.000
Q3UJU9	RMD-3	Regulator of microtubule dynamics protein 3	0.1	-0.6	-1.6	0.000
Q8R1S4	Mtss1	Metastasis suppressor protein 1	-0.1	-0.9	-2.4	0.000
Cytosolic and remaining proteins						
P28651	CARP	Carbonic anhydrase-related protein	-0.1	-1.1	-2.6	0.000
Q3TGF2	Fam107b	Protein FAM107B	-0.1	-1.1	-2.0	0.000
Q99K30	EPS8-LP2	Epidermal growth factor receptor kinase substrate 8-like protein 2	-0.1	-1.1	-2.5	0.000
Q9ERG2	Striatin-3	Striatin-3 / Cell cycle autoantigen SG2NA	0.2	-0.7	-1.6	0.000
P62748	Hpcal1	Hippocalcin-like protein 1 / Neural visinin-like protein 3	0.0	-0.7	-1.3	0.017
Q61704	ITI-HC3	Inter-alpha-trypsin inhibitor heavy chain H3	0.1	-0.6	-1.7	0.020
Q9ERQ8	CA-VII	Carbonic anhydrase 7	0.2	-0.8	-1.4	0.028
Q8BKX1	BAI-ap2	Brain-specific angiogenesis inhibitor 1-associated protein 2 / Insulin receptor substrate p53	-0.1	-0.8	-1.2	0.030
Q1RLL3	Copine-9	Copine-9	-0.1	-0.7	-1.4	0.081
Q9CRD4	Dbndd2	Dysbindin domain-containing protein 2	0.2	-0.9	-1.5	0.082
Q80ZW2	Them6	UPF0670 protein THEM6 homolog	0.0	0.0	-1.1	0.084
Q3UVC0	Ksr2	Kinase suppressor of Ras 2	-0.4	-0.9	-1.1	0.086
Q8BIZ1	AIDA-1	Amyloid-beta protein intracellular domain-associated protein 1	0.0	-0.7	-1.3	0.097
P28828	R-PTPμ	Receptor-type tyrosine-protein phosphatase mu	0.2	-0.2	-1.1	0.078

Table 1 Continued

Accession	Protein Name	Description	8w FC	16w FC	26w FC	q-value
Q02105	C1qc	Complement C1q subcomponent subunit C	ND*	ND*	2.0	0.082
P20152	Vim	Vimentin	0.1	0.2	1.1	0.082
P20060	Hexb	Beta-hexosaminidase subunit beta	0.4	0.7	1.2	0.082
O70370	Ctss	Cathepsin S	0.3	0.8	1.2	0.082
Q06890	Clu / ApoJ	Clusterin / Apolipoprotein J	0.0	1.0	1.4	0.082
Q8BVW0	Ganc	Neutral alpha-glucosidase C	-0.2	0.1	1.0	0.082
Q08642	Padi2	Protein-arginine deiminase type-2	0.0	0.8	1.3	0.082
Q8VDD9	PHIP	PH-interacting protein / Neuronal differentiation-related protein	0.0	-0.7	1.0	0.082
Q8BVW0	Ganc	Neutral alpha-glucosidase C	-0.2	0.1	1.0	0.082
Q9R0M4	PCLP-1	Podocalyxin-like protein 1	-1.2	0.3	1.2	0.082
Q8VHQ3	Ppp1r16b	Protein phosphatase 1 regulatory inhibitor subunit 16B	-0.1	-1.4	ND*	ND*
P16110	Gal-3	Galectin-3, Mac-2 antigen	ND*	1.6	3.0	0.081
P23927	Cryab	Alpha-crystallin B chain	-0.1	0.4	1.1	0.082
P42574	CASP-3	Caspase 3	0.6	ND	3.8	ND*

The protein Calbindin is commonly used as a marker for Purkinje cells and was therefore used as a positive control for our immunohistochemistry procedures. It should be clear that Calbindin is indeed only present in the Purkinje cell body, axon and dendrites residing in the Purkinje, granular and molecular layer respectively (Figure 3). For the DNA-repair knockout mice, both the proteomics data and immunohistochemistry data showed a strong reduction in Calbindin levels and number of Calbindin-positive cells, over time (Figure 3 and Supplementary Figure S3). Furthermore, the staining for the Inositol 1,4,5-trisphosphate activated calcium channel type 1 (IP3R1), cGMP-dependent protein kinase 1 (cGK1), Protein kinase C gamma type (PKCγ) and Glutamate receptor delta-2 subunit (GluRδ2) confirmed the strong abundance and down-regulation in Purkinje cells observed by the proteomics screening (Figure 3, Supplementary figure S3, S4 and S5A). Whereas Calbindin, and PKCγ are present in all compartments of the Purkinje neuron (cell body, axon and dendrite), localization of cGK1 and GluRδ2 is more prominent in the dendrites and cell body and nearly absent in the Purkinje axons. Already after 16 weeks a clear decrease in the number of IP3R1 and cGK1 positive cell bodies in KO tissue is observed. An explanation for the seemingly random pattern of affected Purkinje cell bodies is the stochastic nature of DNA damage that does not stress all cells equally. After 26 weeks, the reduction in IP3R1 and cGK1 protein levels is even more eminent, again confirming the proteomics results. Note the increase in Calbindin, IP3R1 and PKCγ in the few remaining Purkinje cells at 26 weeks.

While Calbindin and IP3R1 are still partly present in Purkinje cell bodies and the ML, a near complete absence of cGK1 is observable. In contrast, while the PKCγ staining shows a clear reduction of axon numbers and a strong shrinkage of most Purkinje cells, PKCγ is still present after 26 weeks. In summary, this indicates that there is a strong effect of DNA-damage on neuronal morphology. As the functions of neurons are heavily dependent on its structural contact and communication with other neurons, it's clear that the observed morphological changes will also effect Purkinje neuron functions.

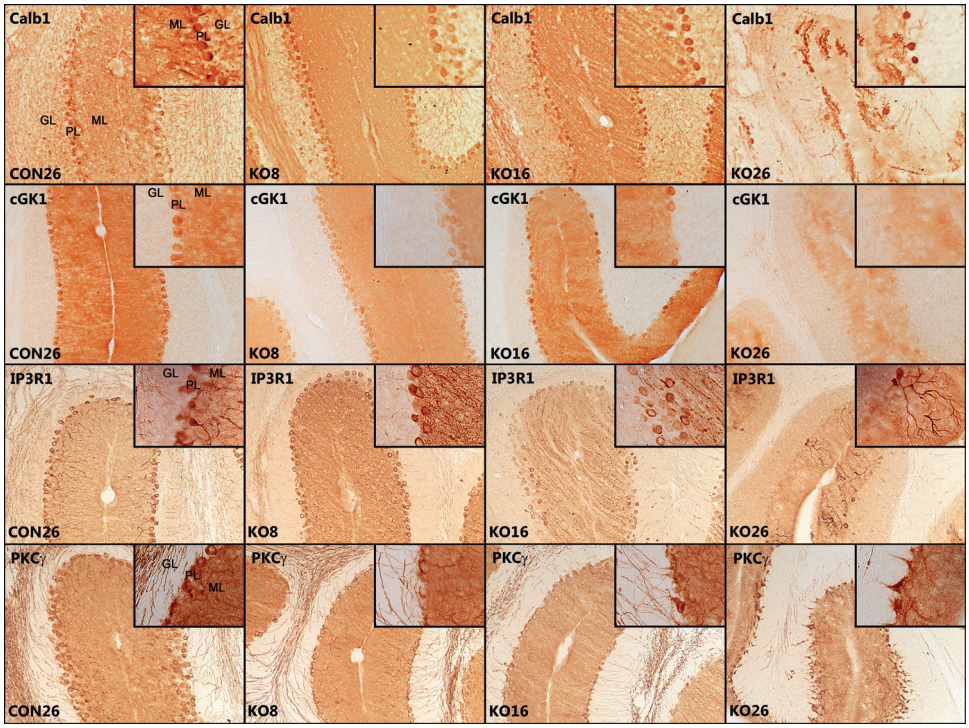

Figure 3: *Immunohistochemical staining of regulated proteins reveals a decrease in synaptic proteins of DNA-repair compromised Purkinje cells. Calbindin (Calb1), cGMP-dependent protein kinase 1 (cGK1), Inositol 1,4,5-trisphosphate receptor type 1 (IP3R1) and Protein kinase C gamma (PKCγ) were found to be only present in Purkinje cells. When comparing 26 weeks old control (CON26) with 8, 16 and 26 weeks old knockout mice (KO8, KO16 and KO26 respectively), an overall decrease of all four proteins is observed in Ercc1 KO mice over time. The organization of down-regulation is different among the proteins. cGK1 is completely removed after 26 weeks, whereas Calbindin and IP3R1 are still present in a few Purkinje cells. PKCγ is still present in most Purkinje cell bodies after 26 weeks, indicating the enormous morphological changes that Purkinje neurons undergo upon DNA-damage. Photographs were taken using 100x and 400x (insets) magnifications.*

As mentioned above, a strong reduction of several Purkinje markers (L7, PEP-19 and Calbindin) was observed. The down-regulation of these markers can be explained in different ways. One simple explanation could be that Purkinje cells with defects in its DNA repair machinery contain a high amount of DNA-damage and therefore the cells will enter apoptosis and disappear. An alternative explanation could be that DNA damage results in a compromise of neuronal cell functioning inducing morphological and functional changes of the affected Purkinje cells. In the second case, the loss of Purkinje markers could indicate the loss of Purkinje characteristic functions rather than cell death/depletion. To test these two hypothesises, the rate of apoptosis was monitored using staining for activated Caspase 3. In addition a GFAP, Galectin-3 (Mac-2), beta-galactosidase and silver staining was performed to visualize neuronal inflammation, phagocytosing macrophages, senescence and neuronal degeneration, respectively.

The protein stainings confirmed the up-regulation of Caspase 3, GFAP and Galectin-3 in

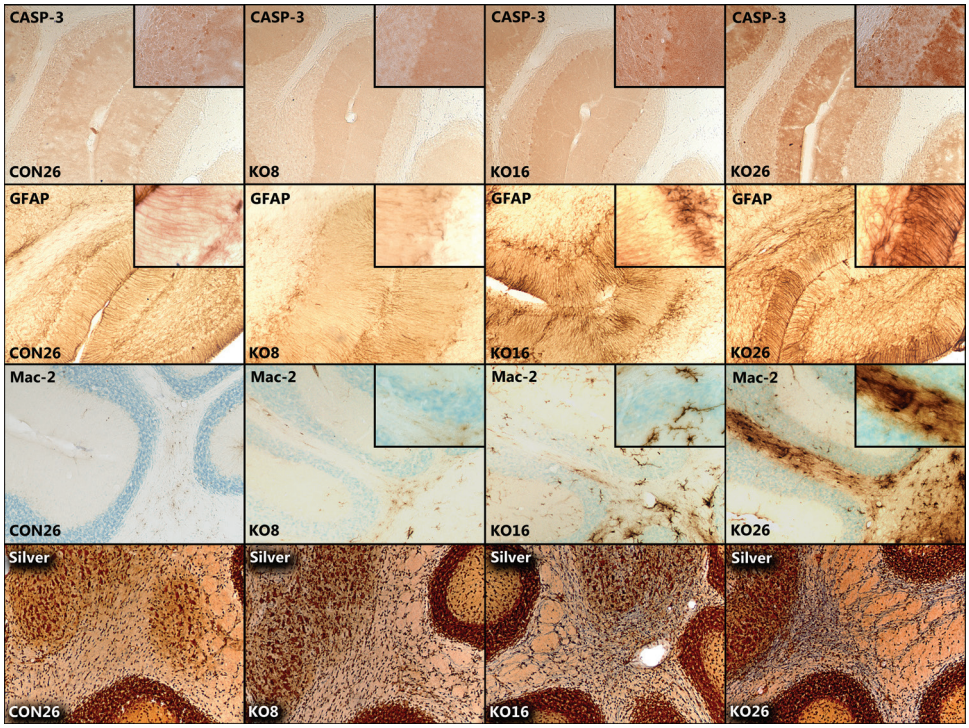

Figure 4: *Immunohistochemical staining of regulated proteins reveals protein regulation in different layers of mice cerebella. A strong increase in activated Caspase 3 (CASP-3) and the astrocyte marker (GFAP) is observed in the molecular layer of Purkinje KO mice cerebella over time (KO8, KO16 and KO26 respectively), indicating dendritic reorganization. The probing of activated macrophages (Mac-2) and staining of neurodegenerative tissue by a silver staining (Silver), indicates neuronal damage is primarily located to the white matter layer (WML) of mice bearing a defective DNA-repair mechanism in their Purkinje cells. Photographs were taken using 100x and 400x (insets) magnifications.*

Purkinje specific Ercc1$^{f/-}$ cerebella as found by the proteomics screening (Figure 4, Table 1, Supplementary figure S3). In addition, IHC revealed that DNA-damage induced a strong localization of Caspase 3 towards the molecular layer, suggesting a strong negative effect of DNA-damage on dendrite synapse maintenance. It has indeed been reported previously that high Caspase activity in dendrites could be responsible for synaptic breakdown and plasticity.(28) The dendrite retraction in Purkinje specific Ercc1 KO mice is further supported by an increase in inflammation in the molecular layer, measured by the up-regulation of the astrocyte markers GFAP and Cx-43 (Table 1, Figure 4, Supplementary Figure S4 and S5B). Furthermore, as a result of these changes, the thickness of the molecular layer of Ercc1 KO cerebella is significantly decreased when compared to control tissues (Supplementary Figure S6). When looking at phagocytosing macrophages (Gal-3 staining) and degenerative tissue (as indicated by the silver stain), there is a strong almost exclusive increase in the white matter layer of Ercc1$^{f/-}$ mice (Figure 4). The white matter layer (WML) is made up largely of myelinated nerve fibres running to and from the cortex, including Purkinje axons (Figure 3D, Supplementary Figure S7). The neurodegenerative staining could possibly in-

dicate a less controlled retraction of Purkinje axons, resulting in the rejection of dead axon structures that need to be cleared by macrophages. In summary, the IHC stainings confirmed a molecular neurodegenerative phenotype of the cerebellum matching to the behavioural phenotype, mimicking aspects of ageing. In addition, the staining indicated different mechanisms could be involved in Purkinje axon and dendrite breakdown, respectively. The beta galactosidase staining did not show any significant staining in any brain tissue and a positive control was not included, therefore it remains unclear if DNA-damaged induced senescence in Purkinje neurons.

Discussion

The link between DNA-damage and ageing(2) has been firmly associated in different organs(1) including the brain(26, 30). Moreover, genetic defects in genome maintenance and DNA repair pathways are responsible for multiple premature symptoms of ageing.(31) Therefore, the study of these progeroid syndromes can shed light on the molecular mechanisms behind the process of ageing.

In this study a cell type specific Ercc1 KO was used to screen for molecular changes in aged cerebella. From time course experiments it was clear that DNA-damage in Purkinje neurons primarily affected the synaptic areas. Both our proteomics screening and immunohistochemistry stainings showed a strong decrease in synaptic protein levels. These proteins include neurotransmitter receptors (i.e. GluRδ2, mGluR1) signalling molecules (i.e. PKCγ, cGK1) and receptor-signal transduction scaffolds (i.e. Homer-3, Delphilin) (Figure 5). Several proteins involved in synaptic signalling were found to be the first to be affected and regulated after 16 weeks of aberrant DNA-repair. Concomitantly, mouse bearing a defective DNA-repair system in Purkinje cells were found to have impaired motor function starting at the age of 16 weeks (van der Vaart et al., in preparation). In addition to decreased motor function, the capacity for motoric learning was also found to be decreased after 26 weeks. The proteomics data suggests that this may be associated with a strong down-regulation of a large number of synaptic proteins encompassing synaptic protein networks (Figure 5). In addition, the immunohistochemistry data allowed us to visualize the neuron structure and its protein localization, revealing a strong retraction of proteins from synaptic areas (both dendrites and axons). As a result of this, a drastic morphological change of the affected Purkinje cells was observed.

As proteins are more actively transcribed, their gene locus becomes less structured and more vulnerable to DNA-damaging processes (i.e. oxidative stress, random errors). The lack of an efficient nucleotide excision repair pathway therefore sensitizes the areas of the genome that are most actively transcribed. It is known that the maintenance of synapse plasticity requires gene expression and de novo protein synthesis, therefore in retrospect the down regulation of synaptic proteins in Ercc1 KO Purkinje cells is not unexpected. Furthermore, the loss of protein expression in strongly affected Purkinje cells was in some cases compensated by neighbouring less affected Purkinje cells that seemed to, in response, express the same synaptic proteins to a higher degree (Figure 3C). These microscopic regulations

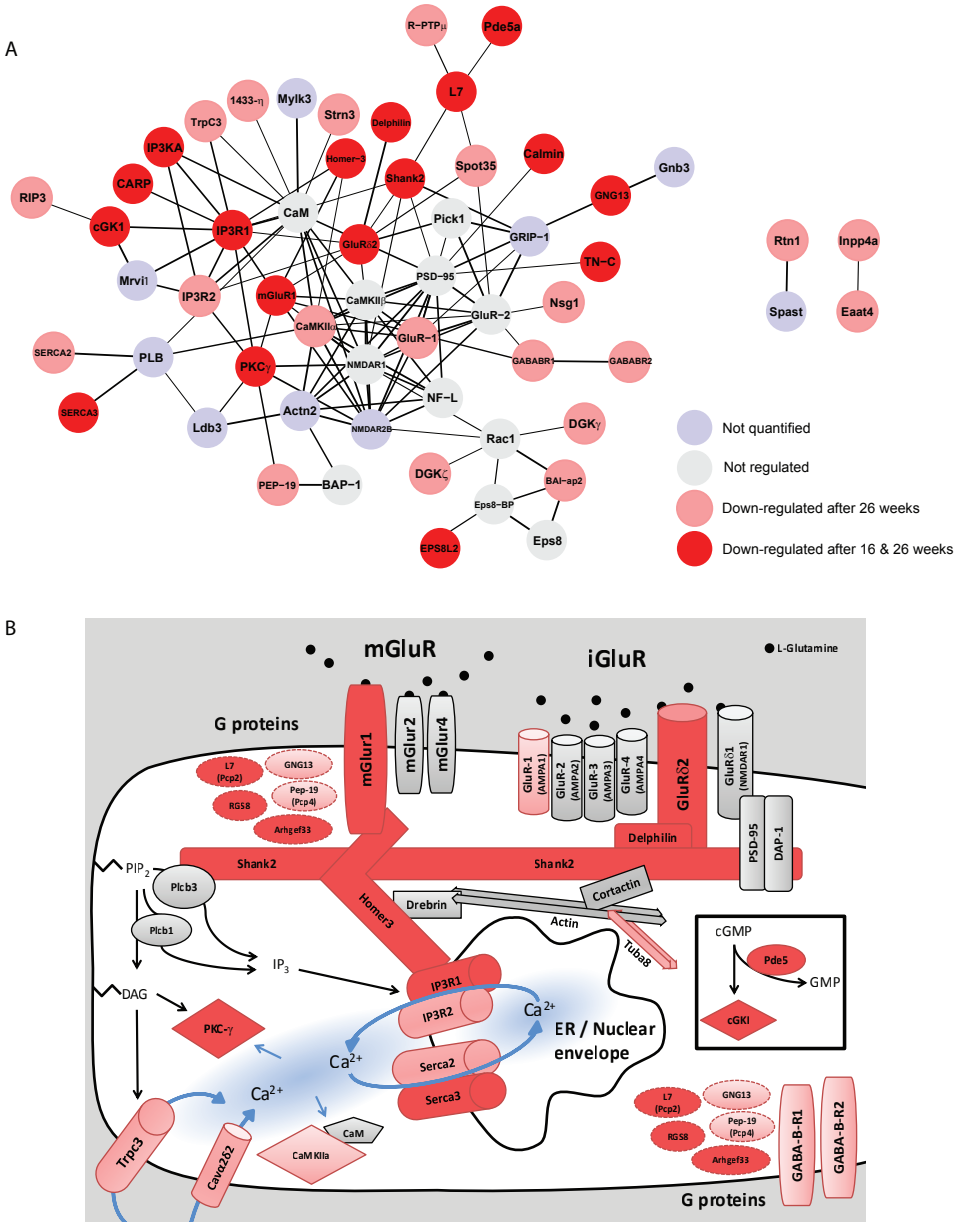

Figure 5: *Effect of DNA-damage on synaptic signalling protein networks. A) A very densely interconnected protein interaction network was calculated by submitting all down-regulated proteins to the STRING database. Interactions of higher confidence are indicated by thicker lines. B) Illustration of possible protein interactions/ pathways that have been reported in literature, which potentially can occur in Purkinje cells. The dotted outlined proteins are placed near G-protein coupled receptors to show their hypothetical interaction with the receptors and their downstream signalling proteins. Both metabotropic (mGluR) and ionotropic (iGluR) up- and downstream synaptic signalling pathways display a strong sensitivity for DNA-damage. Color coding in B is the same as in A.*

cannot be deduced from high-throughput proteomics data or western blot validation, as in these experiments the protein levels are averaged out over all cells present in the cerebellum, indicating a good complementary between proteomics and immunohistochemistry.

Our results also indicate that DNA-damage does not result in direct Purkinje cell removal but gradual functional decline and tissue degeneration. Although, a set of Purkinje specific markers are depleted in the DNA-repair defective Purkinje cells (Table 1), the localisation of Protein kinase C gamma (PKCγ) (Figure 3D) illustrates that DNA repair knockout Purkinje cell bodies have not all disappeared but at least some have drastically changed their morphology and protein localization. Therefore, increased reduction of Purkinje markers over time, indicates functional decline rather than neuron loss upon increasing amounts of DNA-damage. Moreover, the loss of neuronal functions is in line with previous reports studying ageing mechanisms that showed a decrease in neuronal signal processing capabilities(32), reduction of the post-synaptic density(33) and a reduced number of synapses(34) in aged rats.

Finally, in addition to the resemblance with the ageing phenotype, the results presented here could potentially shed some light on the development of motoric diseases such as spinocerebellar ataxia (SCA). IP3R1, mGluR1, Cavα2δ2 and PKC-γ gene mutations are responsible for different types of SCA and considering the motoric impairment of our model system and the strong regulation of these proteins, the list of regulated proteins could also be used as an extended target list to study cerebellar motor diseases, such as ataxia.

Supporting Information

Supplementary Figures are located in Chapter 8; Supplementary Figures to Chapter 2. Supplementary Figures and Tables can also be accesed online from http://mcponline.com.

Acknowledgements

The Netherlands Proteomics Centre, embedded in The Netherlands Genomics Initiative is acknowledged for financial support as is the Centre for Biomedical Genetics. This work was further supported in part by the PRIME-XS project, Grant Agreement Number 262067, funded by the European Union Seventh Framework Program. We acknowledge financial support of the European commission FP7 Markage (FP7-Health-2008-200880), DNA Repair (LSHG-CT-2005-512113) and LifeSpan (LSHG-CT-2007-036894), National Institute of Health (NIH)/National Institute of Ageing (NIA) (1PO1 AG-17242-02), NIEHS (1UO1 ES011044), the Royal Academy of Arts and Sciences of the Netherlands (academia professorship to JHJH) and a European Research Council Advanced Grant to JHJH. The research leading to these results has received funding from the European Community's Seventh Framework Programme (FP7/2007-2013) under grant agreement No. HEALTH-F2-2010-259893.

References

1. Hoeijmakers, J. H. (2009) DNA damage, aging, and cancer. N Engl J Med 361, 1475-1485.

2. Wang, J., Clauson, C. L., Robbins, P. D., Niedernhofer, L. J., and Wang, Y. (2012) The oxidative DNA lesions 8,5'-cyclopurines accumulate with aging in a tissue-specific manner. Aging Cell.

3.Kyng, K. J., and Bohr, V. A. (2005) Gene expression and DNA repair in progeroid syndromes and human aging. Ageing Res Rev 4, 579-602.

4. McWhir, J., Selfridge, J., Harrison, D. J., Squires, S., and Melton, D. W. (1993) Mice with DNA repair gene (ERCC-1) deficiency have elevated levels of p53, liver nuclear abnormalities and die before weaning. Nat Genet 5, 217-224.

5. Weeda, G., Donker, I., de Wit, J., Morreau, H., Janssens, R., Vissers, C. J., Nigg, A., van Steeg, H., Bootsma, D., and Hoeijmakers, J. H. (1997) Disruption of mouse ERCC1 results in a novel repair syndrome with growth failure, nuclear abnormalities and senescence. Curr Biol 7, 427-439.

6. Niedernhofer, L. J., Garinis, G. A., Raams, A., Lalai, A. S., Robinson, A. R., Appeldoorn, E., Odijk, H., Oostendorp, R., Ahmad, A., van Leeuwen, W., Theil, A. F., Vermeulen, W., van der Horst, G. T., Meinecke, P., Kleijer, W. J., Vijg, J., Jaspers, N. G., and Hoeijmakers, J. H. (2006) A new progeroid syndrome reveals that genotoxic stress suppresses the somatotroph axis. Nature 444, 1038-1043.

7. Lawrence, N. J., Sacco, J. J., Brownstein, D. G., Gillingwater, T. H., and Melton, D. W. (2008) A neurological phenotype in mice with DNA repair gene Ercc1 deficiency. DNA Repair (Amst) 7, 281-291.

8. Vegh, M. J., de Waard, M. C., van der Pluijm, I., Ridwan, Y., Sassen, M. J., van Nierop, P., van der Schors, R. C., Li, K. W., Hoeijmakers, J. H., Smit, A. B., and van Kesteren, R. E. (2012) Synaptic proteome changes in a DNA repair deficient ercc1 mouse model of accelerated aging. Journal of proteome research 11, 1855-1867.

9. Schrimpf, S. P., Meskenaite, V., Brunner, E., Rutishauser, D., Walther, P., Eng, J., Aebersold, R., and Sonderegger, P. (2005) Proteomic analysis of synaptosomes using isotope-coded affinity tags and mass spectrometry. Proteomics 5, 2531-2541.

10. Nielsen, P. A., Olsen, J. V., Podtelejnikov, A. V., Andersen, J. R., Mann, M., and Wisniewski, J. R. (2005) Proteomic mapping of brain plasma membrane proteins. Molecular & cellular proteomics : MCP 4, 402-408.

11. Coughenour, H. D., Spaulding, R. S., and Thompson, C. M. (2004) The synaptic vesicle proteome: a comparative study in membrane protein identification. Proteomics 4, 3141-3155.

12. Cheng, D., Hoogenraad, C. C., Rush, J., Ramm, E., Schlager, M. A., Duong, D. M., Xu, P., Wijayawardana, S. R., Hanfelt, J., Nakagawa, T., Sheng, M., and Peng, J. (2006) Relative and absolute quantification of postsynaptic density proteome isolated from rat forebrain and cerebellum. Molecular & cellular proteomics : MCP 5, 1158-1170.

13. Barski, J. J., Dethleffsen, K., and Meyer, M. (2000) Cre recombinase expression in cerebellar Purkinje cells. Genesis 28, 93-98.

14. Doig, J., Anderson, C., Lawrence, N. J., Selfridge, J., Brownstein, D. G., and Melton, D. W. (2006) Mice with skin-specific DNA repair gene (Ercc1) inactivation are hypersensitive to ultraviolet irradiation-induced skin cancer and show more rapid actinic progression. Oncogene 25, 6229-6238.

15. Boersema, P. J., Raijmakers, R., Lemeer, S., Mohammed, S., and Heck, A. J. (2009) Multiplex peptide stable isotope dimethyl labeling for quantitative proteomics. Nat Protoc 4, 484-494.

16. Gauci, S., Helbig, A. O., Slijper, M., Krijgsveld, J., Heck, A. J., and Mohammed, S. (2009) Lys-N and trypsin cover complementary parts of the phosphoproteome in a refined SCX-based approach. Analytical chemistry 81, 4493-4501.

17. Roxas, B. A., and Li, Q. (2008) Significance analysis of microarray for relative quantitation of LC/MS data in proteomics. BMC Bioinformatics 9, 187.

18. Dahlhaus, M., Li, K. W., van der Schors, R. C., Saiepour, M. H., van Nierop, P., Heimel, J. A., Hermans, J. M., Loos, M., Smit, A. B., and Levelt, C. N. (2011) The synaptic proteome during development and plasticity of the mouse visual cortex. Molecular & cellular proteomics : MCP 10, M110 005413.

19. Munoz, J., Low, T. Y., Kok, Y. J., Chin, A., Frese, C. K., Ding, V., Choo, A., and Heck, A. J. (2011) The quantitative proteomes of human-induced pluripotent stem cells and embryonic stem cells. Mol Syst Biol 7, 550.

20. Tusher, V. G., Tibshirani, R., and Chu, G. (2001) Significance analysis of microarrays applied to the ionizing radiation response. Proceedings of the National Academy of Sciences of the United States of America 98, 5116-5121.

21. Jaarsma, D., Haasdijk, E. D., Grashorn, J. A., Hawkins, R., van Duijn, W., Verspaget, H. W., London, J., and Holstege, J. C. (2000) Human Cu/Zn superoxide dismutase (SOD1) overexpression in mice causes mitochondrial vacuolization, axonal degeneration, and premature motoneuron death and accelerates motoneuron disease in mice expressing a familial amyotrophic lateral sclerosis mutant SOD1. Neurobiol Dis 7, 623-643.

22. Nadler, J. V., and Evenson, D. A. (1983) Use of excitatory amino acids to make axon-sparing lesions of hypothalamus. Methods Enzymol 103, 393-400.

23. de Waard, M. C., van der Pluijm, I., Zuiderveen Borgesius, N., Comley, L. H., Haasdijk, E. D., Rijksen, Y., Ridwan, Y., Zondag, G., Hoeijmakers, J. H., Elgersma, Y., Gillingwater, T. H., and Jaarsma, D. (2010) Age-related motor neuron degeneration in DNA repair-deficient Ercc1 mice. Acta Neuropathol 120, 461-475.

24. Walther, D. M., and Mann, M. (2011) Accurate quantification of more than 4000 mouse tissue proteins reveals minimal proteome changes during aging. Molecular & cellular proteomics : MCP 10, M110 004523.

25. Borgesius, N. Z., de Waard, M. C., van der Pluijm, I., Omrani, A., Zondag, G. C., van der Horst, G. T., Melton, D. W., Hoeijmakers, J. H., Jaarsma, D., and Elgersma, Y. (2011) Accelerated age-related cognitive decline and neurodegeneration, caused by deficient DNA repair. J Neurosci 31, 12543-12553.

26. Lu, T., Pan, Y., Kao, S. Y., Li, C., Kohane, I., Chan, J., and Yankner, B. A. (2004) Gene regulation and DNA damage in the ageing human brain. Nature 429, 883-891.

27. Penkowa, M., Carrasco, J., Giralt, M., Moos, T., and Hidalgo, J. (1999) CNS wound healing is severely depressed in metallothionein I- and II-deficient mice. J Neurosci 19, 2535-2545.

28. Gilman, C. P., and Mattson, M. P. (2002) Do apoptotic mechanisms regulate synaptic plasticity and growth-cone motility? Neuromolecular Med 2, 197-214.

29. Jacobs, B., Driscoll, L., and Schall, M. (1997) Life-span dendritic and spine changes in areas 10 and 18 of human cortex: a quantitative Golgi study. J Comp Neurol 386, 661-680.

30. Esiri, M. M. (2007) Ageing and the brain. J Pathol 211, 181-187.

31. Garinis, G. A., van der Horst, G. T., Vijg, J., and Hoeijmakers, J. H. (2008) DNA damage and ageing: new-age ideas for an age-old problem. Nat Cell Biol 10, 1241-1247.

32. Deupree, D. L., Bradley, J., and Turner, D. A. (1993) Age-related alterations in potentiation in the CA1 region in F344 rats. Neurobiol Aging 14, 249-258.

33. Nicholson, D. A., Yoshida, R., Berry, R. W., Gallagher, M., and Geinisman, Y. (2004) Reduction in size of perforated postsynaptic densities in hippocampal axospinous synapses and age-related spatial learning impairments. J Neurosci 24, 7648-7653.

34. Geinisman, Y., deToledo-Morrell, L., Morrell, F., Persina, I. S., and Rossi, M. (1992) Age-related loss of axospinous synapses formed by two afferent systems in the rat dentate gyrus as revealed by the unbiased stereological dissector technique. Hippocampus 2, 437-444.

Improving SRM Assay Development: A Global Comparison between Triple Quadrupole, Ion Trap and Higher Energy CID Peptide Fragmentation Spectra

Erik L. de Graaf[1,2,3], A. F. Maarten Altelaar[1,2], Bas van Breukelen[1,2,4], Shabaz Mohammed[1,2] and Albert J. R. Heck[1,2,3]

published in

Journal of Proteome Research, July 2011

[1] Biomolecular Mass Spectrometry and Proteomics, Bijvoet Center for Biomolecular Research and Utrecht Institute for Pharmaceutical Sciences, Utrecht University, Padualaan 8, 3584 CH Utrecht, The Netherlands

[2] Netherlands Proteomics Centre, Padualaan 8, 3584 CH Utrecht, The Netherlands

[3] Center for Biomedical Genetics

[4] Netherlands Bioinformatics Centre, Padualaan 8, 3584 CH Utrecht, The Netherlands

Abstract

In proteomics selected reaction monitoring (SRM) is rapidly gaining importance for targeted protein quantification. The triple quadrupole mass analyzers used in SRM assays allow for levels of specificity and sensitivity hard to accomplish by more standard shotgun proteomics experiments. Often, an SRM assay is built by in silico prediction of transitions and/or extraction of peptide precursor and fragment ions from a spectral library. Spectral libraries are typically generated from non-ideal ion trap based shotgun proteomics experiments or synthetic peptide libraries, consuming considerable time and effort. Here, we investigate the usability of beam type CID (or 'higher energy CID' (HCD)) peptide fragmentation spectra, as acquired using an Orbitrap Velos, to facilitate SRM assay development. Therefore, peptide fragmentation spectra, obtained by ion-trap CID, triple-quadrupole CID (QqQ-CID) and Orbitrap HCD, originating from digested cellular lysates, were compared. Spectral comparison and a dedicated correlation algorithm indicated significantly higher similarity between QqQ-CID and HCD fragmentation spectra than between QqQ-CID and ion trap-CID spectra. SRM transitions generated using a constructed HCD spectral library increased SRM assay sensitivity up to two-fold, when compared to the use of a library created from more conventionally used ion trap-CID spectra, showing that HCD spectra can assist SRM assay development.

Introduction

In recent years, the proteomics community has started to explore selected reaction monitoring (SRM) as a method to quantitatively assess the abundance of a targeted set of proteins.(1-3) Such SRM assays are typically performed by LC-MS/MS with triple quadrupole (QqQ) instruments, in which the selectivity of two mass filters is applied to target a specific peptide via the relatively unique combination of peptide m/z and multiple diagnostic CID fragment ions, along with peptide retention time. The main benefits of using a triple quadrupole as a sophisticated filter is that it allows one to overcome the issue of sample complexity and to remove background chemical noise, increasing both sensitivity and dynamic range.

In order to create SRM assays, the required specific peptide transitions can be extracted from high-throughput MS/MS fragmentation spectra acquired by shot-gun proteomics experiments and/or data repositories.(4-8) Up to now, in this discovery phase of SRM assay development, peptide fragmentation spectra were most often created using relatively low energy CID ion-trap (IT-MS) instruments.(9) However, it has been well-documented that peptide fragmentation spectra may differ significantly when switching from an IT-MS in the discovery phase to a QqQ instrument for the SRM assay phase. The energy deposited into a peptide differs between IT (resonance) CID and QqQ (beam-type) CID. In CID, an ion is accelerated to a user defined velocity (a defined level of kinetic energy) and is then smashed into gas atoms. The kinetic energy is converted to internal, mostly vibrational, energy which subsequently leads to fragmentation. The extent of fragmentation depends on the size of the gas atom (larger leads to a higher level of conversion) and the kinetic energy of the ion. The differences between IT- and QqQ-CID can be attributed to the gas conditions and the defined window of kinetic energy.(10) The latest developments on the Orbitrap instrument now allow peptide fragmentation spectra to be acquired using beam type CID, so-called HCD.(11) All three fragmentation methods (IT, QqQ and HCD) are considered 'low-energy' and under such conditions, peptide fragmentation adheres to the 'mobile proton model'.(12) The model states that fragment abundance is primarily dictated by the proton affinity of the fragment. Beam-type fragmentation (HCD and QqQ) occurs in a slightly higher energy regime than resonance excitation CID (IT-CID) and as a result fragment ions can retain higher levels of energy leading to further decomposition. For instance, the N-terminal b-ion fragments often dissociate further under beam-type conditions when compared to IT-CID.(13, 14) Another difference between ion trap and beam-type fragmentation is the rate at which energy is transferred to the peptide. IT-CID occurs at slower rates (ms range) when compared to beam type (μs range), allowing for additional rearrangement reactions to occur. As a result of these fundamental differences, CID MS/MS spectra of the same peptide can differ significantly depending on the fragmentation method. Therefore, when using IT-CID data as starting point for SRM assay development, multiple ion species need to be monitored and the collision energy needs to be optimized, making the development of SRM assays, based on IT-CID data, laborious.

It has been shown that synthetic peptides, predicted in silico to be protein specific (proteotypic), can significantly facilitate SRM assay development.(3, 15) By fragmenting these proteotypic peptides on a QqQ instrument, the most favorable SRM coordinates can be

extracted from the peptide MS/MS spectra. Picotti et al. demonstrated that when using this targeted proteomic approach in yeast, low abundant proteins could be quantified in relatively high-throughput assays.(16) A drawback of this approach is that it requires the costly synthesis of a large array of peptides and still a significant amount of QqQ instrument/analysis time whereby only a previously determined small set of proteins can be monitored, hampering the possibility of finding potentially interesting novel proteins. Moreover, if the peptide utilized for a protein assay is only predicted rather than experimentally detected, the developed assay may actually be inappropriate since not all parameters (digestion efficiency, peptide solubility, peptide ionization efficiency and general matrix effects) for these hypothesis-driven experiments are known.

Recent improvements introduced the efficient use of HCD(11) peptide fragmentation to the Orbitrap with a performance in speed and sensitivity starting to rival that of ion traps. (17-21) Here, we use a large set of peptides to explore whether there is an advantage of using HCD over IT-CID peptide fragmentation spectra for SRM assay development. We performed both a qualitative and quantitative comparison of MS/MS spectra generated by IT-CID, QqQ-CID and HCD. Peptides sequenced and identified by each method were selected and compared using an in-house developed cross correlation tool. In order to investigate further the consequences of selecting a SRM transitions from MS/MS spectra, transitions were extracted from IT-CID or HCD derived spectra and subsequently measured on a QqQ instrument for comparison of the peptide fragment intensities.

Experimental Procedures

Sample preparation - HEK293 and HeLa cells were cultured at 37°C and 5% CO_2 in DMEM high glucose medium containing 10% FCS, 10mM L-Glutamine and 5% Penicillin/Streptomycin (all Lonza). Cells were harvested by trypsinization and washed three times with PBS buffer, before lysis with the lysis buffer (8M Urea in 50mM ammonium bicarbonate, 1 tablet Complete mini EDTA-free Cocktail (Roche) and 1 tablet PhosSTOP phosphatase inhibitor Cocktail (Roche)). After centrifugation (20,000g 1.5h at 4°C), the supernatant was assayed for protein content using the BCA-kit standard procedure from Pierce. Protein reduction and alkylation were performed using final concentrations of 2mM dithiothreitol and 4mM iodoacetamide, repsectively. A first enzymatic digestion step was performed in 8M urea lysis buffer using Lys-C at 37°C for 4h (enzyme:substrate ratio 1:75). The second digestion was performed overnight (37°C) with trypsin (enzyme:substrate ratio 1:100) in 2M Urea. Sample cleanup was achieved using 200mg Sep-Pak C18 columns (Waters). The eluate was dried to completion and subsequently reconstituted in 10% formic acid prior to fractionation using strong-cation exchange (SCX) as described previously.(22) In short, SCX was performed using a Zorbax BioSCX-Series II column (0.8-mm inner diameter × 50-mm length, 3.5 μm). SCX Solvent A consists of 0.05% formic acid in 20% acetonitrile, solvent B consists of 0.05% formic acid, 0.5 M NaCl in 20% acetonitrile. The SCX gradient is as follows: 0-0.01 min (0-2% B); 0.01-8.01 min (2-3% B); 8.01-14.01 min (3-8% B); 14.01-28 min (8-20% B); 28-38 min (20-40% B); 38-48 min (40-90% B); 48-54 min (90% B); 54-60 min (0% B). After injection of 1 mg HeLa or 1 mg HEK293 cell lysate, a total of 50 SCX fractions (1 min each, i.e. 50 μL

elution volume) were collected per cell lysate and dried in a vacuum centrifuge. One SCX fraction of the HEK293 cell lysate, containing only doubly charged peptides, was used for MS/MS comparison. MS data of an additional ten SCX fractions of the HeLa cell lysate, corresponding to doubly and triply charged peptides, were acquired and are available for download to aid anyone who would wish to build an extensive HCD spectral library for SRM assay development (Supplementary data 1).

Liquid Chromatography and Mass Spectrometry - Data were acquired using three different mass-spectrometry set-ups. The same online nanoLC configuration was used for all LC-MS experiments.(23) An Agilent 1100 series LC system was equipped with an Reprosil C18-AQ (Dr. Maisch, Ammerbuch, Germany), 100 μm x 20 mm, 3 μm, 120 Å double frit trapping column and a Reprosil C18-AQ (Dr. Maisch, Ammerbuch, Germany), 50 μm x 40 cm, 3 μm, 120 Å analytical column. Trapping was performed at 5 μL/min for 10 min in solvent A (0.1M acetic acid in water), elution was achieved with a non-linear gradient of 10–50% B (0.1M acetic acid in 80%/20% acetonitrile/water) in 267 min, with a total analysis time of 300 min. The flow rate was passively split to 50 nL/min during elution analysis. Nanospray was achieved with an in-house pulled and gold-coated fused silica capillary (o.d. 360 μm; i.d. 20 μm; tip i.d. 10 μm) and an applied voltage of 1.7 kV. Two type of experiments were performed on the Orbitrap Velos (Thermo, San Jose, CA). Ion trap fragmentation analysis was performed using a survey scan in FT mode (350-1500 m/z) with a resolution of 60,000 and accumulation to a target value of 500,000, followed by IT-MS/MS fragmentation (target value of 5,000) of the 20 most intense peaks. The HCD fragmentation method consisted of a FT survey scan from 350-1500 m/z (resolution 30,000) and accumulation to a target value of 500,000 followed by HCD fragmentation (target value of 30,000) of the ten most intense peaks and a readout in the FT analyser (resolution 7500). A normalized HCD collision energy of 35 was used, corresponding to the equations CE= 0.041*m/z - 0.573 and CE= 0.051*m/z + 0.095 for doubly and triply charged precursors. For the triple quadrupole mass analyzer set-up, a TSQ Vantage (Thermo, San Jose, CA) was configured as followed: Q3 Full MS survey scan (m/z 200-1200) at 0.4 Da peak width and 1 sec scan time, followed by two data dependent tandem MS/MS scans using a scan time of 0.6 sec and Q1 and Q3 peak width set at 3.0 and 0.4 Da, respectively. The QqQ collision energy was calculated by one equation for both doubly and triply charged peptides and was optimized on CE = 0.048*m/z.

Data analysis - Peak lists were generated from the raw data files using proteome discoverer version 1.2. Peptide identification was performed by searching the peak lists against the Swissprot v56.2 human taxonomy database including a concatenated decoy version and common contaminants by using mascot software version 2.3.02 (Matrix Science). Trypsin was used with a maximum of two missed cleavages and carbamidomethylation (C) and oxidation (M) was set as constant and variable modification, respectively. Searches with IT-CID and HCD data were performed using a peptide and fragment tolerance of 15 ppm and 0.6 Da, respectively. QqQ-CID data was searched using 0.9 Da precursor and 0.9 Da fragment mass tolerance. After applying the Scaffold filter described below, the amount of decoy hits in the list of peptides identified by all three methods was zero, thus we had an apparent FDR of zero percent for the data used (see supplementary table 1 for a more complete summary of FDR calculations). To facilitate data analysis, an in-house developed

software tool built in JAVA (v1.6) was used to filter, visualize and calculate cross-correlations (Xcorr). MS2Xcorrelation takes as an input Scaffold 2.3 (Proteome Software Inc., Portland, OR) spectral report outputs and the corresponding mascot generic file (MGF) files for each dataset. In Scaffold, all Mascot search files (.DAT) were imported as 3 separate experiments: IT-CID, HCD and QqQ-CID. Subsequently, a spectrum report was generated, in which only peptides with a peptide prophet score of greater than 90% were included. In addition, three MGF files were also exported from Scaffold using the mzDATA export option. The MGF files and spectrum report serve as input for MS2Xcorrelation, which reads and stores all peptide annotations with their corresponding spectrum in a local database. MS2Xcorrelation subsequently finds all peptides that have been identified by all 3 methods, IT-CID, HCD and QqQ-CID, and allows the user to select each of the peptides to visualize the spectra, and calculates an Xcorr between any two selected spectra. Moreover, MS2Xcorrelation allows a fully automated Xcorr calculation of all peptides found in at least 3 different experiments, and exports all Xcorr correlation values and other statistics to a tab delimited (excel readable) tabular file. Xcorrelation was calculated by Fast Fourier Transform (FFT) and correlation, using the Flanagans Java Scientific library v. nov 2010, similar to that described by Frewen *et al*..(8) Before calculating the Xcorr between two spectra, the spectra were normalized to the most intense peak and filtered to remove noise, firstly by retaining the 2 most abundant peaks in each window of 40 Th, and then by reducing the total spectral peaks to a maximum of 15. The Xcorr resolution can be set (as an order of 2n). In our analysis, the Xcorr resolution was selected to be 1024, resulting in ~1 Th bins in the m/z range, reflecting the resolution of the TSQ QqQ.

SRM analysis - SRM transitions were created and evaluated using Skyline(24). First, a IT-CID and HCD spectral library were created with Skyline from the HeLa whole cell lysate sample. Next, the seven highest ranking b- and y-ions for approximately 80 peptides were selected from the two libraries to create two SRM assays (~450 transitions per assay). Collision energy was calculated based on the precursor charge state and mass-to charge ratio using the standard equations CE= 0.03*m/z + 2.905 and CE= 0.038*m/z + 2.281 for doubly and triply charged precursors, respectively. After data acquisition on the TSQ Vantage, only those peptide transition sets adhering to the following requirements were kept: a clear difference between signal and background noise, identical peptide retention times between the HCD and IT-CID built SRM assay, co-elution of at least 3 transitions and a dotp value of at least 0.83 in one of the two assays. Collision energy (CE) optimization was performed on fragment ions using a range of -5 to +5 V around the calculated CE value in steps of 1V.

Results

Defining transitions that occur in triple quadrupole MS/MS peptide fragmentation is a crucial step in the development of SRM assays. In order to compare triple quadrupole fragmentation (beam type CID or QqQ-CID) data with data from other mass analysers, MS/MS spectra were obtained on two tandem MS platforms: a TSQ Vantage QqQ and an Orbitrap Velos capable of performing IT-CID (resonance excitation CID) and HCD (beam type CID). Two complex peptide mixtures were analyzed using nanoLC-MS/MS. Firstly, an SCX

fraction of a HEK293 cell lysate was analyzed, enriched for potential proteotypic doubly charged peptides containing no miscleavages or post-translational modifications(22, 25), mimicking the peptides often used for SRM assays. Secondly, a HeLa whole cell lysate was used, being more representative of samples typically subjected to SRM-based proteomics. This complete cell lysate digest contains more basic and more highly charged peptides that we consider as a valuable alternative for SRM assays when few smaller tryptic peptides are available for a target protein.

Comparison of peptide fragmentation spectra acquired by IT-CID, QqQ-CID and HCD

A global computational analysis of all common product ion spectra generated by each of the three platforms was performed using an in-house developed script (MS2Xcorrelator). The script calculates the cross correlation (Xcorr score)(8, 26, 27) between spectra for the same peptide sequence generated using the different fragmentation platforms. A higher Xcorr value indicates a better correlation between the two spectra. The maximum Xcorr score is unique for every peptide. As SRM assays are measured on a TSQ Vantage QqQ, we defined the maximal cross correlation score as the score associated with the comparison of a QqQ spectrum with itself, which we refer to as the autocorrelation. MS/MS spectra containing noise can 'artificially' generate high Xcorr values due to simply random chance matching. To account for such phenomena we compared spectra which should not match. Systematically shifting all peaks to a different mass in one of the two spectra would allow evaluation of false positive matches. Thus, Xcorr values were calculated at different mass

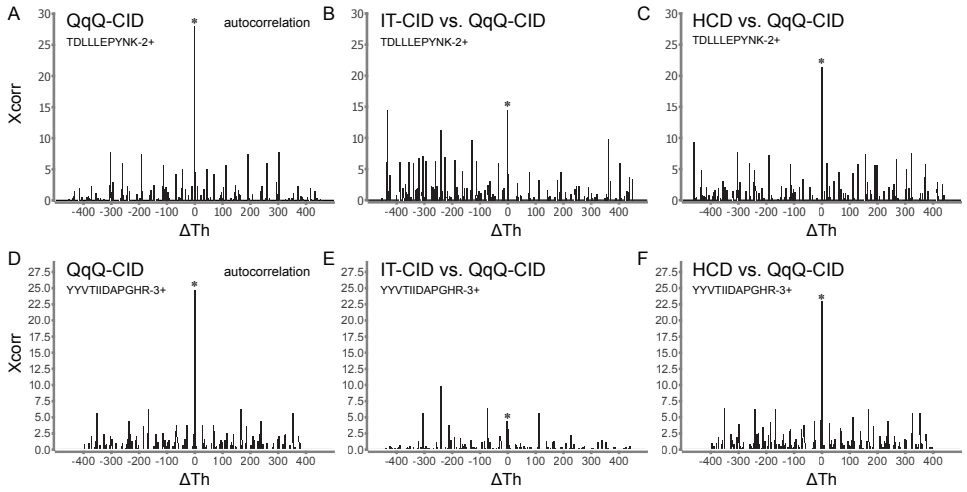

Figure 1: *Comparison of IT-CID and HCD spectra to TSQ spectra for the doubly charged peptide TDLLEPYNK (A-C) and the triply charged peptide YYVTIIDAPGHR (D-F). A) TSQ self-correlation Xcorr plot, indicating a Xcorr score of 28.0. B) IT-CID vs TSQ Xcorr plot indicating a correlation between the two fragmentation methods (rank 1 at 0 ΔTh) with a Xcorr score 14.3. C) HCD vs TSQ Xcorr plot showing a Xcorr score of 21.3 D) TSQ self-correlation Xcorr plot, indicating a maximum Xcorr of 25.8. E) IT-CID vs TSQ Xcorr plot showing a low ranking peak at 0 ΔTh with a Xcorr score 4.4. F) HCD vs TSQ Xcorr plot showing a rank 1 peak at 0 ΔTh and an Xcorr score of 23.1 close to the TSQ self-correlation. Xcorr peaks at zero thomson shift are indicated by an asterix.*

displacements between the two spectra and these values are plotted against the mass displacement (the mass shift). If the peak with the highest Xcorr value is located at a mass shift of zero Thomson then the best match corresponds to the two unadjusted original spectra and allows a context to be placed on the strength of the XCorr value. An example of such an Xcorr plot is shown in Figure 1 for the doubly charged peptide TDLLEPYNK. The QqQ-CID autocorrelation shows the highest Xcorr score (rank 1) at no mass shift (ΔTh=0). This indicates that the QqQ-CID spectrum correlates best with itself when the spectra are not shifted in mass relative to each other. Therefore, optimal Xcorr scores of peptide fragmentation spectra obtained by IT-CID or HCD should be close to the QqQ-CID autocorrelation values.

There is a significant number of IT-CID spectra showing a rank 1 Xcorr peak at no mass shift, indicating a reasonable correlation with the QqQ-CID spectra, in agreement with previous studies. (9, 28, 29) However, when we evaluated many HCD versus QqQ-CID Xcorr plots and IT-CID versus QqQ-CID Xcorr plots, generally the Xcorr scores at ΔTh=0 were found to be significantly higher for HCD, indicating a higher correlation between QqQ-CID and HCD data than between QqQ-CID and IT-CID. An illustrative example is shown for the peptide TDLLEPYNK in Figure 1B and 1C and this trend will be further investigated below.

Typically IT-CID Xcorr plots for a triply charged peptide showed a low ranking Xcorr at ΔTh=0, often associated with a low Xcorr score. In comparison, the HCD Xcorr plots showed a rank 1 peak at a zero mass shift and an Xcorr close to the QqQ-CID autocorrelation values. This trend is illustrated by the Xcorr plots for the triply charged peptide YYVTIIDAPGHR shown in Figure 1D, E and F, where a poor correlation can be seen between IT-CID and QqQ-CID (rank 5 at ΔTh=0, Xcorr score 4.4), while the HCD versus QqQ-CID correlation (rank 1 at ΔTh=0, Xcorr score 23.1) is close to the QqQ-CID autocorrelation (rank 1 at ΔTh=0, Xcorr score 25.8). It has been reported that IT-CID fragmentation of higher charged peptides is more distinct when compared to beam-type CID fragmentation (17, 21), and this may, in part, explain why HCD shows a higher correlation to QqQ-CID. This could imply that in the context of developing SRM assays for triply charged peptides, HCD fragmentation could prove to be more informative than IT-CID fragmentation and therefore will be further investigated.

The high similarity between HCD and QqQ-CID can easily be ascertained by visual inspection of the spectra (Figure 2). When comparing IT-CID to QqQ-CID and HCD fragmentation spectra it was clearly noticeable that there were more high mass b-ions present in the IT-CID generated spectra, as exemplified in Figure 2A, C and E. The loss of large b-ions is well-known and is most likely due to multiple stages of fragmentation in beam-type CID. (13, 14) For triply charged peptides, the spectra also show clearly higher similarity between HCD and QqQ-CID than for IT-CID and QqQ-CID (Figure 2B,D and F).

Incidentally, HCD and IT-CID fragmentation spectra typically exhibited a higher signal and signal-to-noise ratio when compared to QqQ-CID spectra that were obtained on the TSQ Vantage. A simple explanation for these superior spectra is the use of automatic gain control (AGC), by the Velos, which attempts to gather a statistically significant amount of ions before fragmentation takes place. In contrast, the TSQ Vantage acquires spectra for a fixed amount of time, therefore the amount of ions per spectra can be less optimal, resulting

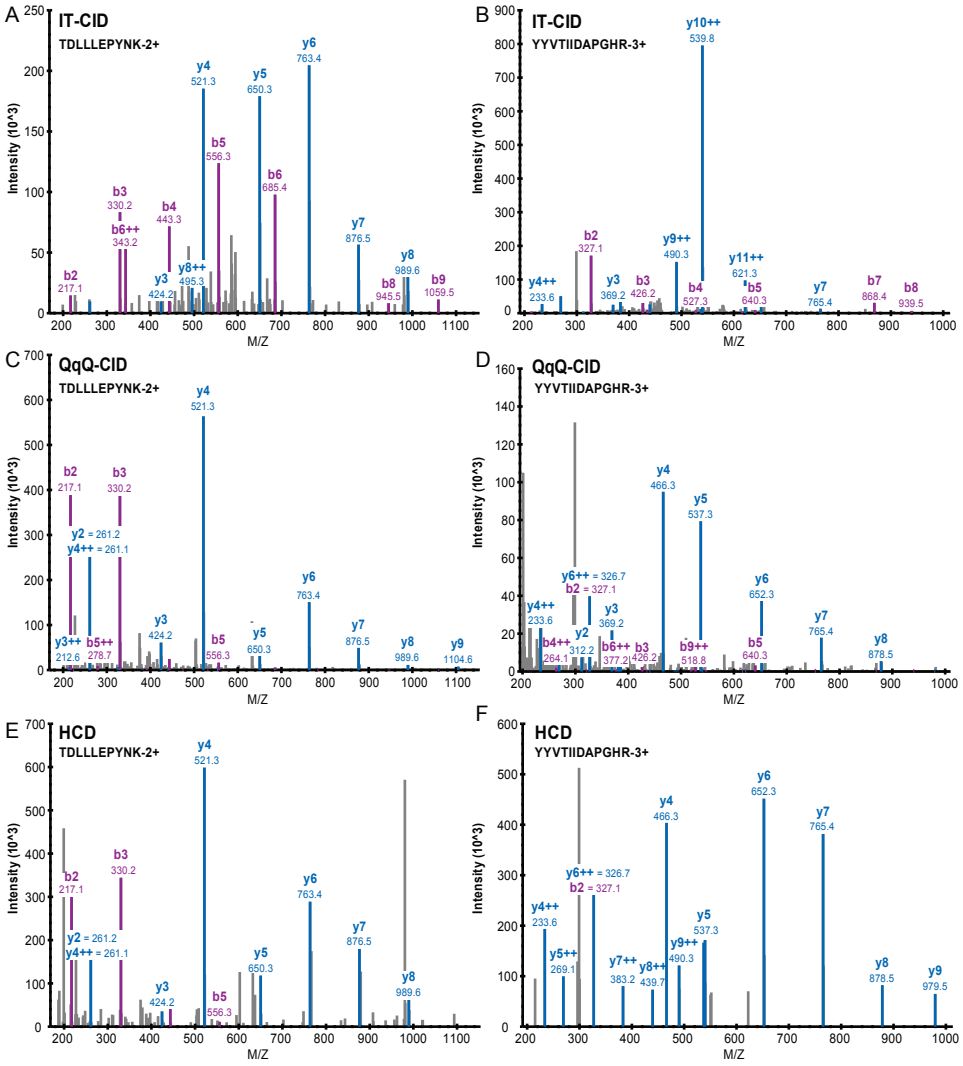

Figure 2: *Comparison of MS/MS spectra from different mass analyzer set-ups. Fragmentation pattern of the doubly charged peptide TDLLEPYNK and the triply charged peptide YYVTIIDAPGHR generated in an ion trap (A,B), triple quadrupole (C,D) and HCD cell (E,F).*

in spectra of less quality.

Statistical analysis of MS/MS spectra

The Xcorr plot results and visual inspection of individual spectra reveal clear patterns in similarity between spectra. Next, in order to assess the average degree of similarity of IT-CID and HCD with QqQ-CID for many spectra, the frequency of rank 1 peaks at ΔTh=0 were calculated for each fragmentation method. Initially, a χ^2 test was performed to investi-

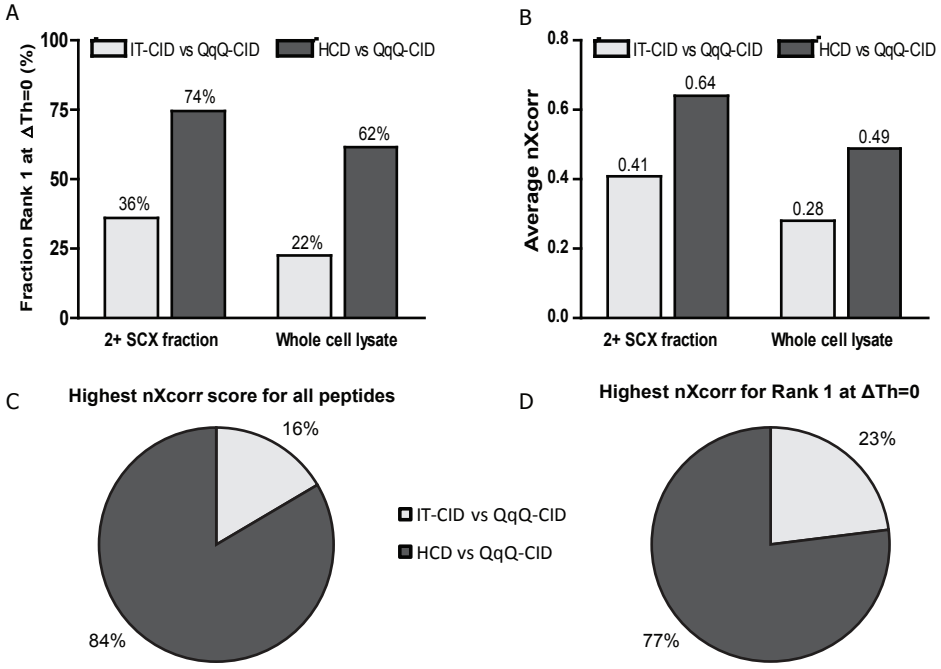

Figure 3: *IT-CID vs. QqQ-CID and HCD vs. QqQ-CID Xcorrelation comparison. A) The fraction of peptides in which the Xcorr calculation between IT-CID vs. QqQ-CID and HCD vs. QqQ-CID showed a maximum Xcorr score peak at a zero Thomson shift for the 2+ SCX fraction and whole cell lysate samples. B) Average Xcorr score normalized by the average TSQ self-correlation Xcorr score plotted for both peptide fragmentation methods and samples. Percentage of peptides that show a higher nXcorr score for HCD than IT-CID (red sector) and vice versa (blue sector) for (C) all peptides and (D) only those peptides with a rank 1 peak at zero Thomson shift.*

gate if the rank 1 distribution of the QqQ-CID versus HCD dataset is significantly different from the QqQ-CID versus IT-CID dataset. A p-value of 4.3E-63 indicated that the two datasets are significantly different. The fraction of total spectra in which the rank 1 peak resided at zero mass shift is largest for the QqQ-CID to HCD comparison (74% compared to 36% for IT-CID) validating the above proposed higher similarity between HCD and QqQ-CID (Figure 3A, Suppl. Table 1). When comparing the two samples, HCD outperforms IT-CID both on the fully tryptic peptides, as well as on the more heterogeneous peptide population originating from the whole cell lysate (61% rank 1 at ΔTh=0 for HCD against 22% for IT-CID). Contrary to expectation, both HCD and IT-CID were decreased equally in their QqQ-CID correlation. One explanation would be that the high sample complexity is pushing the QqQ-CID performance to its boundaries, resulting in noisier peptide fragmentation spectra.

Alongside the rank 1 distribution profiles, the Xcorr score is another important parameter for spectral correlation. Since the Xcorr scores vary between different peptides, we normalized the Xcorr scores using the score associated with the QqQ-CID autocorrelation. The average normalized Xcorr (nXcorr) score for HCD is higher than for IT-CID, indicating a higher similarity between the HCD and QqQ-CID spectra (Figure 3B, Suppl. Table 2).

Furthermore, around 80% of spectra generated from both samples showed a higher nXcorr score for HCD than IT-CID (Figure 3C and D, Suppl. Table 1).

SRM assay intensity comparison

In SRM assay development, selecting the most optimal ions and related transitions is of crucial importance. The higher similarity between HCD and QqQ-CID than IT-CID and QqQ-CID suggest that HCD libraries are more appropriate for defining transitions for SRM analysis. To verify this hypothesis, we built SRM assays based on either IT-CID or HCD spectra for 24 proteins (~450 transitions). After extracting the top 6 or 7 most intense b and y-ions from the spectra, the corresponding SRM transitions were measured on the triple quadrupole. Subsequently, the relative intensities of the peptide transitions were compared with their corresponding MS/MS library signals. The peptides used for this comparison are listed in Supplemental Table 2. The dotp value, calculated by Skyline(28), is a measure for complete orthogonality or identity between transition peak area and MS/MS library peak intensities, indicated by scores ranging from 0 to 1, respectively. For doubly charged peptides, the dotp values are slightly higher for HCD spectra compared to IT-CID spectra (Table 1, Suppl. Table 3). The high similarity between IT-CID and their SRM transitions (dotp 0.96) further supports the findings of Sherwood et al. in which the authors conclude that the y-ions from doubly charged peptides in IT-CID spectra correlate well with their SRM transitions.(28) However, triply charged peptides show a much lower relative intensity correlation in IT-CID when compared to HCD (0.80 versus 0.93, respectively) (Table 1). It is noteworthy that more than a third of all sequenced precursors from the HeLa cell lysate sample had a charge of 3+ (data not shown). Although doubly charged proteotypic peptides have been preferred for SRM assay development so far, these peptides might not always be accessible and therefore higher charged peptides might serve as valuable alternatives, indicating another possible advantage of using HCD MS/MS libraries.

In addition to a higher transition correlation, the HCD library also increased the average peak area of the SRM transitions (Table 1). In order to improve the signal/sensitivity of an SRM assay, collision energy optimization is often performed. Since collision energy opti-

Table 1: *Comparison of SRM assays built from IT-CID or HCD spectral libraries.*

	IT-CID Spectral Library		**HCD** Spectral Library	
	2+ peptides	3+ peptides	2+ peptides	3+ peptides
SRM *vs.* MS/MS Corr.[1]	0.96	0.8	0.98	0.93
Average Signal[2]	2.3E6	4.5E5	2.9E6	8.7E5
Norm. Avg. Signal[3]	1	1	1.26	1.93

[1]Average dot product between SRM transition peak area and MS/MS library peak intensity. [2]Average SRM peak area of peptides observed in both set-ups. [3]Average SRM peak area normalized to the peak area obtained by using the IT-CID spectral library-based assays.

mization adds significantly to instrument time and requires relatively large quantities of sample, this process is not always desirable. Our data suggest that picking the right transitions can already significantly increase the sensitivity of an SRM assay without the use of a labour intensive collision energy optimization procedure. The sensitivity was found, on average, to be increased two-fold by simply using the HCD library instead of the IT-CID library for triply charged peptides. In one specific protein example, we found two triply charged peptides for which one peptide showed a 1.6-fold signal increase whereas another showed a 6.1-fold increase. Therefore, the extent of signal improvement is most likely to be determined by the physicochemical properties of the peptide rather than the peptide abundance level. Moreover, we carried out the SRM analyses using a range of collision energies. The most optimum collision energy for each peptide transition, at best, gave an increase of only 8% over the original suggested collision energy which is in line with the observation of MacLean et al..(30) The total HCD spectral library generated in this work contains ~29,000 peptides and 6,200 proteins and has been made available as part of the supplementary information so that the community can assess the usefulness of HCD shotgun datasets for their SRM assay development.

Conclusion

By performing spectral comparisons, spectra cross correlation and relative intensity comparisons, we show both qualitatively and quantitatively that peptide fragmentation spectra generated on a triple quadrupole instrument (QqQ-CID) are much more similar to HCD than to conventional IT-CID fragmentation spectra. These results are confirmation of previously expressed expectations about the similarity of HCD and QqQ-CID.(11, 20, 21) We show that the HCD versus QqQ-CID spectral correlation is better than the IT-CID versus QqQ spectral correlation for doubly as well as for triply charged peptides. Additionally, we established that HCD spectral libraries are more beneficial than spectral libraries derived from IT-CID when building high-throughput data facilitated SRM assays. Our data also indicate that by using this approach, little or no instrument optimization is required. Thus, when using the above stated approach, more optimal SRM coordinates can be extracted while still saving significant instrument and analysis time.

Supporting Information

Supplementary Tables and network links to the raw data and scaffold data files can be accesed online at http://pubs.acs.org.

Acknowledgements

This research was performed partly within the framework of CTMM, the Center for Translational Molecular Medicine (www.ctmm.nl), project CIRCULATING CELLS (grant 01C-

102), and supported by the Netherlands Heart Foundation. Additionally, the Netherlands Proteomics Centre, a program embedded in the Netherlands Genomics Initiative, is kindly acknowledged for financial support. EdG and AJRH acknowledge support from the Centre for Biomedical Genetics. We would like to thank Rebecca Rose for critically evaluating the manuscript.

References

1. Kirkpatrick, D. S., Gerber, S. A., and Gygi, S. P. (2005) The absolute quantification strategy: a general procedure for the quantification of proteins and post-translational modifications. Methods 35, 265-273.

2. Wolf-Yadlin, A., Hautaniemi, S., Lauffenburger, D. A., and White, F. M. (2007) Multiple reaction monitoring for robust quantitative proteomic analysis of cellular signaling networks. Proc Natl Acad Sci U S A 104, 5860-5865.

3. Gerber, S. A., Rush, J., Stemman, O., Kirschner, M. W., and Gygi, S. P. (2003) Absolute quantification of proteins and phosphoproteins from cell lysates by tandem MS. Proc Natl Acad Sci U S A 100, 6940-6945.

4. Craig, R., Cortens, J. P., and Beavis, R. C. (2004) Open source system for analyzing, validating, and storing protein identification data. J Proteome Res 3, 1234-1242.

5. Craig, R., Cortens, J. P., and Beavis, R. C. (2005) The use of proteotypic peptide libraries for protein identification. Rapid Commun Mass Spectrom 19, 1844-1850.

6. Desiere, F., Deutsch, E. W., King, N. L., Nesvizhskii, A. I., Mallick, P., Eng, J., Chen, S., Eddes, J., Loevenich, S. N., and Aebersold, R. (2006) The PeptideAtlas project. Nucleic Acids Res 34, D655-658.

7. Prince, J. T., Carlson, M. W., Wang, R., Lu, P., and Marcotte, E. M. (2004) The need for a public proteomics repository. Nat Biotechnol 22, 471-472.

8. Frewen, B. E., Merrihew, G. E., Wu, C. C., Noble, W. S., and MacCoss, M. J. (2006) Analysis of peptide MS/MS spectra from large-scale proteomics experiments using spectrum libraries. Anal Chem 78, 5678-5684.

9. Prakash, A., Tomazela, D. M., Frewen, B., Maclean, B., Merrihew, G., Peterman, S., and Maccoss, M. J. (2009) Expediting the development of targeted SRM assays: using data from shotgun proteomics to automate method development. J Proteome Res 8, 2733-2739.

10. Xia, Y., Liang, X., and McLuckey, S. A. (2006) Ion trap versus low-energy beam-type collision-induced dissociation of protonated ubiquitin ions. Anal Chem 78, 1218-1227.

11. Olsen, J. V., Macek, B., Lange, O., Makarov, A., Horning, S., and Mann, M. (2007) Higher-energy C-trap dissociation for peptide modification analysis. Nat Methods 4, 709-712.

12. Paizs, B., and Suhai, S. (2005) Fragmentation pathways of protonated peptides. Mass Spectrom Rev 24, 508-548.

13. Lau, K. W., Hart, S. R., Lynch, J. A., Wong, S. C., Hubbard, S. J., and Gaskell, S. J. (2009) Observations on the detection of b- and y-type ions in the collisionally activated decomposition spectra of protonated peptides. Rapid Commun Mass Spectrom 23, 1508-1514.

14. Morris, M., Thibault, P., and Boyd, R. K. (1994) Characterization of a high-pressure quadrupole collision cell for low-energy collision-induced dissociation. J Am Soc Mass Spectrom 5, 1042-1063.

15. Picotti, P., Rinner, O., Stallmach, R., Dautel, F., Farrah, T., Domon, B., Wenschuh, H., and Aebersold, R. (2010) High-throughput generation of selected reaction-monitoring assays for proteins and proteomes. Nat Methods 7, 43-46.

16. Picotti, P., Bodenmiller, B., Mueller, L. N., Domon, B., and Aebersold, R. (2009) Full dynamic range proteome analysis of S. cerevisiae by targeted proteomics. Cell 138, 795-806.

17. Frese, C. K., Altelaar, A. F. M., Hennrich, M. L., Nolting, D., Zeller, M., Griep-Raming, J., Heck, A. J. R., and Mohammed, S. (2011) Improved Peptide Identification by Targeted Fragmentation Using CID, HCD and ETD on an LTQ-Orbitrap Velos. J Proteome Res.

18. McAlister, G. C., Phanstiel, D., Wenger, C. D., Lee, M. V., and Coon, J. J. (2010) Analysis of tandem mass spectra by FTMS for improved large-scale proteomics with superior protein quantification. Anal Chem 82, 316-322.

19. Nagaraj, N., D'Souza, R. C., Cox, J., Olsen, J. V., and Mann, M. (2010) Feasibility of Large-Scale Phosphoproteomics with Higher Energy Collisional Dissociation Fragmentation. J Proteome Res.

20. Olsen, J. V., Schwartz, J. C., Griep-Raming, J., Nielsen, M. L., Damoc, E., Denisov, E., Lange, O., Remes, P., Taylor, D., Splendore, M., Wouters, E. R., Senko, M., Makarov, A., Mann, M., and Horning, S. (2009) A dual pressure linear ion trap Orbitrap instrument with very high sequencing speed. Mol Cell Proteomics 8, 2759-2769.

21. Zhang, Y., Ficarro, S. B., Li, S., and Marto, J. A. (2009) Optimized Orbitrap HCD for quantitative analysis of phosphopeptides. J Am Soc Mass Spectrom 20, 1425-1434.

22. Gauci, S., Helbig, A. O., Slijper, M., Krijgsveld, J., Heck, A. J., and Mohammed, S. (2009) Lys-N and trypsin cover complementary parts of the phosphoproteome in a refined SCX-based approach. Anal Chem 81, 4493-4501.

23. Pinkse, M. W., Mohammed, S., Gouw, J. W., van Breukelen, B., Vos, H. R., and Heck, A. J. (2008) Highly robust, automated, and sensitive online TiO2-based phosphoproteomics applied to study endogenous phosphorylation in Drosophila melanogaster. J Proteome Res 7, 687-697.

24. MacLean, B., Tomazela, D. M., Shulman, N., Chambers, M., Finney, G. L., Frewen, B., Kern, R., Tabb, D. L., Liebler, D. C., and MacCoss, M. J. (2010) Skyline: an open source document editor for creating and analyzing targeted proteomics experiments. Bioinformatics 26, 966-968.

25. Mohammed, S., and Heck, A. J. (2011) Strong cation exchange (SCX) based analytical methods for the targeted analysis of protein post-translational modifications. Curr Opin Biotechnol 22, 9-16.

26. Powell, L. A., and Hieftje, G. M. (1978) Computer identification of infrared spectra by correlation-based file searching. Anal Chem Acta 100, 313-327.

27. Eng, J. K., McCormack, A. L., and Yates, J. R., 3rd (1994) An Approach to Correlate Tandem Mass Spectral Data of Peptides with Amino Acid Sequences in a Protein Database. J Am Soc Mass Spectrom 5, 976-989.

28. Sherwood, C. A., Eastham, A., Lee, L. W., Risler, J., Vitek, O., and Martin, D. B. (2009) Correlation between y-type ions observed in ion trap and triple quadrupole mass spectrometers. J Proteome Res 8, 4243-4251.

29. Yates, J. R., 3rd, Morgan, S. F., Gatlin, C. L., Griffin, P. R., and Eng, J. K. (1998) Method to compare collision-induced dissociation spectra of peptides: potential for library searching and subtractive analysis. Anal Chem 70, 3557-3565.

30. Maclean, B., Tomazela, D. M., Abbatiello, S. E., Zhang, S., Whiteaker, J. R., Paulovich, A. G., Carr, S. A., and Maccoss, M. J. (2010) Effect of collision energy optimization on the measurement of peptides by selected reaction monitoring (SRM) mass spectrometry. Anal Chem 82, 10116-10124.

Probing phosphorylation dynamics in signaling networks exploiting Ti^{4+}-IMAC phosphopeptide enrichment

Erik L. de Graaf[1,2,#], Piero Giansanti[1,2,#], Albert J. R. Heck[1,2] and A. F. Maarten Altelaar[1,2]

Manuscript in submission

[1] Biomolecular Mass Spectrometry and Proteomics, Bijvoet Center for Biomolecular Research and Utrecht Institute for Pharmaceutical Sciences, Utrecht University, Padualaan 8, 3584 CH Utrecht, The Netherlands

[2] Netherlands Proteomics Centre, Padualaan 8, 3584 CH Utrecht, The Netherlands

[#] These authors contributed equally to this work.

Abstract

In this study we assessed the quantitative reproducibility of phosphopeptide enrichment in large-scale experiments. First we combined high-resolution mass spectrometry with Ti^{4+}-IMAC phosphopeptide enrichment and label-free quantification to assess column enrichment reproducibility in HeLa cells. Secondly, we monitored the phosphoproteome of Jurkat T-cells stimulated by Prostaglandin E2 over 6 time points. In total, we quantitatively monitored 12,799 unique phosphosites throughout all 6 time points, using only 200 μg starting material per enrichment. The extraordinary efficient, qualitative and quantitative reproducibility (average r>0.9), together with a high sensitivity, demonstrate the feasibility to comprehensively chart phosphorylation changes in signalling networks in a large set of samples.

Introduction

Cellular signaling cascades control regulatory mechanisms responsible for cell functioning, e.g. survival, proliferation and death. Phosphorylation is the most studied post-translational modification (PTM), involved in practically all regulatory processes. The dynamics of phosphorylation is nowadays preferably studied by mass spectrometry (MS)-based proteomics,(1-3) which allows the confident identification of many thousands of phosphorylation sites in a single experiment. As phosphoproteins, and their corresponding peptides after enzymatic digestion, are present at much lower abundances than their unmodified counterparts, enrichment of the phosphopeptide population is a necessity. Methodologies such as immobilized metal ion affinity chromatography (IMAC), titanium dioxide (TiO_2) chromatography, immunoprecipitation, or, more recently, immobilized titanium-IMAC (Ti^{4+}-IMAC) based enrichment have enabled the global analysis of phosphorylation in diverse settings and disease states.(4-9)

Quantification of phosphorylation dynamics poses an additional challenge and often includes labeling of samples with heavy isotopes to prevent biases introduced by sample handling. A popular labeling strategy is SILAC, however this approach is somewhat limited to model systems and provides restricted multiplexing capability. Chemical labeling with isobaric tags (e.g. TMT or iTRAQ) allows multiplexing of up to six or eight channels, however suffers from ratio compression.(10, 11) Recent studies report alternative ways of multiplexing relying on high resolution MS or combinations of labeling strategies, although these methods remain to prove their generic practicality.(12, 13) More straightforward is the label-free quantification approach, which alleviates the restriction on number of samples that can be quantitatively monitored, and can reach a very high dynamic range in quantification.

Label-free quantification requires careful experimental design to prevent introduction of sample biases.(3) This is especially true for the analysis of PTMs, requiring the enrichment step to be reproducible both qualitative and quantitative. Therefore, in this paper we assessed the qualitative and quantitative reproducibility of Ti^{4+}-IMAC phosphopeptide enrichment followed by a large scale proof of principle study in Jurkat T-cells stimulated with Prostaglandin E2 (PGE_2) for different incucation times. The immune regulatory actions of T-cells are mediated by auto- and paracrine PGE_2 secreted by many different cell types including tumor cells. (14, 15) When PGE_2 binds to G protein coupled receptors the secondary messenger cyclic adenosine monophosphate (cAMP) is generated leading to the activation of intracellular signal transduction cascades involving protein kinase A (PKA). The quantitative dataset presented in this work also provides a useful resource of dynamic phosphorylation events that occur after T-cell activation with PGE_2 and PKA activation in general.

Experimental Procedures

Cell Culture and Digest Preparation - HeLa cells were grown to confluence in Dulbecco's modified Eagle's medium supplemented with 10% fetal bovine serum and penicillin/strep-

tomycin (Lonza), then cells were washed twice with PBS and harvested. Jurkat T lymphoma cells were grown in RPMI 1640 medium supplemented with 10% fetal bovine serum and penicillin/streptomycin (Lonza). Before PGE_2 stimulation, cells were centrifuged for 1 min at 1500g, growth medium was removed and the cells were resuspended at a final concentration of 1-2 × 106 cells/ml in RMPI. Next, cells were supplemented with 0 (control) or 10 μM PGE_2 and incubated for 5, 10, 20, 30 or 60 min. After treatment Jurkat cells were washed twice with PBS and harvested. Cell lysis was performed on ice by sonication in buffer containing 50 mM ammonium bicarbonate (pH 8.0), 8 M urea, 1 mM sodium orthovanadate, complete EDTA-free protease inhibitor mixture (Roche) and phosSTOP phosphatase inhibitor mixture (Roche). Cell debris were then removed by centrifugation at 20 000g for 15 min at 4 °C. The total protein concentration was measured using a Bradford Assay (BioRad) and then split into 1 mg aliquots for enzymatic digestion. Proteins were reduced with DTT at a final concentration of 4 mM at 56 °C for 25 min; subsequently samples were alkylated with iodoacetamide at a final concentration of 8 mM at RT for 30 min in the dark. Proteins were then digested using Lys-C (1 μg Lys-C per 75 μg protein) and incubated for 4 h at 37 °C. The solution was then diluted to a final urea concentration of 2 M with 50 mM ammonium bicarbonate, and trypsin was added (1 μg trypsin per 100 μg protein) and incubated at 37 °C overnight. The digestion was quenched by acidification to 5% formic acid. The digests were desalted using Sep-Pak C18 cartridges, dried in vacuo and stored at −80 °C for further use.

Phosphopeptides enrichment by Ti^{4+}-IMAC - Ti^{4+}-IMAC material was prepared and used essentially as previously described.(9) Briefly, the Ti4+-IMAC beads (500 μg of beads/200 μL pipet tip) were loaded onto GELoader tips (Eppendorf) using a C8 plug. To reduce variations of enrichment processes, in parallel spin tip enrichment was used. The Ti4+-IMAC columns were conditioned using 50 μL of loading buffer consisting of 80% acetonitrile (ACN)/6% trifluoroacetic acid (TFA) and centrifugation at 200g for 10 min. The protein digests were dissolved in 80% ACN/6% TFA and split in aliquots corresponding to ~100 μg and ~250 μg for the HeLa and Jurkat cell lysates, respectively. The aliquots were transferred to the spin tips and centrifuged at 100g for 30 min. Then, the columns were sequentially washed with 50 μL of washing buffer 1 (50% ACN, 0.5% TFA containing 200 mM NaCl) and additional washing with 50 μL of buffer 2 of 50% ACN/0.1% TFA, each centrifuged at 170g for 15 min, respectively. The bound peptides were eluted into a new tube (already containing 35 μL of 10% formic acid) with 20 μL of 10% ammonia by centrifugation at 100g for 20 min. A final elution was performed with 5 μL of 80% ACN/2% formic acid for 10 min. The collected eluate was further acidified by adding 3 μL of 100% formic acid prior to nLC-MSMS analysis.

Reverse phase chromatography and mass spectrometry - Peptides were subjected to reversed phase nLC–MSMS analysis using a Proxeon EASY-nLC 1000 (Thermo Scientific, Odense, Denmark) with an analytical column heater (40°C) and an LTQ-Orbitrap Elite (Thermo Fisher Scientific, Bremen, Germany). Peptides were first trapped (Dr Maisch Reprosil C18, 3 μm, 2 cm x 100 μm) at a maximum pressure of 800 bar with 100% solvent A (0.1 % formic acid in water) before being separated on the analytical column (either Agilent Poroshell 120 EC-C18, 2.7 μm, 40 cm x 50 μm for the HeLa cell samples or Agilent Zorbax SB-C18, 1.8 μm, 40 cm x 75 μm for the Jurkat T-cell samples). Peptides were chromatograph-

ically separated by a 150 min gradient from 7% to 30% solvent B (0.1% formic acid in ACN) at a flowrate of 150 or 100 nL/min. The total measurement time for each sample was 180 min. The eluent was sprayed via a distal coated fused silica emitter (360 μm o.d., 20 μm i.d., 10 μm tip i.d.; constructed in-house) butt-connected to the analytical column. The electrospray voltage was set to 1.7 kV. The mass spectrometer was operated in a data-dependent mode to automatically switch between MS and MS/MS. Briefly, survey full-scan MS spectra were acquired in the Orbitrap analyzer, scanning from m/z 350 to m/z 1500 at a resolution of 60,000 at m/z 400 using an AGC setting of 1e6 ions. Charge state screening was enabled and precursors with either unknown or 1+ charge states were excluded. After the survey scan the 20 most intense precursors were selected for subsequent decision tree-based iontrap CID or ETD fragmentation.(16) The normalized collision energy for CID was set to 35% and supplemental activation for ETD and dynamic exclusion were enabled (exclusion size list 500, exclusion duration 40 s).

Data processing **-** Out of 108 nLC-MSMS runs 8 runs were not consistent with the other replicates and were discarded prior to data analysis. Raw data were processed with MaxQuant version 1.3.0.5,(17) and MS and MSMS spectra were searched against a concatenated forward-decoy Swissprot Homo sapiens database version 2012_09 (40,992 sequences) using the Andromeda search engine. The database search was performed with the following parameters: an initial mass tolerance of ±20 ppm for precursor masses; final mass tolerance of ±6 ppm, ±0.6 Da for CID and ETD ion trap fragment ions, allowing two missed cleavages. Cysteine carbamidomethylation was used as a fixed modification and methionine oxidation, protein N-terminal acetylation and serine, threonine and tyrosine phosphorylation as variable modifications. The false discovery rate was set to 0.01 for peptides, proteins and phosphosites, the minimum peptide length allowed was six amino acids and a minimum Andromeda peptide score of 60 was required. The match between run feature was enabled. A site localization probability of at least 0.75 and a score difference of at least 5 were used as threshold for the localization of phosphoresidues. Normalization was performed by subtracting the median of log transformed intensities for each nLC-MSMS run. To identify significantly regulated phosphorylation sites a two sample t-test was performed with a permutation-based FDR of 0.005 (randomizations were set to 500 while s0 was tuned to achieve a min 2-fold regulation and varied between 0.35-0.4). Phosphorylation sites regulated in at least one of all fifteen possible comparisons were included for soft clustering analysis using GProX version 1.1.9.(18) Log base 2 intensities were Z-scored prior to fuzzy clustering using fuzzification value 2, 100 iterations and a minimum membership of 0.35. The number of clusters was empirically derived by varying the number between 4 and 10, where 8 clusters returned the most distinctive clusters. The high confident localized phosphorylation sites were further analyzed using iceLogo(19) using a p-value of 0.01 and as background the entire human proteome.

Results and Discussion

To test the feasibility of our recently described Ti^{4+}-IMAC enrichment strategy(9) for label-free quantification of phosphorylation dynamics, we first interrogated its specificity and reproducibility for consecutive enrichments in HeLa cells. Our data from three independent

Ti^{4+}-IMAC enrichments clearly shows a high specificity and qualitative reproducibility (Fig. 1A and 1B, respectively). Virtually all peptides observed are phosphopeptides (>98%), and the overlap of observed phosphopeptides between independent enrichments is similar to technical replicate nLC-MSMS runs. Importantly, the quantitative reproducibility between the different enrichment replicates is likewise satisfactory (Fig. 1C, Supplementary Fig. 1),

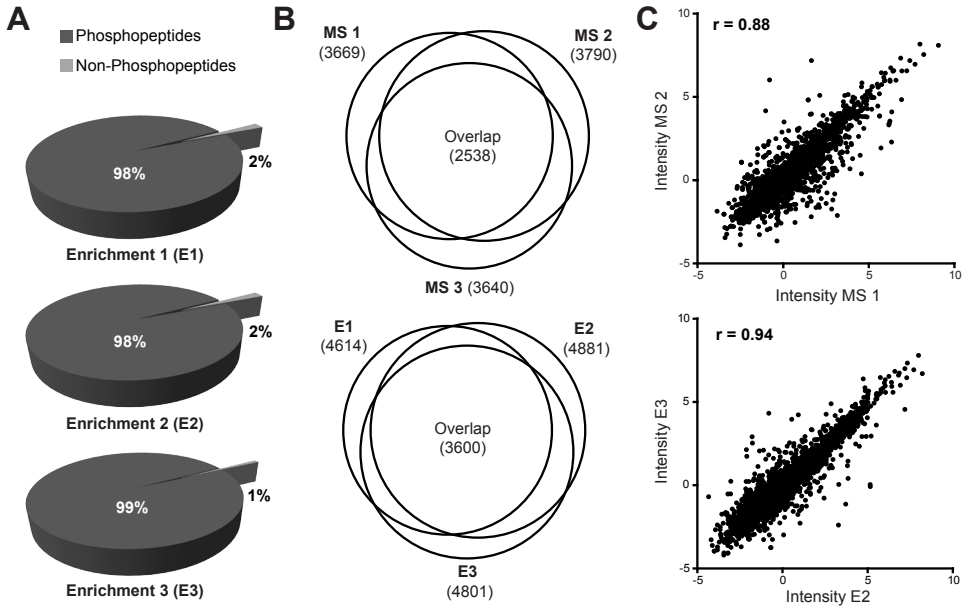

Figure 1: *High specificity and reproducibility of Ti^{4+}-IMAC phosphopeptide enrichment in HeLa cells. A) High percentage of phosphopeptides in three separate enrichments indicating a high specificity of the Ti^{4+}-IMAC enrichment procedure. B) The overlap of identified phosphopeptides in three MS replicates from a single enrichment (MS1-3, upper panel) is similar to the overlap of three separate enrichments (E1-3, lower panel), indicating a highly reproducible phosphopeptide enrichment. C) Phosphopeptide intensity (log base 2) comparisons for MS replicates (upper panel) and enrichment replicates (lower panel), showing high quantitative reproducibility of the Ti^{4+}-IMAC enrichment method.*

with an average pearson correlation of 0.92 and a median CV of 0.2, evidently meeting the above described constraint for label-free quantification of PTMs.

Jurkat T-cell phosphoproteome dynamics

After demonstrating the reproducibility of the enrichment strategy we assessed the potential of the approach in studying large scale phosphorylation dynamics covering extended time series. Therefore, we monitored the temporal phosphorylation signaling of Jurkat T-cells upon stimulation of the G protein coupled receptors, EP1 to EP4, with their ligand PGE_2, thereby affecting several intracellular signaling cascades, including the cAMP/PKA as well as the PI3K-dependent ERK1/2 pathways.(20)

Jurkat cells were grown under normal conditions and samples were harvested after 0, 5, 10,

20, 30 and 60 minutes of PGE_2 stimulation. For every time point three biological replicates were harvested, cells were lysed and proteins digested with Lys-C and Trypsin. Phosphopeptides were enriched using for every biological replicate three independent Ti^{4+}-IMAC enrichment columns, requiring 200 µg of protein digest as input material each. All 54 samples (6 time points, 3 biological replicates, 3 Ti^{4+}-IMAC enrichments, Fig. 2A) were analyzed twice by nLC-MSMS using 2h gradients and decision-tree guided peptide fragmentation,

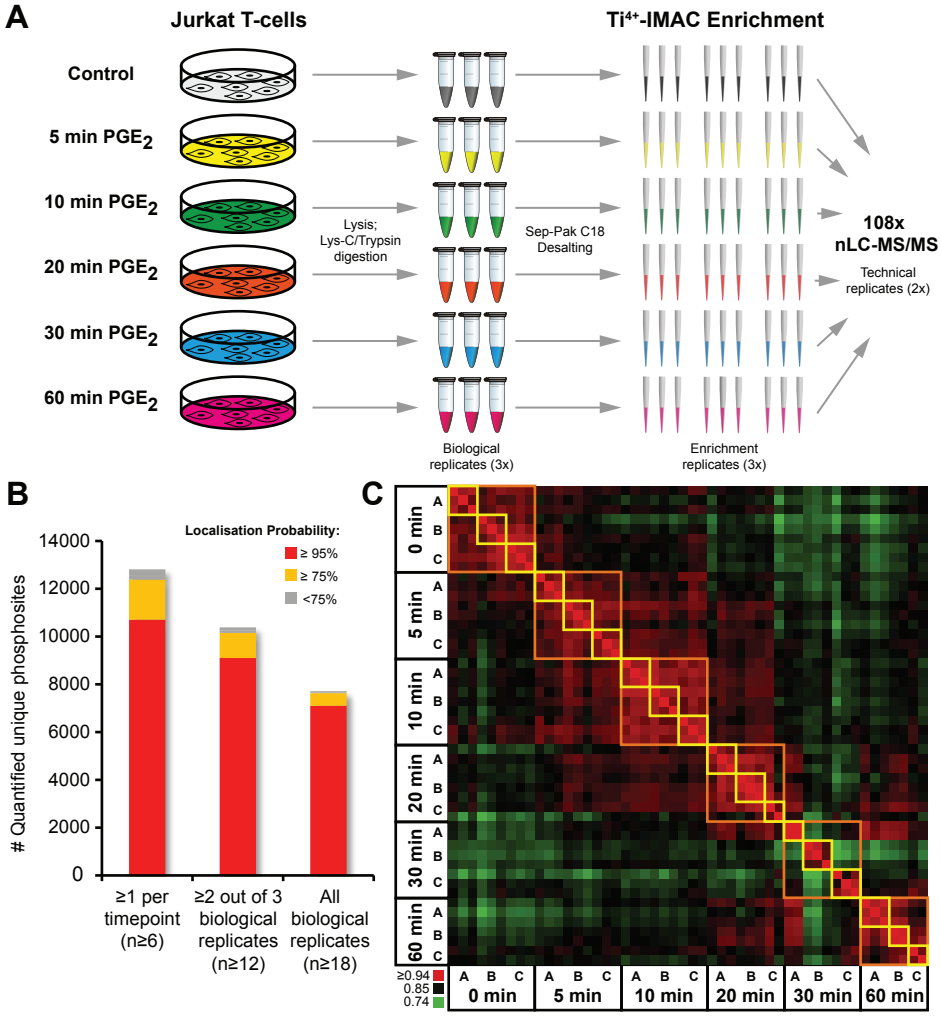

Figure 2: *High-throughput label-free phosphoproteomics. A) Three biological replicates of Jurkat T- lymphocyte cells stimulated with Prostaglandin E2 (PGE_2) and harvested at 6 different time points, are lysed, digested and desalted prior to phosphopeptide enrichment. Each biological replicate was divided over three Ti^{4+}-IMAC phosphopeptide enrichment columns, resulting in 54 samples that were analyzed twice by 2h nLC-MSMS with CID/ETD fragmentation, accumulating to a total of 108 phosphoproteomic datasets. B) Number of unique phosphosites quantified using different cutoff criteria. Between brackets is the minimum number of quantification events. C) Pearson Intensity correlation heat map of all successful runs showing a high reproducibility in phosphosite intensity for the phosphopeptide enrichment replicates (yellow squares) and high similarity between biological replicates A-C (orange squares).*

switching between CID and ETD based on peptide physiochemical properties.(16, 21) This resulted in a total of 108 nLC-MSMS runs, of which 95% were successful, compromising 216 hours of MS analysis time, identifying cumulatively 21,443 unique phosphosites from fifty Ti^{4+}-IMAC enrichments columns.

Of these 21,443 unique phosphopeptides we could quantitatively follow the presence of 12,799 over all time points and 10,375 in at least 2 out of 3 biological replicates (Fig. 2B). To evaluate the quality of the complete experiment we evaluated the correlation and median CV between all Ti^{4+}-IMAC enrichments of the different biological replicates and performed a Pearson's correlation of all phosphopeptide intensities of the successful nLC-MSMS runs, plotted in a heat map (Table 1, Fig. 2C). This heat map illustrates the reproducibility in phosphosite intensities between the different enrichment replicates (yellow squares) and the similarity within the three biological replicates (A, B and C, orange squares). The time point showing the highest biological similarity is 10 min. after activation while the highest variability was observed after 30 min.

Table 1: *Phosphopeptide enrichment triplicate reproducibility*

	Average Pearson Intensity Correlation						**Median Coefficient of Variation**					
Sample	0min	5min	10min	20min	30min	60min	0min	5min	10min	20min	30min	60min
BioA	0.92	0.92	0.93	0.90	0.94	0.94	0.23	0.20	0.21	0.18	0.14	0.18
BioB	0.92	0.92	0.93	0.94	0.91	0.94	0.21	0.21	0.21	0.17	0.18	0.14
BioC	0.94	0.92	0.93	0.90	0.90	0.93	0.19	0.21	0.18	0.16	0.16	0.18

The robustness of our approach allowed us to investigate the dynamics of all phosphosites substantially regulated in at least one of the 6 time points after stimulation by PGE_2. The significance of regulation was determined by conducting a two sample t-test between all the time points using at least three measurements per phosphosite (Fig. 3A). We observed that although a large number of phosphosites are regulated directly upon stimulation with PGE_2 (5 minutes), the majority of phosphosite regulation occurs at later time, hinting at amplified phosphorylation downstream in the triggered signaling cascades (Figure 3A; 5min and 60min, Supplementary Fig. 2). Of all the phosphosites that were quantified in 2 out of 3 biological replicates, 2,983 were significantly regulated and subsequently analyzed for common patterns. Unsupervised clustering of the regulated phosphosites resulted in 8 clusters of distinct temporal dynamics, with 5 clusters displaying patterns related to up regulated phosphosite abundance and 3 clusters displaying primarily down regulation (Fig. 3B). Interestingly, the different clusters show a clear distinction in enriched sequence motifs, portrayed by IceLogo(22) (Fig. 3C and Supplementary Fig. 3), related to distinct kinase activity. Cluster 1, which shows a strong up regulation at the early time point followed by a steady decrease of phosphosite abundance after 5 minutes, represents a clear basophilic kinase motif (e.g. PKA) complemented by the MAPK and CDK motifs. When we look in more detail at the sequence motifs present in this cluster, using motif-X,(23) we can distinguish several distinct sub motifs that are significantly enriched and mutually exclusive (Fig. 3D). These motifs represent the different kinase families containing either proline-directed

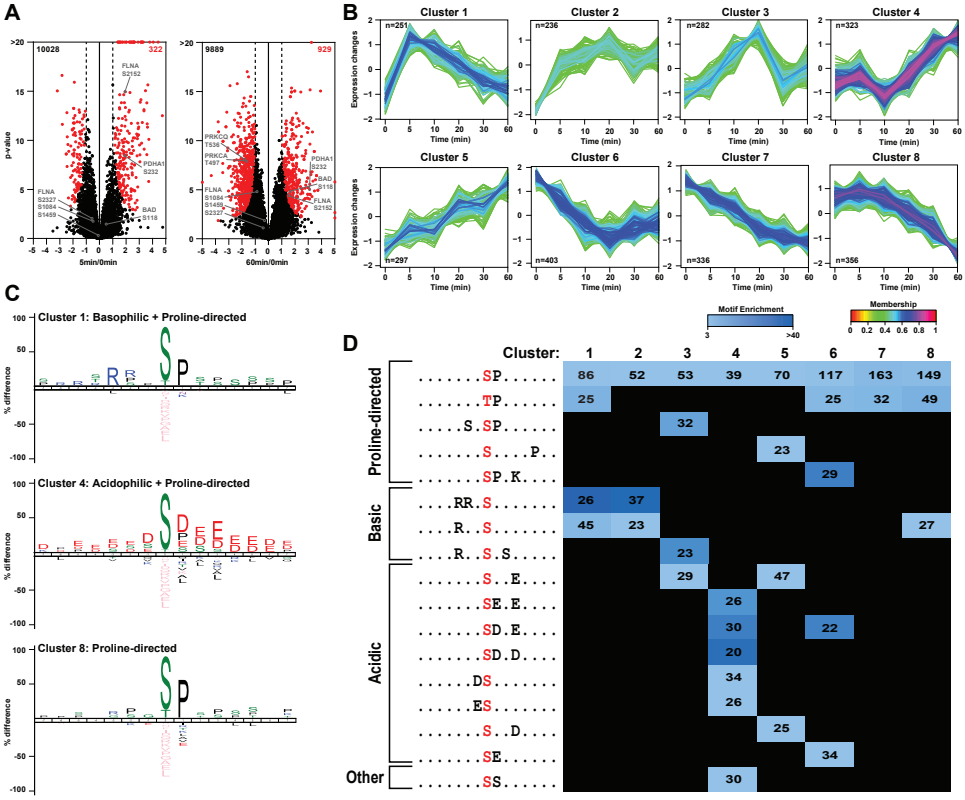

Figure 3: *Analysis of differentially regulated phosphosites. A) Volcano plots of treatment versus control comparisons (two sample t-test). P-values (-log base 10) are plotted as a function of the phosphosite ratio (log base 2) for 5 and 60 min versus 0 min. of PGE_2 treatment. Regulated sites are colored in red (permutation based FDR=0.005, s0 adjusted for 2-fold regulation). B) Soft clustering analysis of significantly regulated phosphosites resulted in 5 differently up regulated and 3 differently down regulated clusters. C) Sequence motifs showing amino acid frequencies flanking regulated phosphosites that are significantly different compared to the reference background (Swiss-Prot human database, P < 0.01) using IceLogo. Early up regulated sites contain basophilic motifs, late up regulated sites acidophilic motifs whereas down regulated sites contain only proline directed consensus sites. D) Linear kinase motifs enriched in each cluster by Motif-X analysis. Significantly overrepresented kinase motifs were determined by querying the sequence motifs of the phosphorylation sites against the IPI human database, using a p-value of E-6, a minimum number of occurrences of 20 and a minimum 3.0-fold enrichment compared to background.*

kinases such as CDK, MAPK, etc., or the more basophilic protein kinase A, C and G families (PKA, PKG, PKC).(24) Cluster 4 contains 323 phosphosites that respond at a later stage (after 20 minutes) to the given stimulus and might therefore represent downstream targets. In the enriched motifs in cluster 4 (Fig. 3C and D), there is a strong preference for acidic amino acids flanking the phosphorylation site, corresponding to a casein kinase 2 (CK2) family motif. Finally, the clusters showing mainly down regulation, such as cluster 8, primarily contain the proline directed motif.

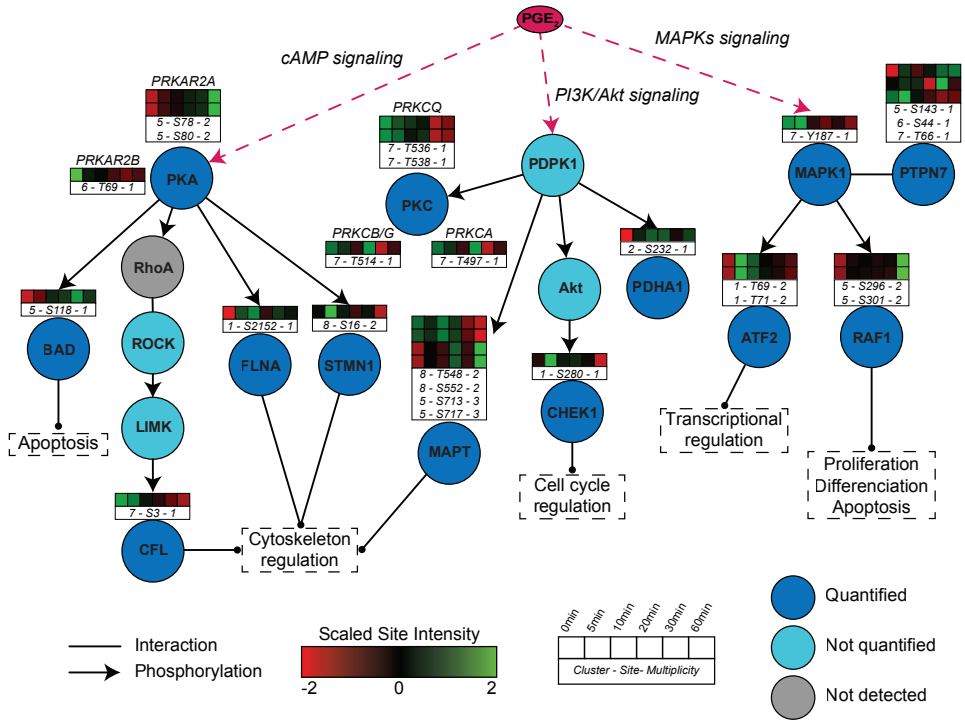

Figure 4: *PGE_2 downstream signaling; early, intermediate and late responders. Arrows represent known phosphorylation events between kinases and their substrates. The network was manually curated using the reference databases UniProt (http://www.uniprot.org/) and PhosphoSitePlus (http://www.phosphosite.org).*

Our data presents a comprehensive resource for PKA, MAPK and PI3K/Akt directed phosphorylation. Although details on substrate specific phosphorylation sites are still scarce, we can define early, intermediate and late responders to PGE_2 stimulation, manually curated from the online resources UniProt Knowledgebase and phosphositeplus (Fig. 4).(25) PGE_2 is known to increase cellular cAMP levels, activating PKA and initiating phosphorylating of its substrates. This activation is apparent in our data through several examples, such as the early up regulation of S2152 on FLNA and S16 on STMN1, which are direct substrates of PKA, and have functional implications in cytoskeleton reorganization. More downstream substrates were also identified, such as S3 on CFL, which shows decreased phosphosite abundance in the late time points because of the inactivation of intermediate substrates RhoA, ROCK and LIMK. Interestingly, the abundance of the phosphosite S118 on BAD, a direct PKA substrate, is gradually up regulated at later time points, indicating additional indirect regulatory mechanisms.

The MAPK and PI3K/Akt signaling pathways are also represented among the regulated phosphosites. MAPK1/ERK2 can be activated upon increase of cAMP levels through its phosphorylation by the upstream kinase activators (MEK1 and MEK2) and the dissociation of MAPK1 from phosphorylated PTPN7, a protein phosphatase that acts preferentially on tyrosine-phosphorylated MAPK1. This results in the activation of MAPK1, and we see its

downstream substrates ATF2 phosphorylated at T71, which nicely shows a slightly delayed increase in phosphosite abundance. Moreover, the negative feedback phosphorylation of MAPK1 on RAF1 (S296 and S301) yielding an inactive, desensitized kinase, was also profiled. The PI3K/Akt pathway is activated through PGE_2 stimulation of PDK1. This kinase activates Akt signaling (site was detected but not quantified), which is observed downstream through phosphorylation of CHEK1 at S280.

Conclusion

Overall our work demonstrates the high-potential of combining Ti^{4+}-IMAC phosphopeptide enrichment, high resolution nLC-MSMS and label-free quantification to interrogate in a comprehensive manner temporal phosphosite abundance regulation. The presented method is highly robust, specific and reproducible, both qualitatively and quantitatively. The quality of the data and strength of the approach was validated by previously curated phosphosite regulations.

The presented dataset encompasses a rich resource of downstream PGE_2 signaling dynamics in T-Cells and PKA signaling dynamics in general, providing a valuable resource for researchers working on T-cell and/or cAMP signaling.

Supporting Information

Supplementary Figures are located in Chapter 8; Supplementary Figures to Chapter 4. Also, all raw mass spectrometry data files and MaxQuant output files have been deposited to the ProteomeXchange Consortium (http://proteomecentral.proteomexchange.org) via the PRIDE partner repository(26) with the following dataset identifier: PXD000293.

Acknowledgements

This work was supported by The Netherlands Proteomics Centre, The Netherlands Genomics Initiative, and the Centre for Biomedical Genetics as well as the Netherlands Organization for Scientific Research (NWO) with the VIDI grant for AFMA (723.012.102). Additionally, the PRIME-XS Project, Grant Agreement 262067, funded by the European Union Seventh Framework Program, is kindly acknowledged for financial support. We thank Shabaz Mohammed, Houjiang Zhou and Serena Di Palma for fruitful discussion and help with the Ti4+-IMAC enrichment protocol.

References

1. Altelaar, A. F., Munoz, J., and Heck, A. J. (2013) Next-generation proteomics: towards an integrative view of proteome dynamics. Nat Rev Genet 14, 35-48.

2. Cox, J., and Mann, M. (2011) Quantitative, high-resolution proteomics for data-driven systems biology. Annu Rev Biochem 80, 273-299.

3. Zhang, Y., Fonslow, B. R., Shan, B., Baek, M. C., and Yates, J. R., 3rd (2013) Protein analysis by shotgun/bottom-up proteomics. Chem Rev 113, 2343-2394.

4. Pinkse, M. W., Uitto, P. M., Hilhorst, M. J., Ooms, B., and Heck, A. J. (2004) Selective isolation at the femtomole level of phosphopeptides from proteolytic digests using 2D-NanoLC-ESI-MS/MS and titanium oxide precolumns. Analytical chemistry 76, 3935-3943.

5. Villen, J., and Gygi, S. P. (2008) The SCX/IMAC enrichment approach for global phosphorylation analysis by mass spectrometry. Nat Protoc 3, 1630-1638.

6. Huang, P. H., and White, F. M. (2008) Phosphoproteomics: unraveling the signaling web. Mol Cell 31, 777-781.

7. Huttlin, E. L., Jedrychowski, M. P., Elias, J. E., Goswami, T., Rad, R., Beausoleil, S. A., Villen, J., Haas, W., Sowa, M. E., and Gygi, S. P. (2010) A tissue-specific atlas of mouse protein phosphorylation and expression. Cell 143, 1174-1189.

8. Bodenmiller, B., Wanka, S., Kraft, C., Urban, J., Campbell, D., Pedrioli, P. G., Gerrits, B., Picotti, P., Lam, H., Vitek, O., Brusniak, M. Y., Roschitzki, B., Zhang, C., Shokat, K. M., Schlapbach, R., Colman-Lerner, A., Nolan, G. P., Nesvizhskii, A. I., Peter, M., Loewith, R., von Mering, C., and Aebersold, R. (2010) Phosphoproteomic analysis reveals interconnected system-wide responses to perturbations of kinases and phosphatases in yeast. Sci Signal 3, rs4.

9. Zhou, H., Ye, M., Dong, J., Corradini, E., Cristobal, A., Heck, A. J., Zou, H., and Mohammed, S. (2013) Robust phosphoproteome enrichment using monodisperse microsphere-based immobilized titanium (IV) ion affinity chromatography. Nat Protoc 8, 461-480.

10. Ting, L., Rad, R., Gygi, S. P., and Haas, W. (2011) MS3 eliminates ratio distortion in isobaric multiplexed quantitative proteomics. Nature methods 8, 937-940.

11. Altelaar, A. F., Frese, C. K., Preisinger, C., Hennrich, M. L., Schram, A. W., Timmers, H. T., Heck, A. J., and Mohammed, S. (2012) Benchmarking stable isotope labeling based quantitative proteomics. J Proteomics in press.

12. Everley, R. A., Kunz, R. C., McAllister, F. E., and Gygi, S. P. (2013) Increasing throughput in targeted proteomics assays: 54-plex quantitation in a single mass spectrometry run. Analytical chemistry 85, 5340-5346.

13. Hebert, A. S., Merrill, A. E., Bailey, D. J., Still, A. J., Westphall, M. S., Strieter, E. R., Pagliarini, D. J., and Coon, J. J. (2013) Neutron-encoded mass signatures for multiplexed proteome quantification. Nature methods 10, 332-334.

14. Kalinski, P. (2012) Regulation of immune responses by prostaglandin E2. J Immunol 188, 21-28.

15. Chemnitz, J. M., Driesen, J., Classen, S., Riley, J. L., Debey, S., Beyer, M., Popov, A., Zander, T., and Schultze, J. L. (2006) Prostaglandin E2 impairs CD4+ T cell activation by inhibition of lck: implications in Hodgkin's lymphoma. Cancer Res 66, 1114-1122.

16. Frese, C. K., Altelaar, A. F., Hennrich, M. L., Nolting, D., Zeller, M., Griep-Raming, J., Heck, A. J., and Mohammed, S. (2011) Improved peptide identification by targeted fragmentation using CID, HCD and ETD on an LTQ-Orbitrap Velos. J Proteome Res 10, 2377-2388.

17. Cox, J., and Mann, M. (2008) MaxQuant enables high peptide identification rates, individualized p.p.b.-range mass accuracies and proteome-wide protein quantification. Nat Biotechnol 26, 1367-1372.

18. Rigbolt, K. T., Vanselow, J. T., and Blagoev, B. (2011) GProX, a user-friendly platform for bioinformatics analysis and visualization of quantitative proteomics data. Mol Cell Proteomics 10, O110 007450.

19. Colaert, N., Helsens, K., Martens, L., Vandekerckhove, J., and Gevaert, K. (2009) Improved visualization of protein consensus sequences by iceLogo. Nat Methods 6, 786-787.

20. Oberprieler, N. G., Lemeer, S., Kalland, M. E., Torgersen, K. M., Heck, A. J., and Tasken, K. (2010) High-resolution mapping of prostaglandin E2-dependent signaling networks identifies a constitutively active PKA signaling node in CD8+CD45RO+ T cells. Blood 116, 2253-2265.

21. Swaney, D. L., McAlister, G. C., and Coon, J. J. (2008) Decision tree-driven tandem mass spectrometry for shotgun proteomics. Nature methods 5, 959-964.

22. Colaert, N., Helsens, K., Martens, L., Vandekerckhove, J., and Gevaert, K. (2009) Improved visualization of protein consensus sequences by iceLogo. Nat Methods 6, 786-787.

23. Schwartz, D., and Gygi, S. P. (2005) An iterative statistical approach to the identification of protein phosphorylation motifs from large-scale data sets. Nature biotechnology 23, 1391-1398.

24. Miller, M. L., Jensen, L. J., Diella, F., Jorgensen, C., Tinti, M., Li, L., Hsiung, M., Parker, S. A., Bordeaux, J., Sicheritz-Ponten, T., Olhovsky, M., Pasculescu, A., Alexander, J., Knapp, S., Blom, N., Bork, P., Li, S., Cesareni, G., Pawson, T., Turk, B. E., Yaffe, M. B., Brunak, S., and Linding, R. (2008) Linear motif atlas for phosphorylation-dependent signaling. Sci Signal 1, ra2.

25. Hornbeck, P. V., Kornhauser, J. M., Tkachev, S., Zhang, B., Skrzypek, E., Murray, B., Latham, V., and Sullivan, M. (2012) PhosphoSitePlus: a comprehensive resource for investigating the structure and function of experimentally determined post-translational modifications in man and mouse. Nucleic acids research 40, D261-270.

26. Vizcaino, J. A., Cote, R. G., Csordas, A., Dianes, J. A., Fabregat, A., Foster, J. M., Griss, J., Alpi, E., Birim, M., Contell, J., O'Kelly, G., Schoenegger, A., Ovelleiro, D., Perez-Riverol, Y., Reisinger, F., Rios, D., Wang, R., and Hermjakob, H. (2013) The PRoteomics IDEntifications (PRIDE) database and associated tools: status in 2013. Nucleic Acids Res 41, D1063-1069.

Phosphoproteome dynamics in onset and maintenance of oncogene-induced senescence

Erik L. de Graaf[1,2], Joanna Kaplon[1,2], Albert J. R. Heck[1,2], Daniel S. Peeper and A. F. Maarten Altelaar[1,2]

Manuscript in submission

[1] Biomolecular Mass Spectrometry and Proteomics, Bijvoet Center for Biomolecular Research and Utrecht Institute for Pharmaceutical Sciences, Utrecht University, Padualaan 8, 3584 CH Utrecht, The Netherlands.

[2] Netherlands Proteomics Centre, Padualaan 8, 3584 CH Utrecht, The Netherlands

[3] Center for Biomedical Genetics.

[4] Division of Molecular Oncology, The Netherlands Cancer Institute, Plesmanlaan 121, 1066 CX Amsterdam, the Netherlands.

Abstract

In human skin, the $BRAF^{V600E}$ mutation is known to cause benign lesions, for example melanocytic nevi (moles). These lesions are arrested by a cell-autonomous mechanism called Oncogene-Induced Senescence (OIS). Infrequently, nevi can progress to malignant melanoma, through mechanisms that are incompletely understood. To gain more insight in this vital tumor suppression mechanism, we performed a mass spectrometry-based screening of the proteome and phosphoproteome in cycling and senescent cells as well as cells that have abrogated senescence. Proteome analysis of senescent cells revealed the up regulation of established senescence markers, including specific cytokines, but also several novel proteins including extracellular matrix-interacting and transforming factors. Using both general and targeted phosphopeptide enrichment by Ti^{4+}-IMAC, as well as phosphotyrosine antibody enrichment, we identified over 15,000 phosphorylation sites. Among the regulated phosphorylation sites we encountered components of the interleukins, BRAF and CDK-retinoblastoma (Rb) pathways and other unexpected factors. The combined extensive proteome and phosphoproteome dataset of $BRAF^{V600E}$-expressing senescent cells provides molecular clues as to how OIS is initiated, maintained or evaded.

Introduction

In order to sustain their reproductive lifespan multicellular organisms require several safeguard mechanisms to maintain cell/tissue homeostasis. Growth and cell replication are essential processes; however, uncontrolled growth can be detrimental resulting in cancer and eventually death. Therefore, several tumor suppressive mechanisms have evolved including apoptosis (1) and replicative senescence (2, 3) that can lead to cell self-destruction, nutrient deprivation or cell proliferation arrest, respectively. Recently it has been shown both in vitro (4-6) and in vivo (7-9) that oncogene activation (e.g. through BRAF, Ras, Raf and E2F) can induce an irreversible cell growth inhibition mechanism termed Oncogene-Induced Senescence (OIS).

Human skin cells are naturally exposed to multiple stress factors that can induce gene mutations, which potentially lead to constitutive protein activation and ultimately tumor formation. Benign skin tumors that may remain dormant for decades, are manifested throughout the human skin in the form of melanocytic nevi (moles) and rarely progress into a malignant state. Nevi show markers of senescence including growth arrest, increase in senescence associated (SA)-β-galactosidase activity and induction of tumor suppressor p16 (7). Strikingly, the $BRAF^{V600E}$ activating mutation (10) is found with very high frequency (~50%) in both nevi as well as primary melanomas (11). This suggests that this BRAF mutation alone is insufficient for melanoma development and additional mutations or other post-transcriptional alterations are needed for transformation. This idea is supported by $BRAF^{V600E}$ knock-in mouse models, which develop nevi that infrequently progress to melanomas (12, 13). Transcriptomic analysis has previously shown that the maintenance of $BRAF^{V600E}$ induced senescence is dependent on an inflammatory network governed by the transcription factor CCAAT-enhancer-binding protein β (C/EBP β) (14). Senescence mechanisms defying tumor outgrowth are currently heavily investigated to understand endogenous tumor suppressive pathways involved and to provide alternative drug solutions to cancer treatment.

To better understand at the molecular level the mechanisms underlying the onset and maintenance of OIS in human fibroblasts, we used multiple complementary proteomics techniques to achieve a high coverage of both the proteome and the phosphoproteome (15). Each protocol has been optimized previously (16-18) to maximize the number of proteins and phosphorylation events quantified. Strong cation exchange peptide fractionation together with both global phosphopeptide enrichment as well as phosphotyrosine site specific enrichment techniques were applied to allow for a deep coverage of the senescence phosphoproteome. Using stable isotope dimethyl labeling 5,997 proteins, 12,547 phosphoserine, 2,361 phosphothreonine and 590 phosphotyrosine sites could be quantified.

Experimental Procedures

Cell culture and cell assays - The human diploid fibroblast (HDF) cell line Tig3 expressing the ectopic receptor, hTERT and sh-p16INK4A (Tig3 (et)-16i) was maintained in DMEM

with 4.5 mg/ml glucose and 0.11 mg/ml sodium pyruvate, supplemented with 9% fetal bovine serum (PAA), 2 mM glutamine, 100 units/ml penicillin and 0.1 mg/ml streptomycin (GIBCO). The Phoenix packaging cell line was used for the generation of ecotropic retroviruses. The plasmids pMSCV-blast and pMSCV-blast-BRAF^{V600E} as well as pRS-puro and pRS-puro-C/EBPβ#1 were previously described (14). For infections, filtered (pore size 0.45 mm) viral supernatant, supplemented with 4–8 μg/ml polybrene was used. In general, a single infection round of 6 h was sufficient to infect at least 90% of the population. Cells infected with shRNA-encoding retrovirus were selected pharmacologically (puromycin or blasticidin) and subsequently infected with BRAF^{V600E}-encoding or control virus. After selection, HDF were seeded for cell proliferation assays into a six-well plate or 6 cm plate (2×10^5, 4×10^5 or 6×10^5 cells, resp.) and maintained in the selection medium. Fixation and staining with crystal violet was performed at day 3 and 9 after the last infection. Images of cell proliferation assays reflect representative results of several independent experiments. Senescence-associated (SA) β-galactosidase was stained using the 'Senescence β-Galactosidase Staining Kit' from Cell Signaling at pH 6 according to the manufacturer's protocol. Images reflect representative results of several independent experiments.

Proteomics sample preparation - Frozen cell pellets were lysed by sonication in lysis buffer (8M Urea in 50 mM ammonium bicarbonate, 1 tablet Complete mini EDTA-free Cocktail (Roche) and 1 tablet PhosSTOP phosphatase inhibitor Cocktail (Roche)), samples for phophotyrosine peptide pulldowns contained an additional 1mM sodium orthovanadate. After centrifugation (20 000x g 30min at 4C), the supernatant was assayed for protein content using the BCA kit standard procedure (Pierce). Protein reduction and alkylation were performed using final concentrations of 5 mM dithiothreitol and 10 mM iodoacetamide, respectively. A first enzymatic digestion step was performed in 8 M urea lysis buffer using Lys-C at 37 °C for 4 h (enzyme:substrate 1:75). The second digestion was performed overnight (37 °C) with trypsin (enzyme:substrate 1:100) in 2 M Urea. Resulting peptides were chemically labelled and washed on Sep-Pak C18 columns (Waters, USA, Massachusetts) using stable isotope dimethyl labelling as described before.(19) Cycling HDF cells were labelled with light, OIS cells with medium and OIS bypass cells with heavy dimethyl isotopes. In the replicate experiment the medium and heavy labels were swapped. The labelling efficiency for all labels was higher than 95%. An aliquot of each label was measured on a regular LC-MS/MS run and samples were mixed 1:1:1 (L:I:H) based on their peptide intensities. This was found to result in a more precise ratio than using the total protein amounts as determined by a BCA assay. After mixing, peptides were dried to completion in vacuo.

Phosphotyrosine peptide immunoprecipitation - Peptides were re-suspended in 800 μL of cold immuno-affinity purification buffer consisting of 50 mM Tris-HCl (pH 7.4), 150 mM NaCl, protease inhibitors (Roche Diagnostics, Germany) and 1% n-octyl-β-D-glucopyrano¬side (NOG) (Sigma, Germany). The peptide mixture was agitated on a shaker for 30 minutes to dissolve the peptides thoroughly and the pH was adjusted, if necessary, to pH 7.4. The immuno-affinity purification was performed as previously described (20, 21). In summary, the peptides were added to 50 μL slurry of PY99 antibody beads (Santa Cruz biotechnology; CA USA) and the peptide-antibody-bead mixture was incubated overnight at 4°C on a rotator. The beads were spinned down and supernatant was used for a second

iteration of immunoprecipitation using fresh PY99 beads. The pelleted beads were washed and bound peptides were eluted twice using 0.15% trifluoroacetic acid (TFA) (Sigma, Germany). Sample desalting was performed using home-made tips with C18 material (Aqua™ C18, 5 μm, Phenomenex, Torrance, CA) as described elsewhere (22). Finally, peptides were dried in vacuo and reconstituted in 40 μL of 10% formic acid prior to LC-MS/MS analysis.

Strong cation exchange fractionation - SCX was performed on two separate systems optimized for sample amount and type. For the protein identification experiment, 150-200 μg sample was injected on SCX system 1 and for the Ti^{4+}-IMAC phosphopeptide identification experiment, 3 mg was loaded on SCX system 2. For SCX system 1, peptides were fractionated as described elsewhere (Helbig et al., 2010). Briefly, the SCX system consisted of an Agilent 1100 HPLC system (Agilent Technologies) with a Strata X 33u (Phenomenex, The Netherlands; 50 x 4.6 mm) trapping cartridge and a polysulfoethyl A SCX column (PolyLC, Columbia, MD; 200 mm x 2.1 mm inner diameter, 5 μm, 200-Å). Labeled peptides were reconstituted in 10% formic acid and loaded onto the trap column at 100 μl/min and subsequently eluted onto the SCX column with 80% acetonitrile (ACN) (Biosolve, The Netherlands) and 0.05% FA. SCX buffer A was made out of 5 mM KH_2PO_4 (Merck, Germany), 30% ACN and 0.05% FA, pH 2.7; SCX buffer B consisted of 350 mM KCl (Merck, Germany), 5 mM KH_2PO_4, 30% ACN and 0.05% FA, pH 2.7. The gradient was as follows: 0% B for 10 min, 0–85% B in 35 min, 85–100% B in 6 min and 100% B for 4 min. A total of 45 fractions were collected for each set and dried in a vacuum centrifuge. The second SCX system (Pinkse et al., 2008) was performed using a Opti-Lynx (Optimized Technologies, Oregon OR) trapping cartridge and a Zorbax BioSCX-Series II column (0.8-mm inner diameter × 50-mm length, 3.5 μm). SCX Solvent A consists of 0.05% formic acid in 20% acetonitrile while solvent B was 0.05% formic acid, 0.5 M NaCl in 20% acetonitrile. The SCX salt gradient as follows: 0-0.01 min (0-2% B); 0.01-8.01 min (2-3% B); 8.01-14.01 min (3-8% B); 14.01-28 min (8-20% B); 28-38 min (20-40% B); 38-48 min (40-90% B); 48-54 min (90% B); 54-60 min (0% B). A total of 50 SCX fractions (1 min each, i.e. 50-μl elution volume) were collected and dried in a vacuum centrifuge.

Ti^{4+}-IMAC phosphopeptide enrichment - Prior to phosphopeptide enrichment, SCX fractions were desalted using Sep-Pak C18 columns and dried to completion using a speed vac. Ti^{4+}-IMAC columns were prepared and used as described previously (18, 23). Briefly, microcolumns were created by loading Ti^{4+}-IMAC beads onto GELoader tips (Eppendorf) with a C8 plug to approximately 1–2 cm length. The enrichment procedure for all SCX fractions was as follows: Ti^{4+}-IMAC columns were pre-equilibrated two times with 30 μL of loading buffer (80% ACN, 6% trifluoroacetic acid (TFA)). Next, each SCX fraction was resuspended in 30 μL of loading buffer and loaded onto the equilibrated GELoader tips. Ti^{4+}-IMAC columns were washed with 40 μL washing buffer A (50% ACN, 0.5% TFA, 200 mM NaCl) and subsequently with 40 μL washing buffer B (50% ACN, 0.1% TFA). Bound peptides were eluted by 30 μL of 10% ammonia into 30 μL of 10% formic acid. Finally, the remainder of the peptides was eluted with 4 μL of (80% ACN, 2% formic acid). The collected eluate was further acidified by adding 6 μL of 100% formic acid and subsequently stored at -20 °C for LC-MS/MS analysis.

Liquid Chromatography and Mass Spectrometry - For protein identification and quantification the SCX fractions containing doubly and triply charged peptides (approx. 20 fractions each SCX) were reconstituted in 10% formic acid and analysed directly using nano flow reverse phase liquid chromatography using an Agilent 1200 coupled to a LTQ-Orbitrap Velos mass spectrometer (Thermo, San Jose, CA). Depending on the SCX UV-trace, 1-10% of each fraction was injected. Densely populated 2+ fractions were injected twice to minimize undersampling of the mass spectrometer. About half of each phosphopeptide sample was injected to allow for erroneous events. Peptides were trapped on a trap column (ReproSil-Pur C18-AQ, 3μm, (Dr. Maisch GmbH, Ammerbuch, Germany); 20 mm x 100 μm ID, packed in house) at 5μl/min in 100% solvent A (0.1 M acetic acid in water). Next, peptides were eluted from the trap column onto the analytical column (ReproSil-Pur C18-AQ, 3μm (Dr. Maisch GmbH, Ammerbuch, Germany); 40 cm x 50-μm ID, packed in house) at ˜100 nl/min in 1h, 2h or 3 h linear gradients from 10 to 50% solvent B (0.1 M acetic acid in 8:2 (v/v) acetonitrile:water). Nanospray was achieved with an in-house pulled and gold-coated fused silica capillary (360 μm OD; 20 μm ID; 10 μm tip ID) and an applied voltage of 1.7 kV. In all experiments the mass spectrometer was configured to perform a FT survey scan from 350 to 1500 m/z at a resolution of 30 000. For the protein identification/quantification experiment the top 10 most intense peaks were fragmented by HCD fragmentation (35% normalized CE at a target value of 50 000 ions, resolution 7500). The phosphopeptide samples were analysed using the HCD/ETcaD decisiontree as described previously (24), fragmenting the top 10 or top 5 most abundant peaks for the Ti^{4+}-IMAC and phosphotyrosine IP samples respectively.

Data analysis - All MS data were processed with Proteome Discover (Thermo Scientific) using a standardized workflow. Peak lists, generated in Proteome Discover, were searched against a concatenated forward-decoy(reverse) Uniprot (v2010-12, taxonomy Homo Sapiens, 41 008 protein entries) database, supplemented with frequently observed contaminants, using Mascot (Matrix Science, UK). The following search parameters were used: 50 ppm precursor mass tolerance, 0.05 Da fragment ion tolerance for FT analysed spectra, 0.6 Da fragment ion tolerance for IT analysed spectra, trypsin cleavage with maximum of 2 miscleavages, cysteine carbamidomethyl static modification and methionine oxidation and dimethyl labelling (L,M,H) of lysine residues and the peptide N-termini as dynamic modification. For the phosphopeptide experiments, Phospho (STY) was used as an additional dynamic modification and the phosphoRS (25) node was used in the workflow to calculate site occupation probabilities. Triplex dimethyl labelling was used as a quantification method, with a mass precision of 2 ppm for consecutive precursor mass scans and a normalisation on the median peptide ratio. A RT tolerance of 0.5 min was used to account for the effect of deuterium on the retention time in long gradients. To filter for high quality protein data and to control the false discovery rate on identifications, only the PSMs adhering to the following criteria were kept for analysis: minimal Mascot score of 20, minimal peptide length of 7, no inconsistently labelled peptides (i.e. no light n-terminus and heavy lysine on a single peptide) only unique rank 1 peptides and a mass deviance between 20 ppm and 15 ppm. As a result we obtained peptide FDRs of 1% and lower for all experiments. To distinguish between proteins with high similarities, only unique peptides were considered for protein identification and quantification. Proteins with a minimum average 2-fold change

were considered regulated. The criteria used for phosphopeptide analysis were a 1 % FDR calculated by percolator (26) and an additional minimum mascot score of 20. Phosphopeptide quantification was performed by using an in-house developed script that calculated the average ratio, intensity and localisation probability for every non-redundant phosphopeptide. The script mapped all the identified phosphopeptides on their protein keeping doubly/triply phosphorylated peptides together to potentially map differential crosstalk of close-by sites.

Protein extraction and immunoblot analysis - Cycling, OIS or OISb cells were harvested 9 days after $BRAF^{V600E}$ infection. Cells were scraped with 1X PBS, centrifuged @ 4000 rpm for 4 minutes at 4°C and the pellets were frozen or used immediately. Fresh or frozen pellets were lysed on ice in RIPA buffer supplemented with protease inhibitor cocktail (Roche) and phosphatase inhibitors (10 mM β-glycerophosphate, 2 mM sodium fluoride, 0.2 mM sodium orthovanadate, 1 mM sodium pyrophosphate). Lysates were sonicated for 1 minute (5 sec on/off interval), centrifuged at 4°C and 1200 rpm for 10 minutes and supernatants were transferred to fresh Eppendorf tubes. Protein concentrations were determined using Bradford assay (Bio-Rad). Protein samples were prepared in 4X sample buffer (Invitrogen) supplemented with 2.5% β-mercaptoethanol. Proteins were separated on 4-12% polyacrylamide gels (Invitrogen), transferred onto a nitrocellulose membrane (Whatman), and blocked in blocking buffer (4% milk in 1X TBS-Tween) for 1h at room temperature. The membrane was probed with the indicated primary antibodies (overnight at 4°C in 4% milk in 1X TBS-Tween), followed by 1 h incubation with the corresponding secondary antibodies conjugated with horseradish peroxidase (HRP) enzyme. For detection of the signal the membrane was incubated 1 minute with ECL reagent (Amersham Biosciences) and visualized on films (GE Healthcare).

Antibodies - The primary antibodies used for immunoblot analysis were C2CD2 (D01P; H00025966; MaxPab), phospho-cdc2 (known as CDK1) (Tyr15) (#9111; Cell Signaling), cdc2 (known as CDK1) (#9112; Cell Signaling), FGF-2 (C-18; sc-1360; Santa Cruz), Hsp90 (4874; Cell Signaling), STAT-3 phospho (Tyr705) (3E2; #9138; Cell Signaling), STAT-3 (sc-482; Santa Cruz). Secondary antibodies used were goat anti-mouse IgG (H+L) HRP conjugated (G21040; Invitrogen), goat anti-rabbit IgG (H+L) HRP conjugated (G21234; Invitrogen) and rabbit anti-goat IgG (H+L) HRP conjugated (R-21459; Invitrogen).

Results and Discussion

To study the mechanisms underlying Oncogene-Induced Senescence (OIS) we employed a previously described cell system (14) comprising in vitro cultured human diploid fibroblasts (HDF) in three differential states: cycling, OIS and OIS bypass (OISb). HDF transduced with lentivirus harboring an empty vector were used as normal cycling control cells. OIS was induced by infecting HDF with a lentivirus carrying the constitutively active $BRAF^{V600E}$ gene. Lastly, OIS was abrogated by depleting C/EBP β, a crucial component of the inflammatory pathway in OIS, in the presence of mutated BRAF. Cells expressing the $BRAF^{V600E}$ oncogene typically showed an initial burst of growth followed by induction of senescence.

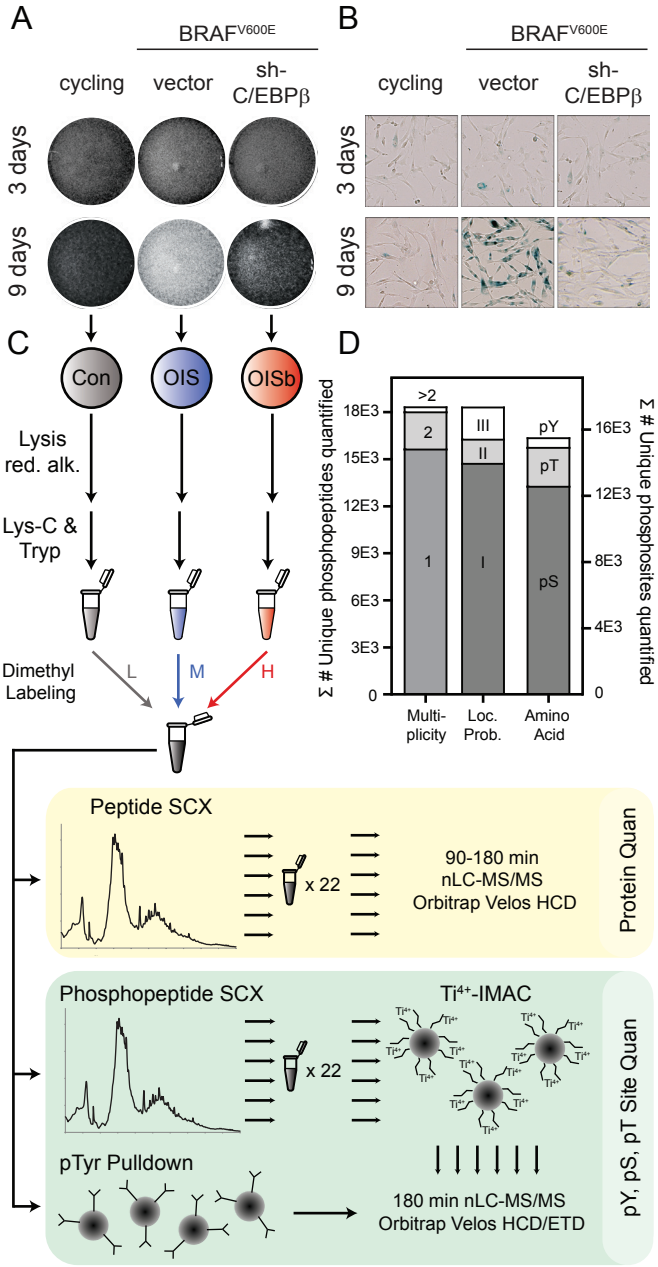

Figure 1: *Cell proliferation assay (A) and SA-β-Galactosidase activity (senescence marker) assay (B) of HDF cells transduced with empty vector, BRAF^{V600E} or both BRAF^{V600E} and a shRNA targeting C/EBPβ that were subsequently fixed and stained 3 and 9 days after induction of BRAF^{V600E}. Senescence is evident only 9 days after BRAF^{V600E} induction (OIS), but not in the cells knocked-down for C/EBPβ (OISb) C) Proteomics workflow optimized for the quantification of proteins and phosphopeptides. D) Number of unique phosphopeptides quantified. Distributions of phosphopeptide multiplicity, site localisation probability Class (I:≥95%, II:≥ 50% and III:<50%) and type of residue phosphorylated.*

Three days after infection, all cells showed equal growth (Fig. 1A). However, after nine days, cells expressing BRAFV600E displayed a strong growth arrest accompanied by the induction of the senescence marker SA-β-galactosidase, while the cycling control cells continued growth without displaying the senescence marker (Fig. 1B; Cycling and $BRAF^{V600E}$/Vector). As expected, the C/EBP β-depleted OIS bypass (OISb) cells, in contrast to OIS cells but similarly to cycling cells, failed to undergo cell cycle arrest or induction of senescence markers nine days following $BRAF^{V600E}$ introduction (Fig. 1A and B; $BRAF^{V600E}$/sh-C/EBPβ). Therefore, we hypothesized that the comparison of cycling, OIS and OISb samples in the early onset and final phenotype would allow us to find proteins and phosphosites regulated specifically in the onset and maintenance of senescence. Moreover, the comparison between OISb and OIS cells could potentially reveal processes involved in the bypass of senescence into tumor progression/malignancy.

To uncover protein expression specific for cycling, OIS or OISb, related lysates were digested and resulting peptides were desalted and labeled with different dimethyl isotopes. After mixing the three labeled cell lysates, peptides were processed through three distinct pipelines. For full proteome analysis, peptides were fractionated by strong cation exchange (SCX) and for unbiased global phosphopeptide analysis an additional SCX was ran followed by Ti^{4+}-IMAC enrichment as described before (18). For the specific analysis of tyrosine phosphorylation dynamics the protein digest was directly treated with phosphotyrosine antibody coupled beads (Fig. 1C). The workflow was applied to cells harvested 3 days and 9 days post transduction and the replicate experiment was performed using a label-swap.

The first experiment was focused on the analysis of unmodified peptides to assess protein identity and protein expression differences and resulted, following stringent filtering as described in the methods section, in the robust quantification of 5,997 proteins. The extensive fractionation before the highly selective Ti^{4+}-IMAC phosphopeptide enrichment together with the pTyr immunoprecipitation resulted in the quantification of 18,320 unique phosphopeptides (15,498 unique phosphosites) with a localization certainty of on average 90% (Fig. 1D). Due to the different type of data distributions for different sample comparisons, we used a threshold of 2-fold up- or downregulated in both replicates to pinpoint differentially regulated proteins/phosphopeptides. The full lists of all quantified proteins and phosphopeptides can be found in Supplementary Tables 1 and 2. The determined ratios in our parallel triple experiments were used to define two sets of biologically distinct changes. The first set of regulated events is so called BRAF specific and is observed when comparing both OIS cells and OISb cells to cycling cells. The second and most interesting scenario was termed OIS specific and consisted of features up or down regulated for the OIS versus cycling and OIS versus OISb comparisons. In order to find consistent regulations the logarithmic base 2 ratio was plotted for both replicates in each sample comparison in figure 2.

Proteome changes upon BRAF transfection

As displayed in the OIS versus OISb protein ratio plots, minor changes were observed after 3 days that became more apparent after 9 days when the full phenotypes were displayed

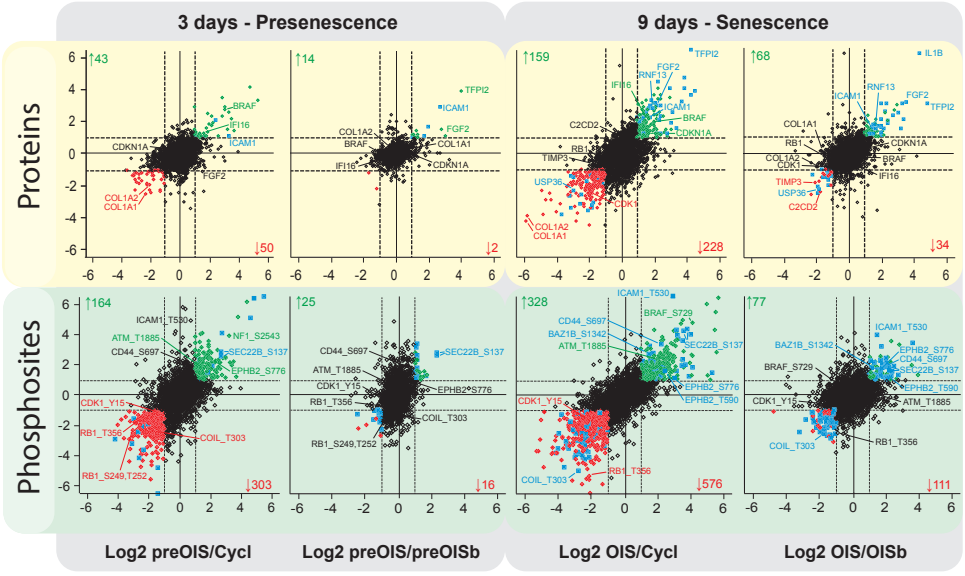

Figure 2: *Proteins and phosphosites comparisons between different conditions. A) Protein ratios determined in both replicates highlighting consistently regulated proteins regulated (replicate one plotted versus replicate two). In total 5997 proteins were quantified. B) Ratio plots of unique phosphopeptides quantified in both Ti^{4+}-IMAC phosphopeptide enrichments and phosphotyrosine immunoprecipitation replicates. Down or up regulated proteins/phosphopeptides (>2-fold) are highlighted with red and green diamonds, respectively. Proteins and phosphosites regulated specifically in OIS (and not in OISb) are highlighted with blue squares.*

(16 and 102 substantially regulated protein after 3d and 9d, respectively) (Fig. 2, Suppl. Fig. 1). Interestingly, a larger number of differentially regulated proteins was observed for the OIS versus cycling and OISb versus cycling cells, that increased over time (93 and 387 proteins regulated for 3d and 9d, respectively)(Fig. 2, Suppl. Fig. 1). Although the OIS cells are cell cycle arrested after 9 days and both the OISb and cycling cells are still growing, a substantially higher similarity in protein expression was observed between OISb cells and OIS cells than between OISb and cycling cells. This illustrates that oncogene activation has clearly a dominant effect on global protein expression levels. In line with expectations all BRAF transduced cell lines display similarly elevated levels of BRAF both after 3 days and 9 days of growth, suggesting equal transduction levels (Fig. 2, Suppl. Fig. 1).

Ontology enrichment analysis revealed that proteins concurrently down regulated in all BRAF transduced cells (OIS and OISb, 3 days and 9 days), are primarily involved in extracellular matrix interactions (Fig. 3A). Extracellular matrix constituents such as collagens (I,III,VI,XII), fibronectin and cytoskeletal proteins filamin A and C were amongst the most strongly down regulated proteins in both OIS and OISb (Fig. 2, Suppl. Fig. 1). Interestingly, there is a difference between the early and the late stage of oncogenic insult for some down regulated protein categories. DNA replication and cell cycle regulatory proteins were differentially enriched in the preOIS stage at day 3, whereas cell contact gene ontologies were found to be specifically enriched in OIS. The only significantly enriched protein group that was up regulated in both OIS and OISb belonged to the lysosome compartment. When

looking more specifically into the data, it is interesting to note that IFI16 was up regulated in all BRAF expressing cells after three and nine days (Fig. 2, Suppl Fig. 1). It has been shown previously that IFI16 is able to induce senescence via up regulation of p21(CDKN1A) and that subsequent p21 knockdown can bypass senescence (27). In our model we see p21 up regulated in both OIS and OISb but only after nine days, suggesting a delayed tumor suppressive response to oncogene activation. Moreover, in line with the up regulation p21, CDK1 protein expression levels showed a down regulation after 9 days in OIS that seems to be less strong in OISb (Fig. 2, Suppl Fig. 1, Fig. 4B).

Phosphoproteome changes upon BRAF transfection

Similar to the proteome changes, the phosphoproteome changes were found to be the largest when comparing either the OIS or OISb cells to cycling cells (Fig. 2, Suppl. Fig. 2). In line with an increase of BRAF protein expression we also observed a strong increase in BRAF Ser729 phosphorylation in OIS and OISb cells. At the same time, phosphorylation on NF1 (S2543, T2565, S2802) and sprouty (related) proteins (Spry2_S167, Spry4_S125, Spred1_S238 and Spred2_S168) that are known to negatively affect BRAF activation, were similarly up regulated after BRAF oncogene introduction. This is in line with a previous study by Courtois-Cox et al. demonstrating a negative feedback signalling loop involved in OIS (28). In addition, up regulation of ATM (T1885) phosphorylation was observed, while ATM protein levels remained unchanged. Although biological information about this specific site is absent, this site could be potentially interesting as the ATM stress kinase is known to phosphorylate important components of the stress response, such as p53 (29).

As described above the increase in protein levels of the CDK inhibitor p21 was observed both in OIS and OISb only after 9 days. Therefore, we searched for the effects of p21 upregulation on the phosphorylation levels of CDK targets. Indeed our global phosphorylation screen revealed a strong reduction in CDK dependent phosphorylation sites as observed in the sequence motif analysis (Fig. 5A). When compared to background, the CDK phosphorylation motif (SPxK) was strongly and solely enriched in the set of phosphosites down regulated in both BRAF transduced cells after 9 days (56-fold increase). When CDKs bind to Cyclins their t-loop is repositioned exposing the active site of CDK. Tyrosine 15 is part of the active site and gets phosphorylated by Wee1 and Myt1, and is essential to control the activity of the CDK-cyclin complexes (30, 31). Interestingly, CDK1/2/3 Y15 phosphorylation was strongly down regulated in OIS and to a lesser extent in OISb (Fig. 2B, Fig. 4B). Hence the down regulation of CDK phosphorylation in OIS and OISb correlates well with the elevated p21 levels, inhibiting the Cyclin-CDK complex formation.

When looking at individual CDK targets, RB1 is one of the major Cyclin/CDK substrate proteins controlling cell cycle progression; RB1 hyper-phosphorylation is associated with transcriptional repression. In line with this, a strong down regulation was observed of RB1 phosphorylation sites in senescence, while the proteome screening indicated unchanged RB1 protein levels. After 3 days of growth, phosphorylated T356 as well as the dual phosphorylated sites S249, T252 and T821, T826 were only down regulated in the preOIS cells

(Fig. 2, Suppl. Fig. 2). After 9 days, phosphorylation levels of both dual phosphorylation sites were near absent in both OIS and OISb. Furthermore, T356 phosphorylation was found to be down regulated in OIS and into a lesser extent in OISb (Fig. 2B, Suppl. Fig. 2). These changes in phosphorylation indicate a strong RB1 mediated growth suppression upon BRAF transfection, albeit more pronounced in the (pre-) senescent cells.

Another CDK1 substrates is histone acetyltransferase Myst2 T88 that was indeed strongly down regulated in the OIS and OIS bypass cells, whereas protein levels remained unregulated. Interestingly, Wu et al. has shown that phosphorylation of this site by CDK1 can be crucial for cell cycle progression (32). Similarly, another CDK1 substrate, S38 on Stathmin1 (Stmn1) (33), showed a down regulation of phosphorylation in both OIS and OISb.

In summary, these findings indicate a strong tumor suppressive response to BRAF oncogene activation, involving CDK inhibition, potentially leading to cell cycle arrest. However,

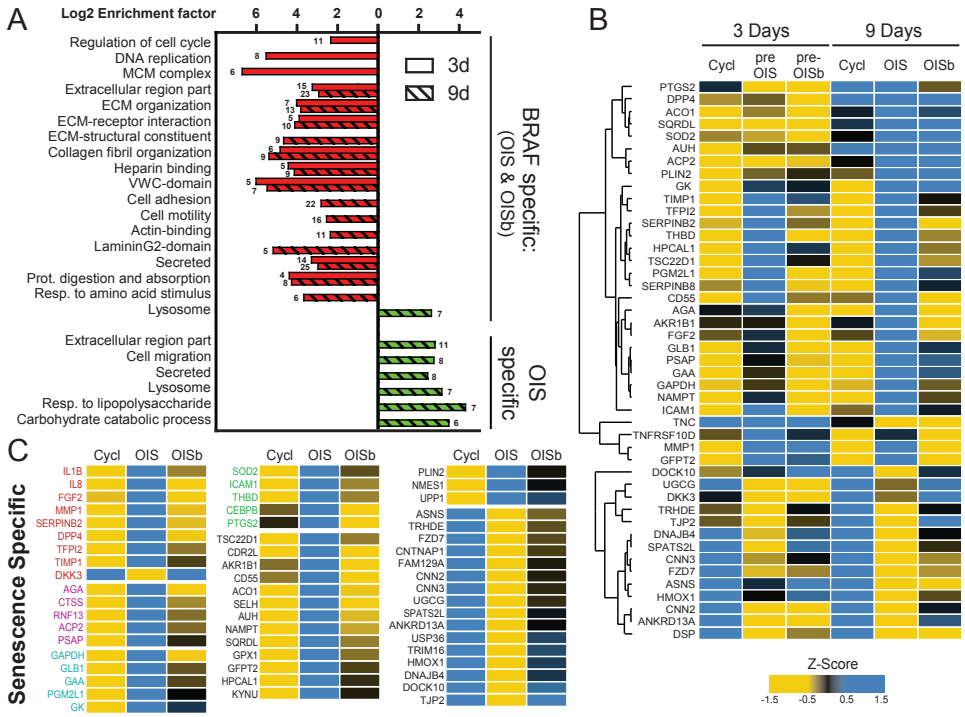

Figure 3: *Subset of proteins potentially involved in oncogene-induced senescence as they are specifically regulated in OIS cells (and not in OISb), i.e. a senescence signature proteome. A) Functional classification of the proteins in the oncogene transduced and senescence signature proteome by GO categories (Benjamin-Hochberg corrected FDR<0.02). The numbers indicate the amount of proteins in each category and red and green bars indicate down and up regulated categories, respectively. B) Hierarchically clustered normalized average intensities displayed as a heat map for proteins specifically regulated in preOIS and/or OIS (3 and/or 9 days post-transfection of BRAF). Only proteins quantified in all replicates of all samples were plotted. C) Normalized average intensities displayed as a heat map of proteins specifically regulated 9 days post BRAF transfection. Functional categories were colored as followed: secreted proteins (red), lysosomal (purple), LPS response (green), carbohydrate processing (cyan). Proteins present in multiple categories were colored only according to their first category.*

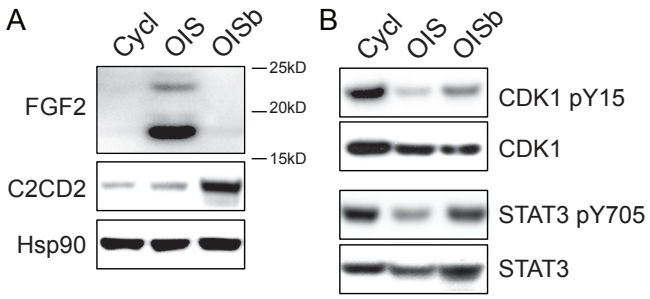

Figure 4: *Western blot validation of a selection of regulated proteins and phosphosites. A) Protein blots probed with antibodies for FGF2 and C2CD2 confirm their regulation observed in the proteomics data. Hsp90 was used as a loading control. B) CDK1 and STAT3 phosphorylation and total protein levels confirm a strong reduction of phosphorylation in OIS and a smaller reduction in the OIS bypass cells.*

although all BRAF activated cells display a theoretically higher sensitivity for cell cycle arrest, only the cells with endogenous CEBP/β levels intact have the ability to enter a full cell cycle arrest. These senescence specific events were studied next.

Proteome changes specific for Senescence

Even though the proteomes of (pre-) OIS and OISb are very much alike at a global scale, the former cells are cell cycle arrested while OISb keep proliferating. This implies there are additional mechanisms or molecular markers that distinguish the OIS cells from OISb cells. Therefore, we narrowed down our focus to the small set of proteins that also displayed differential expression between OIS and OISb. After 3 days the OIS cells are still in an early or pre-senescence state, while after 9 days they are senescent. Therefore, of interest are proteins that are regulated both in the early and full OIS stage after 3 and 9 days, respectively. Moreover, other proteins relevant to OIS are those only regulated after 9 days when the full phenotype is displayed. Therefore, we can distinguish between markers of senescence onset and maintenance.

In figure 3 an overview is given of the senescence specific proteins, whereby in 3A they are clustered into gene ontology groups. Upregulated enriched gene ontologies include mainly lysosomal, inflammatory and extracellular acting proteins (Fig. 3A). To visualize the dynamics of senescence associated proteins the intensities for proteins regulated in OIS cells after 3 days and/or 9 days were clustered in figure 3B. Interestingly, some extracellular matrix processing proteins such as SerpinB2/PAI-2 and TFPI2 are regulated both after 3 and 9 days, whereas other proteins belonging to the same category were specifically upregulated after 3 days (SerpinB8) or 9 days only (MMP1, TIMP1). This could indicate different phases of extracellular modifications are required in order to progress into a senescence like state.

Consistent with our previous results on mRNA profiling (14), a strong up regulation was found for the cytokines IL-1b and IL-8 in senescent cell lines at day 9. At the preOIS stage, IL-1b was already clearly visible in both BRAF transduced cells whereas IL-8 was only identified in one replicate. Additional proteins specifically regulated in the final senescent

phenotype (Fig. 3C) include CEBP/B and TSC22D1 that have been shown previously to stimulate the production of cytokines such as IL-1b and IL-8, and to be crucial to maintain the senescence phenotype in OIS (14, 34, 35). Similarly, PTGS2 (COX2) the main enzyme in the inflammation-linked prostaglandin production was found to be up regulated in OIS in our data and was shown previously to be critical for the senescence onset and maintenance (36). Other up regulated proteins include glucose linked biomolecule-processing enzymes (i.e. AGA, GFPT2) and lysosomal (i.e. GAA, GLB1/beta-galactosidase) proteins.

Another factor strongly up regulated specifically in OIS was the fibroblast growth factor FGF2. The finding of this growth promoting factor was at first sight counterintuitive but western blotting validated a strong presence of both the short excretable para-/autocrine FGF2 and the longer NLS containing intracrine FGF2 variant of 18kDa and 24kDa, respectively (Fig. 4B). In accordance with our data, FGF2 was shown previously in oncogenic RAS cells to induce an irreversible cell cycle arrest that could be negated by a tyrosine kinase inhibitor (37).

One of the most strongly down regulated proteins in senescence was the deubiquitinating (DUB) enzyme USP36 (Fig. 2, Fig 3C). Interestingly, this enzyme was previously shown by Richardson et al. to positively control cell growth by stabilizing rRNA production and ribosome biogenesis (38). Moreover, knockdown of both yeast and fly USP36 homologs were shown to result in an inhibition of cell growth (38, 39). Therefore, in the light of our model system, USP36 can provide a novel link between DUBs and growth arrest in OIS. Conversely, an ubiquitin E3 ligase (RNF13) was found to be up regulated in OIS, suggesting an important role for ubiquitin mediated proteasomal degradation in OIS.

In summary, the protein expression analysis revealed a lot of different processes regulated in OIS previously shown to be directly or indirectly linked to growth arrest but also hinting at mechanisms yet to be further explored in the context of senescence.

Phosphoproteome changes specific for Senescence

To specifically get insight into the signalling involved in OIS we focus on the OIS cell specific regulations in phosphopeptides (i.e. differential regulation between OIS and OISb or Cycl). Among the most interesting phosphorylation events specifically regulated in OIS are phosphosites on COIL, STAT3, p53, BAZ1B, EPHB2 and CD44. For Coilin/p80 (COIL) total protein levels remained unchanged, while its phosphorylation at Thr303 was down regulated already in preOIS after 3 days that was even more strongly and specifically pronounced in OIS after 9 days. Previously, Coilin phosphorylation levels were shown to increase in mitosis (40), suggesting a role of Coilin phosphorylation in the growth inhibition in OIS.

Signal transducer and transcription activator 3 (STAT-3) is a well-known mediator of cellular responses activated by several growth factors and cytokines. Tyrosine phosphorylated STAT3 dimerizes to form an activated signal transducer that has was shown to be an important mediator in a whole variety of human cancers (41, 42). Moreover, it is interesting to note that inhibition of STAT3 activity also lead to the induction of pro-inflammatory cytokine expression, including IL-6 (43). In our study, we indeed observed down regulation of STAT-3

Tyr705 phosphorylation specifically in OIS while protein levels remained unchanged (Fig. 4). This suggests that STAT3 may be an important mediator of OIS by reducing growth and by inducing the inflammatory response needed to maintain the OIS signature.

Surprisingly, the phosphorylation of the tumor suppressor protein p53 at Ser315 was specifically down regulated in OIS cells (9d) only. This site was demonstrated to be phosphorylated by two different kinases. Blaydes et al. demonstrated that CDK inhibition reduced Ser315 phosphorylation levels and that the p53 S315A mutant showed reduced transcriptional activity, resulting in lowered expression levels of p21 in cancer cell lines (44). However, in our data similar levels of p21 were observed in OIS and OISb and in contrast to the observation in OIS, Ser315 phosphorylation was not down regulated in OISb, which suggests that in our model p21 levels are independent of p53 Ser315 phosphorylation. The second kinase reported to phosphorylate p53 Ser315 is Aurora kinase A. In a study by Katayama et al., phosphorylation of Ser315 by Aurora A lead to p53 ubiquitination and subsequent protein degradation (45). Therefore, reduced levels of S315 would suggest a greater stability of p53 in OIS cells compared to cycling and OIS bypass cells potentially leading to a prolonged p53 effect in OIS.

Another potentially interesting phosphosite found up regulated in OIS was Ser1342 of the tyrosine protein kinase WSTF (or BAZ1B) whereas its total protein levels remain similar in all conditions (Fig. 2, Suppl. Fig 3). WSTF is part of the chromatin remodeling complexes WICH (46) and WINAC (47) and was shown to tyrosine phosphorylate histone H2AX thereby regulating the formation of γ-H2AX foci (48). The appearance of γ-H2AX foci is one of the hallmarks of senescent cells and therefore specific up regulation of phosphorylated WSTF could provide a novel link between foci formation and OIS.

Other heavily phosphorylated proteins with sites specifically up regulated in OIS include the transmembrane proteins Extracellular matrix receptor III (ECMR-III/CD44) and Ephrin type-B receptor 2 (EPHB2) that mediate the communication between the ECM and adjacent cells, respectively. (Fig. 2, Suppl. Fig 3). CD44 contained several sites that remained unregulated (Ser686, Ser706 and Thr720), another site that was found down regulated in both OIS and OISb (Ser718) whereas yet another site (Ser697) was specifically up regulated in OIS. EPHB2 showed enhanced phosphorylation in OIS (T590 and S776) while T585 was down regulated in both BRAF transfected cells. Not much is known about the effect of the specific phosphorylation sites on these receptors. However, CD44 is known to bind to the glycosaminoglycan and ECM component hyaluronic acid (HA), to interact with matrix metalloproteinases (MMPs) (49) that we found to be regulated in OIS as well. Moreover, CD44 was also identified as an important cancer stem cell marker in a variety of different cancers (50, 51). Therefore, the role of CD44 in cell-ECM interaction and its' differential phosphorylation status in senescence and cancer might be interesting functionally validate.

Discussion

In this work we report the most comprehensive dataset of protein and phosphosite regulations associated with the onset and maintenance of oncogene-induced senescence to

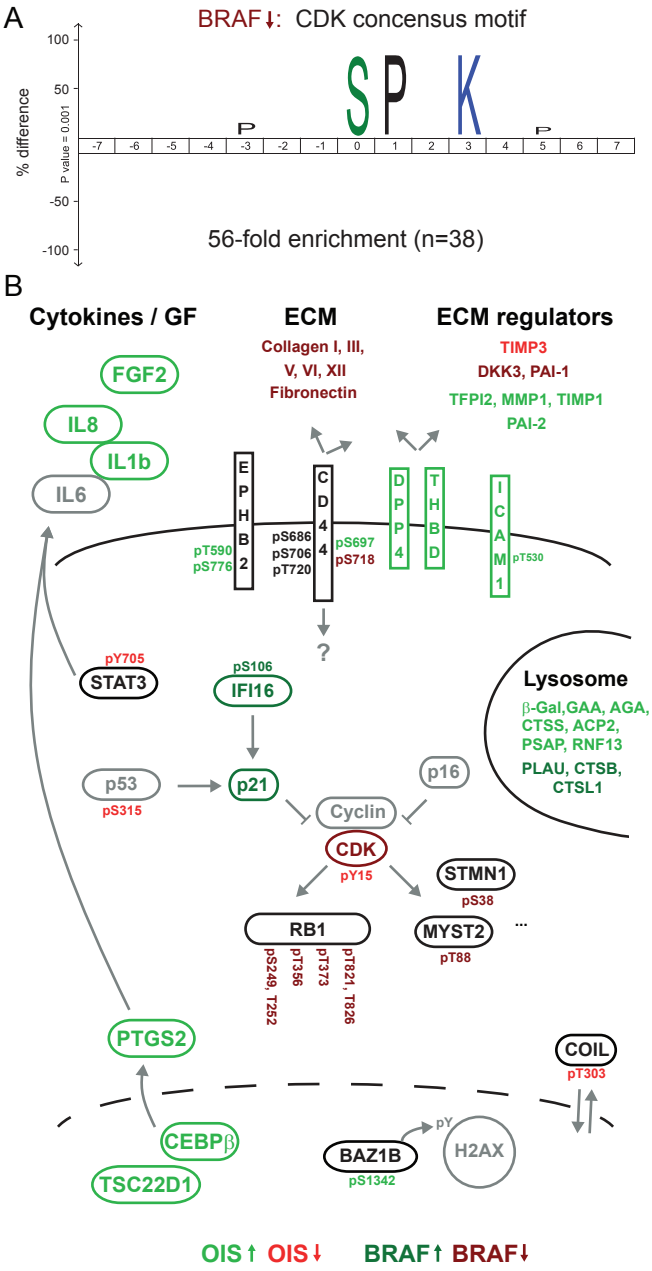

Figure 5: *Phosphosites regulated in Senescence. A) Sequence motif analysis of enriched phosphopeptides down regulated in OIS and OIS bypass by Motif-X, suggesting a specific inhibition of CDK kinase activity. B) (Phospho-)proteins differentially regulated in OIS. Up or down regulated features specific for OIS are depicted in light green and light red, respectively. Proteins and phosphosites regulated by BRAF transformation (regulated in OIS and OIS bypass) are depicted in dark green and dark red, respectively. Unregulated and unquantified proteins are indicated in black and gray, respectively. Gray lines correspond to interactions described in literature.*

date. Using Ti^{4+}-IMAC phosphopeptide enrichment and phosphotyrosine peptide IPs a vast amount of signalling events could be monitored. The high specificity and selectivity of Ti^{4+}-IMAC beads and pTyr antibodies, allowed for a large dynamic range in phosphopeptide detection, including low abundant tumor suppressors (i.e. RB1, p53) and signalling proteins (i.e. STAT3, CDK1).

We induced senescence by ectopic expression of the $BRAF^{V600E}$ oncogene in human primary fibroblasts. The oncogene introduction lead to a substantial reprogramming of the proteome and phosphoproteome (Figure 2) when compared to cycling fibroblasts, both 3 and 9 days following transduction. That not all these observed changes can be attributed to the onset of senescence becomes apparent from our second control, wherein we introduced the $BRAF^{V600E}$ oncogene, but rescued the cells from senescence by shRNA mediated knock-down of the proinflammatory transcription factor C/EBP β essential for $BRAF^{V600E}$ oncogene-induced senescence. The (phospho)proteome data reveal (Figure 2) that substantial less reprogramming of the proteome is taking place when comparing the senescent (OIS) to the bypassed-non-senescent cells (OISb), both after 3 and 9 days. Thus oncogene transduction leads to a large reprogramming of the proteome, while senescence leads to smaller, albeit more specific changes in the (phospho)proteome. That senescence is really salvaged in the OISb cells is evident by some landmarks proteins observed to be differential between OIS and OISb, i.e. the relatively lower levels of interleukins (IL6, IL8) and beta galactosidase, giving credibility to our hypothesis that proteins and phoshosites differentially regulated between OIS and OISb cells are potentially involved in Oncogene-Induced Senescence.

It has been shown previously that senescence can be induced by inhibition of the cell cycle via the p53-p21 and p16INK4A-RB1 pathway (52-54). We also find that BRAFV600E oncogene transduction results in up regulation of p21 and repression of CDK signalling and subsequent RB1 hypophosphorylation. However, this mechanism is not specific for OIS, as p21 levels are enhanced and CDK protein and phosphorylation levels are repressed in both OIS and OISb cells (albeit in OISb to lower extend). This suggests that these tumor suppressive signals were not strong enough in the OISb cells to evoke a full cell cycle arrest. However, we do see other proteins and sites differentially regulated that can be either directly or indirectly linked to the OIS onset and maintenance. A summary of these OIS specific events is displayed in figure 5B.

A frequently observed feature of proteins and sites regulated in OIS is the involvement in the extracellular matrix compartment that is known to play an important role in auto and paracrine cell signalling. The ECM can function as a sequestering binding site and pool of growth factors that can be released upon ECM degradation, as is the case for heparin and heparan sulfate (HS) proteoglycans (HSPG) and fibroblast growth factors (55). Moreover, the interaction of extracellular matrix components heparin and HS with FGF2 result in enhanced receptor tyrosine kinase signalling (56, 57). Interestingly, proteins specifically regulated in the senescence phenotype contain ECM interacting (i.e. FGF2, CD44), ECM protein cleaving (i.e. MMP1,DPP4,THBD), ECM cleavage inhibitor (i.e. TIMP3,TFPI2,SerpinB2/PAI-2), glycan processing (i.e. AGA, GLB1/b-Galactosidase, PGM2L1) and membrane lipid

processing proteins (i.e. GAA, PSAP). From our data it is clear that a substantial amount of OIS regulated proteins are acting on the extra-cellular space and matrix. Moreover, other studies have shown extracellular communication through protein secretion is the driving force behind the OIS phenotype (58, 59). Therefore, functional studies to ECM regulation, organization, ECM-receptor and ECM-growth factor interaction could further increase our understanding of OIS and cancer related cell signalling in general.

In summary, this large dataset describes differential protein and phosphorylation changes upon oncogene transduction and oncogene-induced senescence that allowed for comprehensive monitoring of tumor suppressive mechanisms combining known and identifying several novel interesting links to OIS, providing a new confident target list for cancer research.

Supporting Information

Supplementary Figures are located in Chapter 8; Supplementary Figures to Chapter 5. Supplementary Tables can be accesed online at http://mcponline.org.

Acknowledgements

The Netherlands Proteomics Centre, embedded in The Netherlands Genomics Initiative is acknowledged for financial support as is the Centre for Biomedical Genetics. This work was further supported in part by the PRIME-XS project, Grant Agreement Number 262067, funded by the European Union Seventh Framework Program. We acknowledge financial support of the European commission FP7 Markage (FP7-Health-2008-200880), DNA Repair (LSHG-CT-2005-512113) and LifeSpan (LSHG-CT-2007-036894), National Institute of Health (NIH)/National Institute of Ageing (NIA) (1PO1 AG-17242-02), NIEHS (1UO1 ES011044), the Royal Academy of Arts and Sciences of the Netherlands (academia professorship to JHJH) and a European Research Council Advanced Grant to JHJH. The research leading to these results has received funding from the European Community's Seventh Framework Programme (FP7/2007-2013) under grant agreement No. HEALTH-F2-2010-259893.

References

1. Cotter, T. G. (2009) Apoptosis and cancer: the genesis of a research field. Nat Rev Cancer 9, 501-507.

2. Hayflick, L. (1965) The Limited in Vitro Lifetime of Human Diploid Cell Strains. Exp Cell Res 37, 614-636.

3. Deng, Y., Chan, S. S., and Chang, S. (2008) Telomere dysfunction and tumour suppression: the senescence connection. Nat Rev Cancer 8, 450-458.

4. Serrano, M., Lin, A. W., McCurrach, M. E., Beach, D., and Lowe, S. W. (1997) Oncogenic ras provokes premature cell senescence associated with accumulation of p53 and p16INK4a. Cell 88, 593-602.

5. Zhu, J., Woods, D., McMahon, M., and Bishop, J. M. (1998) Senescence of human fibroblasts induced by oncogenic Raf. Genes Dev 12, 2997-3007.

6. Dimri, G. P., Itahana, K., Acosta, M., and Campisi, J. (2000) Regulation of a senescence checkpoint response by the E2F1 transcription factor and p14(ARF) tumor suppressor. Mol Cell Biol 20, 273-285.

7. Michaloglou, C., Vredeveld, L. C., Soengas, M. S., Denoyelle, C., Kuilman, T., van der Horst, C. M., Majoor, D. M., Shay, J. W., Mooi, W. J., and Peeper, D. S. (2005) BRAFE600-associated senescence-like cell cycle arrest of human naevi. Nature 436, 720-724.

8. Collado, M., and Serrano, M. (2010) Senescence in tumours: evidence from mice and humans. Nat Rev Cancer 10, 51-57.

9. Kuilman, T., Michaloglou, C., Mooi, W. J., and Peeper, D. S. (2010) The essence of senescence. Genes Dev 24, 2463-2479.

10. Davies, H., Bignell, G. R., Cox, C., Stephens, P., Edkins, S., Clegg, S., Teague, J., Woffendin, H., Garnett, M. J., Bottomley, W., Davis, N., Dicks, E., Ewing, R., Floyd, Y., Gray, K., Hall, S., Hawes, R., Hughes, J., Kosmidou, V., Menzies, A., Mould, C., Parker, A., Stevens, C., Watt, S., Hooper, S., Wilson, R., Jayatilake, H., Gusterson, B. A., Cooper, C., Shipley, J., Hargrave, D., Pritchard-Jones, K., Maitland, N., Chenevix-Trench, G., Riggins, G. J., Bigner, D. D., Palmieri, G., Cossu, A., Flanagan, A., Nicholson, A., Ho, J. W., Leung, S. Y., Yuen, S. T., Weber, B. L., Seigler, H. F., Darrow, T. L., Paterson, H., Marais, R., Marshall, C. J., Wooster, R., Stratton, M. R., and Futreal, P. A. (2002) Mutations of the BRAF gene in human cancer. Nature 417, 949-954.

11. Pollock, P. M., Harper, U. L., Hansen, K. S., Yudt, L. M., Stark, M., Robbins, C. M., Moses, T. Y., Hostetter, G., Wagner, U., Kakareka, J., Salem, G., Pohida, T., Heenan, P., Duray, P., Kallioniemi, O., Hayward, N. K., Trent, J. M., and Meltzer, P. S. (2003) High frequency of BRAF mutations in nevi. Nat Genet 33, 19-20.

12. Dankort, D., Curley, D. P., Cartlidge, R. A., Nelson, B., Karnezis, A. N., Damsky, W. E., Jr., You, M. J., DePinho, R. A., McMahon, M., and Bosenberg, M. (2009) Braf(V600E) cooperates with Pten loss to induce metastatic melanoma. Nat Genet 41, 544-552.

13. Dhomen, N., Reis-Filho, J. S., da Rocha Dias, S., Hayward, R., Savage, K., Delmas, V., Larue, L., Pritchard, C., and Marais, R. (2009) Oncogenic Braf induces melanocyte senescence and melanoma in mice. Cancer Cell 15, 294-303.

14. Kuilman, T., Michaloglou, C., Vredeveld, L. C., Douma, S., van Doorn, R., Desmet, C. J., Aarden, L. A., Mooi, W. J., and Peeper, D. S. (2008) Oncogene-induced senescence relayed by an interleukin-dependent inflammatory network. Cell 133, 1019-1031.

15. Altelaar, A. F., Munoz, J., and Heck, A. J. (2013) Next-generation proteomics: towards an integrative view of proteome dynamics. Nat Rev Genet 14, 35-48.

16. de Graaf, E. L., Vermeij, W. P., de Waard, M. C., Rijksen, Y., van der Pluijm, I., Hoogenraad, C. C., Hoeijmakers, J. H., Altelaar, A. F., and Heck, A. J. (2013) Spatio-temporal analysis of molecular determinants of neuronal degeneration in the aging mouse cerebellum. Molecular & cellular proteomics : MCP 12, 1350-1362.

17. Gauci, S., Helbig, A. O., Slijper, M., Krijgsveld, J., Heck, A. J., and Mohammed, S. (2009) Lys-N and trypsin cover complementary parts of the phosphoproteome in a refined SCX-based approach. Anal Chem 81, 4493-4501.

18. Zhou, H., Low, T. Y., Hennrich, M. L., van der Toorn, H., Schwend, T., Zou, H., Mohammed, S., and Heck, A. J. (2011) Enhancing the identification of phosphopeptides from putative basophilic kinase substrates using Ti (IV) based IMAC enrichment. Molecular & cellular proteomics : MCP 10, M110 006452.

19. Boersema, P. J., Raijmakers, R., Lemeer, S., Mohammed, S., and Heck, A. J. (2009) Multiplex peptide stable isotope dimethyl labeling for quantitative proteomics. Nat Protoc 4, 484-494.

20. Zoumaro-Djayoon, A. D., Heck, A. J., and Munoz, J. (2012) Targeted analysis of tyrosine phosphorylation by immuno-affinity enrichment of tyrosine phosphorylated peptides prior to mass spectrometric analysis. Methods 56, 268-274.

21. Di Palma, S., Zoumaro-Djayoon, A., Peng, M., Post, H., Preisinger, C., Munoz, J., and Heck, A. J. (2013) Finding the same needles in the haystack? A comparison of phosphotyrosine peptides enriched by immuno-affinity precipitation and metal-based affinity chromatography. J Proteomics 91C, 331-337.

22. Gobom, J., Nordhoff, E., Mirgorodskaya, E., Ekman, R., and Roepstorff, P. (1999) Sample purification and preparation technique based on nano-scale reversed-phase columns for the sensitive analysis of complex peptide mixtures by matrix-assisted laser desorption/ionization mass spectrometry. J Mass Spectrom 34, 105-116.

23. Zhou, H., Ye, M., Dong, J., Han, G., Jiang, X., Wu, R., and Zou, H. (2008) Specific phosphopeptide enrichment with immobilized titanium ion affinity chromatography adsorbent for phosphoproteome analysis. Journal of proteome research 7, 3957-3967.

24. Frese, C. K., Altelaar, A. F. M., Hennrich, M. L., Nolting, D., Zeller, M., Griep-Raming, J., Heck, A. J. R., and Mohammed, S. (2011) Improved Peptide Identification by Targeted Fragmentation Using CID, HCD and ETD on an LTQ-Orbitrap Velos. J Proteome Res.

25. Taus, T., Kocher, T., Pichler, P., Paschke, C., Schmidt, A., Henrich, C., and Mechtler, K. (2011) Universal and confident phosphorylation site localization using phosphoRS. Journal of proteome research 10, 5354-5362.

26. Kall, L., Canterbury, J. D., Weston, J., Noble, W. S., and MacCoss, M. J. (2007) Semi-supervised learning for peptide identification from shotgun proteomics datasets. Nature methods 4, 923-925.

27. Xin, H., Pereira-Smith, O. M., and Choubey, D. (2004) Role of IFI 16 in cellular senescence of human fibroblasts. Oncogene 23, 6209-6217.

28. Courtois-Cox, S., Genther Williams, S. M., Reczek, E. E., Johnson, B. W., McGillicuddy, L. T., Johannessen, C. M., Hollstein, P. E., MacCollin, M., and Cichowski, K. (2006) A negative feedback signaling network underlies oncogene-induced senescence. Cancer Cell 10, 459-472.

29. Shiloh, Y. (2003) ATM and related protein kinases: safeguarding genome integrity. Nat Rev Cancer 3, 155-168.

30. Morgan, D. O. (1995) Principles of CDK regulation. Nature 374, 131-134.

31. Mueller, P. R., Coleman, T. R., Kumagai, A., and Dunphy, W. G. (1995) Myt1: a membrane-associated inhibitory kinase that phosphorylates Cdc2 on both threonine-14 and tyrosine-15. Science 270, 86-90.

32. Wu, Z. Q., and Liu, X. (2008) Role for Plk1 phosphorylation of Hbo1 in regulation of replication licensing. Proceedings of the National Academy of Sciences of the United States of America 105, 1919-1924.

33. Marklund, U., Brattsand, G., Shingler, V., and Gullberg, M. (1993) Serine 25 of oncoprotein 18 is a major cytosolic target for the mitogen-activated protein kinase. J Biol Chem 268, 15039-15047.

34. Sebastian, T., Malik, R., Thomas, S., Sage, J., and Johnson, P. F. (2005) C/EBPbeta cooperates with RB:E2F to implement Ras(V12)-induced cellular senescence. Embo J 24, 3301-3312.

35. Homig-Holzel, C., van Doorn, R., Vogel, C., Germann, M., Cecchini, M. G., Verdegaal, E., and Peeper, D. S. (2011) Antagonistic TSC22D1 variants control BRAF(E600)-induced senescence. Embo J 30, 1753-1765.

36. Martien, S., Pluquet, O., Vercamer, C., Malaquin, N., Martin, N., Gosselin, K., Pourtier, A., and Abbadie, C. (2013) Cellular senescence involves an intracrine prostaglandin E pathway in human fibroblasts. Biochim Biophys Acta.

37. Costa, E. T., Forti, F. L., Matos, T. G., Dermargos, A., Nakano, F., Salotti, J., Rocha, K. M., Asprino, P. F., Yoshihara, C. K., Koga, M. M., and Armelin, H. A. (2008) Fibroblast growth factor 2 restrains Ras-driven proliferation of malignant cells by triggering RhoA-mediated senescence. Cancer Res 68, 6215-6223.

38. Richardson, L. A., Reed, B. J., Charette, J. M., Freed, E. F., Fredrickson, E. K., Locke, M. N., Baserga, S. J., and Gardner, R. G. (2012) A conserved deubiquitinating enzyme controls cell growth by regulat-

ing RNA polymerase I stability. Cell Rep 2, 372-385.

39. Taillebourg, E., Gregoire, I., Viargues, P., Jacomin, A. C., Thevenon, D., Faure, M., and Fauvarque, M. O. (2012) The deubiquitinating enzyme USP36 controls selective autophagy activation by ubiquitinated proteins. Autophagy 8, 767-779.

40. Carmo-Fonseca, M., Ferreira, J., and Lamond, A. I. (1993) Assembly of snRNP-containing coiled bodies is regulated in interphase and mitosis--evidence that the coiled body is a kinetic nuclear structure. J Cell Biol 120, 841-852.

41. Alvarez, J. V., Febbo, P. G., Ramaswamy, S., Loda, M., Richardson, A., and Frank, D. A. (2005) Identification of a genetic signature of activated signal transducer and activator of transcription 3 in human tumors. Cancer Res 65, 5054-5062.

42. Garcia, R., Bowman, T. L., Niu, G., Yu, H., Minton, S., Muro-Cacho, C. A., Cox, C. E., Falcone, R., Fairclough, R., Parsons, S., Laudano, A., Gazit, A., Levitzki, A., Kraker, A., and Jove, R. (2001) Constitutive activation of Stat3 by the Src and JAK tyrosine kinases participates in growth regulation of human breast carcinoma cells. Oncogene 20, 2499-2513.

43. Wang, T., Niu, G., Kortylewski, M., Burdelya, L., Shain, K., Zhang, S., Bhattacharya, R., Gabrilovich, D., Heller, R., Coppola, D., Dalton, W., Jove, R., Pardoll, D., and Yu, H. (2004) Regulation of the innate and adaptive immune responses by Stat-3 signaling in tumor cells. Nat Med 10, 48-54.

44. Blaydes, J. P., Luciani, M. G., Pospisilova, S., Ball, H. M., Vojtesek, B., and Hupp, T. R. (2001) Stoichiometric phosphorylation of human p53 at Ser315 stimulates p53-dependent transcription. J Biol Chem 276, 4699-4708.

45. Katayama, H., Sasai, K., Kawai, H., Yuan, Z. M., Bondaruk, J., Suzuki, F., Fujii, S., Arlinghaus, R. B., Czerniak, B. A., and Sen, S. (2004) Phosphorylation by aurora kinase A induces Mdm2-mediated destabilization and inhibition of p53. Nat Genet 36, 55-62.

46. Bozhenok, L., Wade, P. A., and Varga-Weisz, P. (2002) WSTF-ISWI chromatin remodeling complex targets heterochromatic replication foci. Embo J 21, 2231-2241.

47. Kitagawa, H., Fujiki, R., Yoshimura, K., Mezaki, Y., Uematsu, Y., Matsui, D., Ogawa, S., Unno, K., Okubo, M., Tokita, A., Nakagawa, T., Ito, T., Ishimi, Y., Nagasawa, H., Matsumoto, T., Yanagisawa, J., and Kato, S. (2003) The chromatin-remodeling complex WINAC targets a nuclear receptor to promoters and is impaired in Williams syndrome. Cell 113, 905-917.

48. Xiao, A., Li, H., Shechter, D., Ahn, S. H., Fabrizio, L. A., Erdjument-Bromage, H., Ishibe-Murakami, S., Wang, B., Tempst, P., Hofmann, K., Patel, D. J., Elledge, S. J., and Allis, C. D. (2009) WSTF regulates the H2A.X DNA damage response via a novel tyrosine kinase activity. Nature 457, 57-62.

49. Marrero-Diaz, R., Bravo-Cordero, J. J., Megias, D., Garcia, M. A., Bartolome, R. A., Teixido, J., and Montoya, M. C. (2009) Polarized MT1-MMP-CD44 interaction and CD44 cleavage during cell retraction reveal an essential role for MT1-MMP in CD44-mediated invasion. Cell Motil Cytoskeleton 66, 48-61.

50. Prince, M. E., Sivanandan, R., Kaczorowski, A., Wolf, G. T., Kaplan, M. J., Dalerba, P., Weissman, I. L., Clarke, M. F., and Ailles, L. E. (2007) Identification of a subpopulation of cells with cancer stem cell properties in head and neck squamous cell carcinoma. Proceedings of the National Academy of Sciences of the United States of America 104, 973-978.

51. Li, C., Heidt, D. G., Dalerba, P., Burant, C. F., Zhang, L., Adsay, V., Wicha, M., Clarke, M. F., and Simeone, D. M. (2007) Identification of pancreatic cancer stem cells. Cancer Res 67, 1030-1037.

52. Hara, E., Smith, R., Parry, D., Tahara, H., Stone, S., and Peters, G. (1996) Regulation of p16CDKN2 expression and its implications for cell immortalization and senescence. Mol Cell Biol 16, 859-867.

53. Shay, J. W., Pereira-Smith, O. M., and Wright, W. E. (1991) A role for both RB and p53 in the regulation of human cellular senescence. Exp Cell Res 196, 33-39.

54. Campisi, J., and d'Adda di Fagagna, F. (2007) Cellular senescence: when bad things happen to good cells. Nat Rev Mol Cell Biol 8, 729-740.

55. Flaumenhaft, R., Moscatelli, D., and Rifkin, D. B. (1990) Heparin and heparan sulfate increase the radius of diffusion and action of basic fibroblast growth factor. J Cell Biol 111, 1651-1659.

56. Moy, F. J., Safran, M., Seddon, A. P., Kitchen, D., Bohlen, P., Aviezer, D., Yayon, A., and Powers, R. (1997) Properly oriented heparin-decasaccharide-induced dimers are the biologically active form of basic fibroblast growth factor. Biochemistry 36, 4782-4791.

57. Rapraeger, A. C., Krufka, A., and Olwin, B. B. (1991) Requirement of heparan sulfate for bFGF-mediated fibroblast growth and myoblast differentiation. Science 252, 1705-1708.

58. Krtolica, A., and Campisi, J. (2002) Cancer and aging: a model for the cancer promoting effects of the aging stroma. Int J Biochem Cell Biol 34, 1401-1414.

59. Kuilman, T., and Peeper, D. S. (2009) Senescence-messaging secretome: SMS-ing cellular stress. Nat Rev Cancer 9, 81-94.

Chapter 6

Outlook

Future perspectives on cell signalling analysis

Decades ago the discovery of new proteins or protein modifications posed a rate limiting step in gaining knowledge on biological processes. The amino acid sequencing of the first protein (Insulin) by Sanger and colleagues was a considerable achievement that took more than ten years. (1) The classical approach used in these days to track the molecular basis of biological processes, started with the isolation of proteins by extensive fractionation of cell extracts followed by purification to obtain single protein species. When this labour intensive work was performed, the actual determination of the protein structure and function could start.

The introduction of so-called 'omics' technologies has greatly revolutionized the discovery of new molecular entities; i.e. DNA, RNA and proteins. Nowadays, whole genome, transcriptome or proteome analysis is performed within a week to a month. As a result large data repositories are accessible to the whole scientific community and literature is overwhelmed with lengthy lists of genes and proteins associated to different phenotypes or diseases. Every list originating from large discovery experiments contains however a number of random false hits. The false discovery of proteins or post translational modified (PTM) protein isoforms can be attributed to random false identifications, uncontrollable biological variability and the large number of hypothesizes that are being tested in a small number of samples. However, to minimize the amount of falsely selected targets several strategies have been developed to control false positive peptide identification (2-4) and quantification rates (5-7). As a result of this, currently published proteomics data is filtered to statistically-sound target lists that need none to little identification or quantification validation by other techniques.

The high confidence putative molecular determinants discovered by shotgun proteomics can be subsequently used as new candidate biomarkers for diagnostic purposes or studied further to investigate their role in biological processes. Paradoxically however, the exponential increase in number and complexity of these candidate lists has yielded little new definite biomarkers and proved to be challenging in contributing new knowledge to biological systems. This gap between target discovery and validation can be largely explained by several challenges that occur in the subsequent evaluation stages of the proteomics' generated target lists.

One of the main challenges is caused by the fact that confidently differentially expressed proteins in disease and control samples are not necessarily linked to the phenotype. In other words, when the protein (de-)regulation is established to be disease specific, it still remains ambiguous if the protein is the cause or simply a result of a disease. Therefore, to explore the biological roles of proteins resident on the target list, classical labor intensive follow-up experiments and several validation steps are required (Figure 1).

The first follow-up stage involves the validation of the lengthy target list. Initially, discovery experiments are performed using a limited number of easily available samples. Therefore targets need to be validated in more samples from different origin. This is especially the case in biomarker validation where candidate biomarkers need to be monitored in a whole

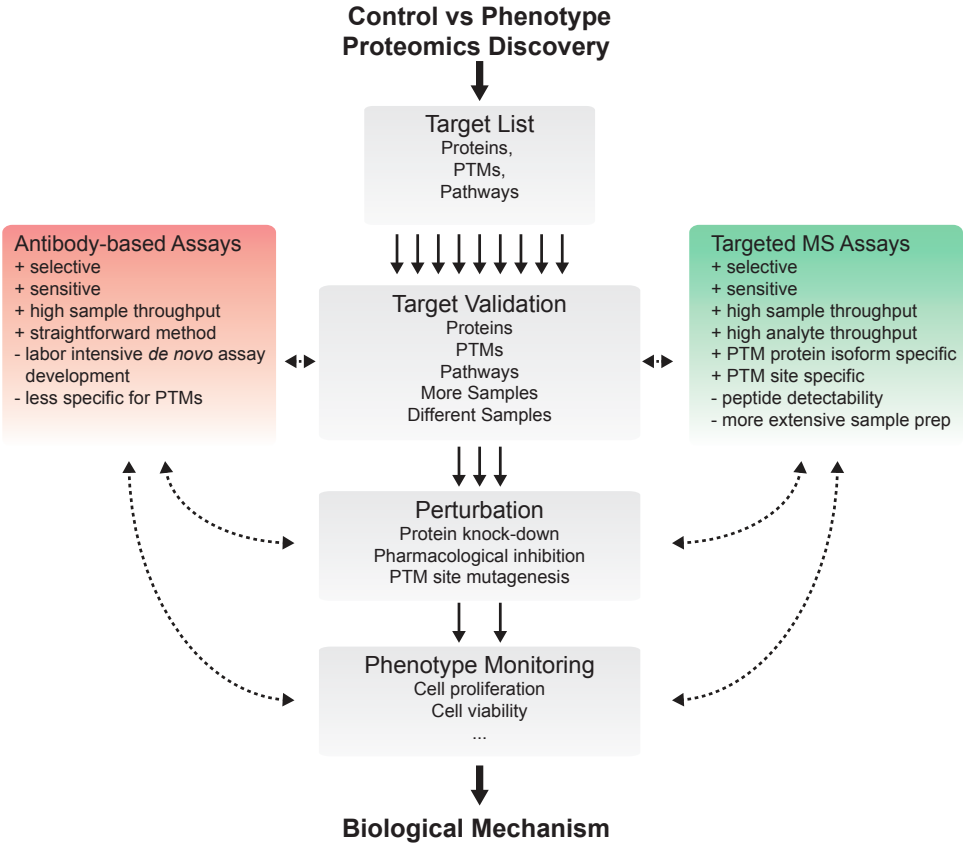

Figure 1: *The bottleneck of gaining insights in biological mechanisms associated with disease is currently formed by the lack of high-quality antibody-based protein and PTM assays. In the near future targeted mass spectrometry-based assays may pose an interesting alternative for protein and PTM regulation validation. Especially, regarding PTM analysis the later technique can offer more selective and precise measurements.*

range of different phenotypes originating from tens to hundreds of samples to evaluate their predictive diagnostic power. This validation is further challenged by low abundant candidates that can contain single protein hits, single PTM sites or full pathways and associated PTMs. As discovery experiments often lack the selectivity and reproducibility to comprehensively measure all important nodes and associated PTM signaling events in a pathway, more specific and selective assays are required to validate their importance.

Other challenges in target validation include monitoring of changes in protein, PTM or pathway expression dynamics after perturbation (Figure 1). The scale of a perturbation experiment can range from low throughput single protein mutagenesis to high throughput RNAi/drug screenings for protein function (8, 9) and pathway screening (10). Using the latter strategies a long list of targets can be tested for phenotype alteration involving a high number of samples. Therefore, to validate the link between phenotype perturbation and biological mechanisms, the molecular entities encompassing the mechanism need to be

monitored in a comprehensive and similarly high throughput manner.

The major technical bottleneck restricting the throughput of these candidate validation steps is caused by the lack of quantitative assays for the majority of human proteins and even more so for protein modifications. (11) Furthermore, the de novo development for protein/PTM assays typically involves the generation of antibody-based assays (i.e. immunoblotting and enzyme-linked immunosorbent assays (ELISA)) that are prohibitively expensive for a large number of targets. On top of this, when studying protein phosphorylation events, several other factors discourage the use of antibodies.

Kinase and regulatory domains are highly conserved in sequence and structure. Therefore, protein isoform specific antibodies are needed, that are directed against unique protein sequences/domains. However, important PTMs are often located in highly conserved functional domains, hampering the generation of antibodies selective for PTMs from different protein isoforms. Exemplary proteins are the two Erk isoforms for which an antibody is directed against the phosphosites Thr202 and Tyr204 on Erk1 that also recognize Thr185 and Tyr187 on Erk2. Therefore, while immunoblotting is highly sensitive, antibodies against the singly phosphorylated ERK forms are often of limited specificity.(12, 13) Moreover, protein phosphorylation often occurs sequentially on different close-proximity sites that cannot be easily distinguished by an antibody. This phosphorylation crosstalk is a common mechanism in many crucial growth signaling proteins including for example CDK1/2/3 (Thr14 and Tyr15) (14), Erk1/2 (Thr202, Tyr204 and Thr185, Tyr187) and Retinoblastoma (Ser249, Thr252 and S807, S811). When using antibodies to study these signalling dynamics it is often unclear which site(s) the antibody will recognize. This could mean that the assay is blind for differences between mutually exclusive singly or doubly phosphorylated domains. Similarly, it remains unclear if the binding affinity of the specific antibody varies when different sites within the recognition domain are changing. Therefore, a signal decrease can be a result of total dephosphorylation or partial (single site) dephosphorylation. Furthermore, little is known about changes in antibody binding affinity when additional unknown close-proximity PTMs appear in vivo that have not been tested in vitro. Although recent efforts have been made to increase antibody specificity for phosphorylated peptides by in silico antibody design (15), the antibody binding pocket, intrinsically recognizing 5-7 amino acid stretches, will always hamper specificity for close proximity phosphosite isoforms.

The field of targeted mass spectrometry forms an interesting and potent antibody-free alternative for protein and PTM validation.(16-18) Multiple reaction monitoring (MRM, or SRM) is a powerful technique that needs relatively low amount of resources for de novo protein/PTM assay development when compared to antibody-based assays.(19, 20) The number of mass-spectrometric protein assays is increasing rapidly and soon to overtake the amount of immunoassays available. The recent documentation of MRM protein assays for the near-complete yeast proteome is illustrative for the ease and pace at which assays can be generated.(21) MRM sample throughput is similar to immunological assays whereas analyte multiplexing can be even much higher for MRM (typically 1-3 and 10-100 proteins per experiment for ELISA and MRM, respectively). Another important property of pre-

cise validation assays is selectivity, which is especially important for PTM analysis. Similar PTMs on different protein isoforms with high sequence homology and similar structural properties can hardly be distinguished by antibody binding. However, due to small differences in protein isoform sequences, protein isoforms can be easily separated by LC-MS-based assays that are based on unique protein sequences. (22, 23) An exemplary case is the phosphorylation of the TEY-motif of the Erk1 and Erk2 paralogues that differ 28 Da in mass and therefore are easily distinguished in a MS-based assay (Erk1: IADPEHDHTGFLpTEpY-VATR; 2331.01 Da, Erk2: VADPDHDHTGFLpTEpYVATR; 2302.99 Da).(22) Finally, contrary to antibody-based methods, close proximity sites can be distinguished by LC-MS, as singly and multiply phosphorylated peptides can be separated in the mass spectrometer by mass increments of ~80 Da and phosphoisomers of identical mass can be distinguished using liquid chromatography and/or peptide fragment-specific analysis.(13)

In summary, MS-based validation approaches can study close-proximity phosphosites and therefore protein phospho-form distributions much more accurately, allowing for more precise signalling measurements and more in depth biological conclusions. (13, 22)

Challenges that remain eminent for MS-based validation assays are peptide/PTM detectability, sensitivity and reproducibility. Not all peptide sequences are directly measurable by using the standard MS workflow. In some cases PTMs can reside in peptide sequences either short or in excess of tryptic cleavage sites, leading to peptide lengths unsuitable for MS analysis. Moreover, when charge directed digestion enzymes such as trypsin are used, nearby negatively charged phosphosites can mask positively charged arginine or lysine cleavage sites, potentially leading to irreproducible miscleavage rates. To overcome this problem different digestion enzymes can be used to generate suitable peptide lengths and unaffected cleavage rates.

Antibody-based assays typically are ultra-sensitive (down to picograms) and need little sample preparation (immunoblot: sample lysis, ELISA: cell/tissue fixation). However, MS-based methods are more prone to signal interference and sample preparation workflows are generally more complex. Highly abundant proteins present in the sample can affect MS-assay sensitivity by suppressing ionization or introducing background interference. Different techniques have been developed to minimize matrix effects and increase analyte sensitivity. These strategies include e.g.; ion transmission optimization (24), multiple peptide fragmentation rounds (MRM3) (25), abundant protein depletion (17) and extensive sample fractionation (PRISM) (26). However, the addition of several sample processing steps potentially compromises assay reproducibility, placing high constrains on sample handling workflows. Moreover, for PTMs, that are present at low endogenous and sub stoichiometric levels, additional enrichment steps are required. As a result of this, only low-throughput phosphosite validation experiments have been described using targeted MS, mainly analysing highly abundant phosphopeptides.(19, 27-29)

Recent developments in quantitative reproducible phosphopeptide enrichment (30) (Chapter 4) should allow the combination of phosphopeptide enrichment protocols with targeted mass spectrometry. Moreover, the generation of PTM-specific MRM assays is straightforward therefore the generation of a comprehensive set of MRM assays is just a matter of

time. Individually, the MRM and phosphopeptide enrichment are well established and therefore a successful marriage in the near future will be highly probable. The combination of both techniques will likely solve the bottleneck of phosphosite validation throughput resulting in routine full cell signalling pathway analysis paving the way for highly comprehensive systems biology screenings. Finally, when combined with the perturbation and phenotype screening, targeted mass spectrometry-based PTM analysis has a high potential to greatly increase the understanding of the biological mechanisms underlying phenotype/disease manifestation.

References

1. Sanger, F. (1988) Sequences, sequences, and sequences. Annu Rev Biochem 57, 1-28.

2. Moore, R. E., Young, M. K., and Lee, T. D. (2002) Qscore: an algorithm for evaluating SEQUEST database search results. Journal of the American Society for Mass Spectrometry 13, 378-386.

3. Peng, J., Elias, J. E., Thoreen, C. C., Licklider, L. J., and Gygi, S. P. (2003) Evaluation of multidimensional chromatography coupled with tandem mass spectrometry (LC/LC-MS/MS) for large-scale protein analysis: the yeast proteome. Journal of proteome research 2, 43-50.

4. Kall, L., Canterbury, J. D., Weston, J., Noble, W. S., and MacCoss, M. J. (2007) Semi-supervised learning for peptide identification from shotgun proteomics datasets. Nature methods 4, 923-925.

5. Galitski, T., Saldanha, A. J., Styles, C. A., Lander, E. S., and Fink, G. R. (1999) Ploidy regulation of gene expression. Science 285, 251-254.

6. Tusher, V. G., Tibshirani, R., and Chu, G. (2001) Significance analysis of microarrays applied to the ionizing radiation response. Proceedings of the National Academy of Sciences of the United States of America 98, 5116-5121.

7. Cox, J., and Mann, M. (2008) MaxQuant enables high peptide identification rates, individualized p.p.b.-range mass accuracies and proteome-wide protein quantification. Nat Biotechnol 26, 1367-1372.

8. Kamath, R. S., Fraser, A. G., Dong, Y., Poulin, G., Durbin, R., Gotta, M., Kanapin, A., Le Bot, N., Moreno, S., Sohrmann, M., Welchman, D. P., Zipperlen, P., and Ahringer, J. (2003) Systematic functional analysis of the Caenorhabditis elegans genome using RNAi. Nature 421, 231-237.

9. Ooi, S. L., Shoemaker, D. D., and Boeke, J. D. (2003) DNA helicase gene interaction network defined using synthetic lethality analyzed by microarray. Nat Genet 35, 277-286.

10. Prahallad, A., Sun, C., Huang, S., Di Nicolantonio, F., Salazar, R., Zecchin, D., Beijersbergen, R. L., Bardelli, A., and Bernards, R. (2012) Unresponsiveness of colon cancer to BRAF(V600E) inhibition through feedback activation of EGFR. Nature 483, 100-103.

11. Anderson, N. L., Anderson, N. G., Pearson, T. W., Borchers, C. H., Paulovich, A. G., Patterson, S. D., Gillette, M., Aebersold, R., and Carr, S. A. (2009) A human proteome detection and quantitation project. Molecular & cellular proteomics : MCP 8, 883-886.

12. Yao, Z., Dolginov, Y., Hanoch, T., Yung, Y., Ridner, G., Lando, Z., Zharhary, D., and Seger, R. (2000) Detection of partially phosphorylated forms of ERK by monoclonal antibodies reveals spatial regulation of ERK activity by phosphatases. FEBS Lett 468, 37-42.

13. Prabakaran, S., Everley, R. A., Landrieu, I., Wieruszeski, J. M., Lippens, G., Steen, H., and Gunawardena, J. (2011) Comparative analysis of Erk phosphorylation suggests a mixed strategy for measuring phospho-form distributions. Mol Syst Biol 7, 482.

14. Morgan, D. O. (1995) Principles of CDK regulation. Nature 374, 131-134.

15. Koerber, J. T., Thomsen, N. D., Hannigan, B. T., Degrado, W. F., and Wells, J. A. (2013) Nature-inspired design of motif-specific antibody scaffolds. Nature biotechnology.

16. Anderson, L., and Hunter, C. L. (2006) Quantitative mass spectrometric multiple reaction monitoring assays for major plasma proteins. Molecular & cellular proteomics : MCP 5, 573-588.

17. Keshishian, H., Addona, T., Burgess, M., Kuhn, E., and Carr, S. A. (2007) Quantitative, multiplexed assays for low abundance proteins in plasma by targeted mass spectrometry and stable isotope dilution. Molecular & cellular proteomics : MCP 6, 2212-2229.

18. Lange, V., Picotti, P., Domon, B., and Aebersold, R. (2008) Selected reaction monitoring for quantitative proteomics: a tutorial. Mol Syst Biol 4, 222.

19. Gerber, S. A., Rush, J., Stemman, O., Kirschner, M. W., and Gygi, S. P. (2003) Absolute quantification of proteins and phosphoproteins from cell lysates by tandem MS. Proc Natl Acad Sci U S A 100, 6940-6945.

20. Picotti, P., Bodenmiller, B., and Aebersold, R. (2013) Proteomics meets the scientific method. Nature methods 10, 24-27.

21. Picotti, P., Clement-Ziza, M., Lam, H., Campbell, D. S., Schmidt, A., Deutsch, E. W., Rost, H., Sun, Z., Rinner, O., Reiter, L., Shen, Q., Michaelson, J. J., Frei, A., Alberti, S., Kusebauch, U., Wollscheid, B., Moritz, R. L., Beyer, A., and Aebersold, R. (2013) A complete mass-spectrometric map of the yeast proteome applied to quantitative trait analysis. Nature 494, 266-270.

22. Hahn, B., D'Alessandro, L. A., Depner, S., Waldow, K., Boehm, M. E., Bachmann, J., Schilling, M., Klingmuller, U., and Lehmann, W. D. (2013) Cellular ERK phospho-form profiles with conserved preference for a switch-like pattern. Journal of proteome research 12, 637-646.

23. Tanner, G. J., Colgrave, M. L., Blundell, M. J., Goswami, H. P., and Howitt, C. A. (2013) Measuring hordein (gluten) in beer--a comparison of ELISA and mass spectrometry. PLoS One 8, e56452.

24. Hossain, M., Kaleta, D. T., Robinson, E. W., Liu, T., Zhao, R., Page, J. S., Kelly, R. T., Moore, R. J., Tang, K., Camp, D. G., 2nd, Qian, W. J., and Smith, R. D. (2011) Enhanced sensitivity for selected reaction monitoring mass spectrometry-based targeted proteomics using a dual stage electrodynamic ion funnel interface. Molecular & cellular proteomics : MCP 10, M000062-MCP000201.

25. Lemoine, J., Fortin, T., Salvador, A., Jaffuel, A., Charrier, J. P., and Choquet-Kastylevsky, G. (2012) The current status of clinical proteomics and the use of MRM and MRM(3) for biomarker validation. Expert Rev Mol Diagn 12, 333-342.

26. Shi, T., Fillmore, T. L., Sun, X., Zhao, R., Schepmoes, A. A., Hossain, M., Xie, F., Wu, S., Kim, J. S., Jones, N., Moore, R. J., Pasa-Tolic, L., Kagan, J., Rodland, K. D., Liu, T., Tang, K., Camp, D. G., 2nd, Smith, R. D., and Qian, W. J. (2012) Antibody-free, targeted mass-spectrometric approach for quantification of proteins at low picogram per milliliter levels in human plasma/serum. Proceedings of the National Academy of Sciences of the United States of America 109, 15395-15400.

27. Domanski, D., Murphy, L. C., and Borchers, C. H. (2010) Assay development for the determination of phosphorylation stoichiometry using multiple reaction monitoring methods with and without phosphatase treatment: application to breast cancer signaling pathways. Analytical chemistry 82, 5610-5620.

28. Unwin, R. D., Griffiths, J. R., and Whetton, A. D. (2009) A sensitive mass spectrometric method for hypothesis-driven detection of peptide post-translational modifications: multiple reaction monitoring-initiated detection and sequencing (MIDAS). Nat Protoc 4, 870-877.

29. Lam, M. P., Lau, E., Scruggs, S. B., Wang, D., Kim, T. Y., Liem, D. A., Zhang, J., Ryan, C. M., Faull, K. F., and Ping, P. (2013) Site-specific quantitative analysis of cardiac mitochondrial protein phospho-

rylation. J Proteomics 81, 15-23.

30. Kettenbach, A. N., and Gerber, S. A. (2011) Rapid and reproducible single-stage phosphopeptide enrichment of complex peptide mixtures: application to general and phosphotyrosine-specific phosphoproteomics experiments. Analytical chemistry 83, 7635-7644.

Summary
Samenvatting
Curriculum Vitae
List of Publications
Acknowledgments

Summary

This thesis describes proteomics technology developments and implementations that have been used to increase knowledge about aging and cancer suppression mechanisms. The core tool in all performed experiments was mass spectrometry-based proteomics. In addition to the application of these techniques, newly developed proteomics techniques have been introduced that helped to increase the analytical power of proteomics experiments.

In **chapter one**, a general introduction is given into proteomics and mass spectrometry-based proteomics workflows. The type of research question strongly determines the type of experimental workflow and equipment needed. Therefore, there is no standard single method resulting in a highly dynamic experimental workflow with a myriad of different techniques/approaches.

In **chapter two**, the influence of DNA-damage on tissue aging was studied over time, for a specific part of the mouse brain. Quantitative shotgun proteomics discovery experiments were performed in order to find biomarkers and mechanisms of aging in the cerebellum. After the identification of several interesting differentially regulated proteins associated with age and neuron-deterioration by proteomics, a more targeted validation was performed using immunohistochemistry. By using this complementary method, changes over time as well as changes in protein localization could be monitored for a set of differentially regulated proteins. Major observations of this study include a DNA-damage induced downregulation of a tightly interconnected synaptic signaling network associated with a morphological transformation of Purkinje cells and a new molecular link between DNA-damage and motoric diseases such as spinocerebellar ataxia.

In **chapter three** a technical method development is described that optimizes the transition from a discovery proteomics experiment to a targeted proteomics approach. In discovery experiments often ion trap CID or higher energy CID (HCD) are used to fragment and identify peptides. These two fragmentation methods were compared to the fragmentation technique used in targeted proteomics single reaction monitoring experiments (i.e. SRM, also referred to as MRM). Peptide fragmentation spectra comparisons as well as fragmentation specific spectral libraries were used to assess the best way to develop discovery-based SRM assays. The work demonstrates that CID spectra acquired on SRM instruments are more similar to spectra acquired in the HCD cell than those acquired in the ion trap of LTQ-Orbitrap instruments. Concomitantly, SRM assays generated using HCD spectra showed a higher sensitivity when compared to ion trap spectra-generated SRM assays. Therefore, when planning a targeted MS experiment, choosing for HCD fragmentation in the discovery phase can help facilitate SRM assay development later on.

In **chapter four**, the phosphopeptide enrichment robustness of a new Ti^{4+}-IMAC method was assessed. First it was established that Ti^{4+}-IMAC enrichment resulted in a highly repro-

ducible quantification of phosphorylation sites in HeLa cells. Subsequently, this strategy was applied to monitor the phosphoproteome of Jurkat T-cells upon Prostaglandin E_2 stimulation over the timespan of 60 minutes. It was demonstrated that using this enrichment strategy and label-free quantification a comprehensive temporal phosphoproteome of Jurkat T-cells could be constructed, indicating differential regulation of different kinases over time. The proved straight-forward yet comprehensive phosphoproteomics workflow and its' applicability to every sample type could form a good alternative method for the phosphoproteomics community.

Chapter five describes a study performed to increase knowledge about the mechanisms involved in $BRAF^{V600E}$ oncogene-induced senescence (OIS). Using both whole proteome and phosphoproteome analysis, differential protein expression and phosphorylation were compared between cycling, OIS and OIS bypassed (OISb) cells. Proteome analysis revealed a strong resemblance between proliferating BRAF-activated (OISb) cells and senescent BRAF-activated (OIS) cells. In both cell lines an up regulation of the retinoblastoma tumor suppressor mechanism was observed that was more pronounced in OIS. Proteins specifically regulated in OIS included senescence markers, inflammatory proteins and extracellular processing factors amongst other previously senescence-unassociated proteins. Phosphorylation screening identified strong reduction in cyclin-dependent kinase activity and numerous sites to be differentially regulated in OIS. This work forms an extensive resource of protein and phosphorylation regulations associated with OIS, benefitting researchers in the field of tumor suppressor research.

In **chapter six** an outlook is described regarding the validation of proteomics-derived target lists with the emphasis on cell signaling analysis. Currently, antibody-based methods are used predominantly to validate and investigate the involvement of proteins and post translation modifications (PTMs) in different biological processes. Limitation of antibody-based methods are described and targeted mass spectrometry-based methods are proposed as an alternative approach to overcome these problems. Especially for PTM analysis, targeted mass spectrometry has a high potential that will likely become a standard methodology for dynamic cell signaling analysis.

Samenvatting

Dit proefschrift beschrijft nieuwe technologische ontwikkelingen voor (fosofo-)proteoom analyse en haar toepassing om de biologische kennis van verouderings en kanker suppressie processen te verhogen. Massaspectrometrie-gebaseerde analyses vormen de kern van alle experimenten in deze thesis. Naast de toepassing van deze technieken op biologische modelsystemen, worden er ook nieuwe (fofsfo-)proteoom technieken geïntroduceerd om de analytische kracht van experimenten te verhogen.

In het **eerste hoofdstuk** wordt een algemene inleiding gegeven over proteomics and massaspectrometrie. Afhankelijk van het soort onderzoeksvraag kan een aantal keuzes worden gemaakt met betrekking tot monster preparatie, chromatografie, specifieke verrijking en type massaspectrometer. Het grote aantal mogelijkheden zorgt er dus voor dat er geen standaard experiment configuratie is voor elk probleem en dus goed moet worden overwogen welke technieken men het beste kan gebruiken voordat men een experiment initieert.

In het **tweede hoofdstuk** is onderzocht wat de gevolgen zijn van DNA beschadiging in de hersenen van muizen van verschillende leeftijden. Om ouderdom-specifieke biologische processen en eiwitten te vinden is een kwantitatieve globale proteomics screening toegepast op het cerebellum van muizen. Na de identificatie van differentieel gereguleerde eiwitten die geassocieerd worden met leeftijd en neurodegeneratie, is een specifieke validatie uitgevoerd voor een aantal eiwitten door middel van immunohistochemie. Door de combinatie van de twee gebruikte complementaire analytische technieken, kon voor een set eiwitten in het cerebellum in kaart worden gebracht wat hun regulatie over tijd en regulatie in eiwit localisatie was. De belangrijkste observaties die gerelateerd kunnen worden aan DNA beschadigingen in de hersenen zijn een sterke reductie van een netwerk van eiwitten belangrijk voor neuronale communicatie, een sterke verandering in morfologie van geaffecteerde neuronen en een nieuwe moleculaire link tussen beschadigingen in het DNA en motorische ziekten zoals cerebellaire ataxia.

In **hoofdstuk drie** wordt de ontwikkeling van een methode beschreven die kan worden gebruikt om de transitie van een globale naar een doelgerichte proteomics aanpak te faciliteren. In deze studie is een vergelijking gemaakt tussen drie verschillende peptide fragmentatie technieken uitvoerbaar op twee massaspectrometrie platformen, namelijk collision induced dissociation (CID) in de ion trap en higher energy CID (HCD) in de HCD-cell van een LTQ-Orbitrap Velos massaspectrometer en CID in een quadrupole van een triple quadrupole massaspectrometer. Vaak wordt CID in de ion trap gebruikt voor globale screening experimenten en worden vervolgens de vergaarde fragmentatiespectra gebruikt voor het opzetten van single reaction monitoring (SRM) experimenten die worden gedraaid op een triple quadrupole massaspectrometer. In deze studie laten we zien dat CID fragmentatie spectra van een triple quadruple meer gelijkenissen vertonen met HCD spectra dan met ion trap CID spectra vergaard op een LTQ-Orbitrap Velos. Daarnaast laten we zien dat als een gevolg van de betere gelijkenissen, SRM assays gemaakt van HCD spectra een hogere gevoeligheid hebben dan SRM assays gebaseerd op ion trap CID spectra. Wanneer men dus een doelgericht proteomics experiment anticipeert na een globale screening adviseren we

het gebruik van HCD fragmentatie i.p.v. ion trap CID fragmentatie.

In **hoofdstuk vier** hebben we Ti^{4+}-IMAC fosfopeptide verrijking onderzocht op quantitatieve robuustheid. Als eerst werd de kwantitatieve reproduceerbaarheid van deze fosfopeptide verrijkingstechniek gedemonstreerd in HeLa cellen. Vervolgens werd met deze techniek het fosfoproteoom van Prostaglandin E_2 gestimuleerde Jurkat T-cellen gemeten over zes verschillende stimulatie duraties (0-60min). De toepassing van Ti-IMAC fosfopeptide verrijking samen met label-free kwantificering op 54 monsters resulteerde in de creatie van een dynamische en veelomvattende fosfoproteoom dataset, die vervolgens werd gebruikt om de activatie van verschillende kinases over tijd te monitoren. De omschreven robuuste techniek voor fosfopeptide analyse die toepasbaar is op grote hoeveelheden monsters van elke origine vormt een volwaardig alternatief voor wetenschappers in de fosfoproteoom gemeenschap.

In **hoofdstuk vijf** wordt een studie beschreven waarin onderzoek is gedaan naar de mechanismen die een rol spelen in BRAF oncogene-induced senescence (OIS). OIS is een lichaamseigen manier om de wildgroei van weefsel (kanker) te voorkomen. Door de analyse van zowel het proteoom als het fosfoproteoom, konden differentieel gereguleerde eiwitten en eiwit phosphorylaties worden gedecteerd en vergeleken tussen normale, OIS en OIS ontvluchte (OISb) cellen. De analyse van het proteoom onthulde een verrassend grotere gelijkenis tussen de prolifererende en niet prolifererende BRAF-geactiveerde cellen (OISb vs. OIS) dan tussen de prolifererende normale en prolifererende BRAF-geactiveerde cellen (normale vs OISb). Zowel in OIS als in OISb was een toename zichtbaar van het retinoblastoma tumor suppressor mechanisme met een grotere mate in OIS. Eiwitten specifiek gereguleerd in OIS bevatten senescence markers, inflammatoire eiwitten, extracellulaire modificerende enzymen alsmede andere eiwitten eerder niet geassocieerd met senescence. Eiwit fosforylatie analyse identificeerde vele phosphosites speciek voor OIS en een sterke reductie in cyclin-dependent kinase activiteit in OIS. De verkregen dataset vormt een uitgebreide bron van eiwit en fosforylatie regulaties die geassocieerd zijn met OIS en kan worden gebruikt door andere onderzoekers die zich bezig houden met antitumor therapieën.

In **hoofdstuk zes** staat een vooruitzicht op de validatie van cel signalering mechanismen geidentificeerd door globale proteomics screenings. Tegenwoordig worden met name antilichamen gebruikt om gereguleerde eiwitten en post translationele modificaties (PTMs) te valideren. De beperkingen van antilichaam-gebaseerde methoden worden beschreven en doelgerichte massa spectrometrie-gebaseerde methoden worden voorgesteld als alternatieve methoden die veel van deze beperkingen kunnen wegnemen. Doelgerichte massaspectrometrie heeft met name een hoge potentie voor PTM analyse en zal waarschijnlijk een standard methodologie worden voor de analyse van signaal transductie routes in cellen.

Curriculum vitae

Erik de Graaf was born on the third of December 1984 in Heesch, Noord-Brabant, The Netherlands. In 2003 he finished his pre-university education specialized in science & technology and science & health. He started to explore his wide interests in science by studying Natural Sciences at the Radboud University Nijmegen. During the completion of his Bachelor in 2007, he had already started his master in molecular biology and biochemistry that elaborated further to his broad background in physics, chemistry and molecular biology. During his first master project at the department of Applied Biology, under supervision of Dr. Esther Piek, he studied signal transduction cascades involved in the differentiation of bone marrow-derived mesenchymal stem cells into the osteo- and adipogenic lineages, respectively. In the final year of his Master studies, he focussed on processes involved in hormone refractory prostate cancer at the research department of hormone dependent disorders at MSD (former Organon/Schering-Plough, Oss, NL) under the supervision of dr. Patrick Groothuis and prof. dr. Sybrand Wijmenga. Here he identified and inhibited active growth signal transduction pathways in an array of (hormone refractory) prostate cancer cells. During this project he was also introduced to mass spectrometry and got impressed by the power of this technique. Therefore, in 2009, Erik joined the Biomolecular Mass spectrometry and Proteomics group of prof. dr. Albert Heck in Utrecht (NL) where he learned all the aspects of LC-MS and proteomics. The research was mainly focussed on method development and application; studying proteins and protein phosphorylation in aging and senescence in collaboration with prof. dr. Jan Hoeijmakers and prof. dr. Daniel Peeper, respectively. The results of this projects are presented in this doctoral thesis.

List of Publications

de Graaf, E.L.; Altelaar, A.F.; van Breukelen, B.; Mohammed, S.; Heck, A.J., Improving SRM assay development: a global comparison between triple quadrupole, ion trap, and higher energy CID peptide fragmentation spectra. *Journal of proteome research* 2011, 10, (9), 4334-41.

De Vos, D.; Frederiks, F.; Terweij, M.; van Welsem, T.; Verzijlbergen, K.F.; Iachina, E.; **de Graaf, E.L.**; Altelaar, A.F.; Oudgenoeg, G.; Heck, A.J.; Krijgsveld, J.; Bakker, B.M.; van Leeuwen, F., Progressive methylation of ageing histones by Dot1 functions as a timer. *EMBO reports* 2011, 12, (9), 956-62.

de Graaf, E.L.; Vermeij W.P.; de Waard, M.C.; Rijksen, Y.; van der Pluijm, I.; Hoogenraad, C.C.; Hoeijmakers, J.H.; Altelaar, A.F.; Heck, A. J., Spatio-temporal analysis of molecular determinants of neuronal degeneration in the aging mouse cerebellum. Molecular & cellular proteomics: *Molecular & Cellular Proteomics* 2013, 12, (5), 1350-62.

de Graaf, E.L.*; Giansanti, P.*; Heck, A.J.R.; Altelaar, A.F.M., Probing phosphorylation dynamics in signaling networks exploiting Ti4+-IMAC phosphopeptide enrichment. Manuscript in submission.

de Graaf, E.L.; Kaplon, J.; Zhou, H.; Heck, A.J.; Peeper, D.S.; Altelaar, A.F.M., Phosphoproteome dynamics in onset and maintenance of oncogene induced senescence. Manuscript in submission.

* These authors contributed equally.

Acknowledgments

Albert, I would like to thank you for giving me the opportunity to work and have fun in your lab. When I began to meet fellow scientist from other labs, I started to feel more and more privileged to be able to work in this lab with so many different high quality facilities and opportunities. From you I learned how to sell my research and how important politics are in the scientific community. I also would like to thank you for providing me with the opportunity to go on a scientific and personal adventure in Victoria.

Maarten, I think your "daily" supervision style was a perfect fit to me. You gave me directions and new projects and gave me the freedom to explore my own directions and interests. Your shrugging policy forced me to figure out my own problems. You got me excited about working with a lot of external collaborators, showing me the importance of working together using different disciplines.

Shabaz, you have been an unofficial supervisor to me. You taught me all the ins and outs of liquid chromatography and mass spectrometry. Also writing an article and having journal club with you opened my eyes for the extra layers, as well as research alone, that need to be present in a paper suitable for publication. You left the group leaving nobody behind to arrange Friday afternoon beers.

I would like to thank Prof. Christoph Borchers for having me in his lab at the the UVic Proteomics Centre together with Angela, Andrew P, Darryl and Melissa. And of course I'd like to thank my new friends, the Vancouver Island weekend crew, Serena, Adrian, Nicole, Andrew C and Karl for the hiking, biking, boating and site-seeing, I truly had an amazing time at Vancouver Island thanks to you. Serena, you should consider becoming a tour guide!

Without collaborations my thesis would be pretty empty. Therefore I would like to thank Prof. Jan Hoeijmakers' group at the Erasmus MC and Prof. Casper Hoogenraad from the Utrecht University for their expertise on aging and mouse models and help in neurobiology, respectively. Wilbert, many thanks for your immunohistochemistry teaching and your strong scientific involvement in the mouse cerebellum project. I would also like to thank my collaborators at the NKI; Prof. Daniel Peeper for scientific discussions and Joanna for the culturing of the cells and her enthusiasm for and involvement in the senescence proteomics project.

My awesome office Z608! Becky my big sister, my mentor, you spent a lot of time during my years here and I had a great time with you and your English humour. I learnt many nuances of British language and how to get excited about cricket (although I'm not sure if I got all the cricket rules right…). Shahram, always helping me with my IT problems and keeping me informed about the internet bargains of the day. Soenita, my surrogate mother, you always kept me going with your delicious food during the lunch club or from the numerous weekend food/cake left-overs. Your warm personality always made me feel at home in the office. Harm, in the beginning I was a bit scared of your angry face behind the computer in front of me, but soon enough I found out that this is how you look when you are concentrating and you're actually a very nice guy. You are a very skilled and dedicated

person and I enjoyed having a dutch sparring partner, exchanging quite some Southern and Eastern Dutch vocabulary as well. Fanster! I'm glad you joined our office as well, you are such a bright happy little bouncing ball, making my day happy as well.

The competing office next door, Z609. Arjan I enjoyed our completely random, meaningless but very funny conversations. However, the times when you were singing/shouting "My Heart Will Go On" I enjoyed less. And Glen, you amazed me with your Canadian tolerance to cold and alcohol. Christian, what a guy, my slow-motion competitor, I enjoyed your smart-ass sarcastic humor, you were a great companion during our stay in Vancouver and Bremen.

Which brings me to my Chinese friend, master of Ti4+-IMAC, Houijang. You are such a great guy and I had lots of fun with you dining in Chinese restaurants and during our stay in Bremen and later in Cambridge as well.

The wolf-pack, Benjamin and Pepijn. I thought nightlife had come to an end when I was not a student anymore. Nope, these two party animals dragged me to a lot of parties and were my social companions for the first years of my PhD. I had loads of fun with you guys, especially during the numerous days we played with our football team and ended up celebrating or remorsing our results with a cold beer in the sports bar and a few hours later in town, going out.

Which brings me to the NPC Football team. I really enjoyed playing in our own "What the Heck" indoor football team. Shabaz, Javier, Hendrik, Harm, Benjamin, Pepijn, Gianluca (when he wasn't in Sweden), Rob, Christian, Marco B, Marco H, Mao, Vikash, Michiel we really got better every match, although that didn't necessarily translate into better scores… Of course without the support of the real NPC team this was not possible, therefore many thanks to Martje and Werner.

Then I would like to thank the people with whom I shared many nice lab days out, Friday afternoon beers and Sinterkerst borrels; Fabio, (Sh)Alba, Ana, Javier, Reinout, Christian Sr., Nadia, Nadine, Natalie, Onno, Jeffrey, Paul, Adja, Vincent, Ayse, Andreas, Arjen, Patrick, Renske, Lucrece, Mirjam, Bas, Salvo, Henk (also for his scripting), Mao, Ioana, Basak, Joost, Esther, Jess, Charlotte, Sarah P, Thin Thin, Monique, Renske, Hongtao, Qingyang, Teck, Nicolas, Violette, Matina, Liana, Andrej, Andrea, Clement, Thierry, Philip, Anja and the Italian table football and Uno sympathisers Sara, Piero and Eleonora.

Also, I would like to thank my students, Lisette and Laura, that helped me with my projects. And all the visiting students/researchers that brought extra colour to the lab.

Special thanks to my saviour Geert who recovered 99% of my files after my harddrive failed. It was a big relief after nearly finishing my thesis with little to none backup.

Of course many thanks go to Corine for always helping me with all my questions, paperwork and PhD thesis planning.

Papa en mamma bedankt voor de vrijheid en openheid die jullie me hebben gegeven in

het maken van mijn keuzes. De vele weekendjes terug in het ouderlijk huis waren als een soort mini-vakanties waar we totaal tot rust konden komen. Deze vakanties werden vaak prettig onderbroken door bezoekjes van mijn favoriete neefje Lucas en mijn broer Paul en zijn vrouw Antoinette, en sinds kort ook Tim. Wat me ook altijd veel goed deed waren de feestjes met mijn vrienden uit het altijd gezellige Brabant, bedankt Remco, Wouter, Jeroen, Julian, Stefan, Bas (ook voor hulp met het cover design), Jacques, Peter, Twan, Robert, Jan, Tom, Merel, Susan, Magdelt, Lonneke, Inge en Erin.

En als laatste maar zeker niet als minste: Richelle, bedankt dat je altijd met liefde goed voor me hebt gezorgd, dat je open stond voor nieuwe avonturen en dat we daardoor samen veel leuke dingen hebben meegemaakt.

Thank you all!

This last 4 exciting years went by very quick and formed an important part of my life. Now it's time for a new oppurtunity!

Appendix:
Supplementary Information

Supplementary Figures to Chapter 2

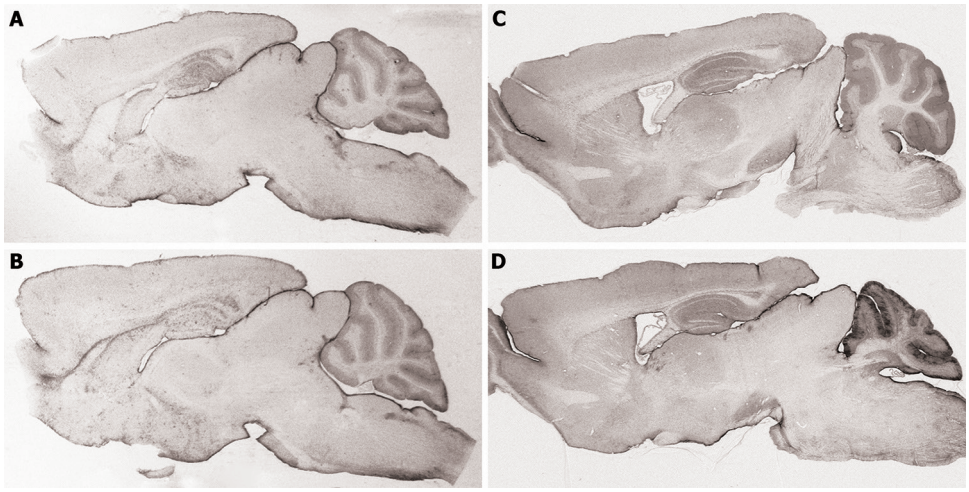

***Figure 1**: Mice model brain morphology comparison. GFAP sagittal IHC stainings of 8 weeks old control (A) & knock-out (B) and 26 weeks old control (C) & knock-out (D) mice.*

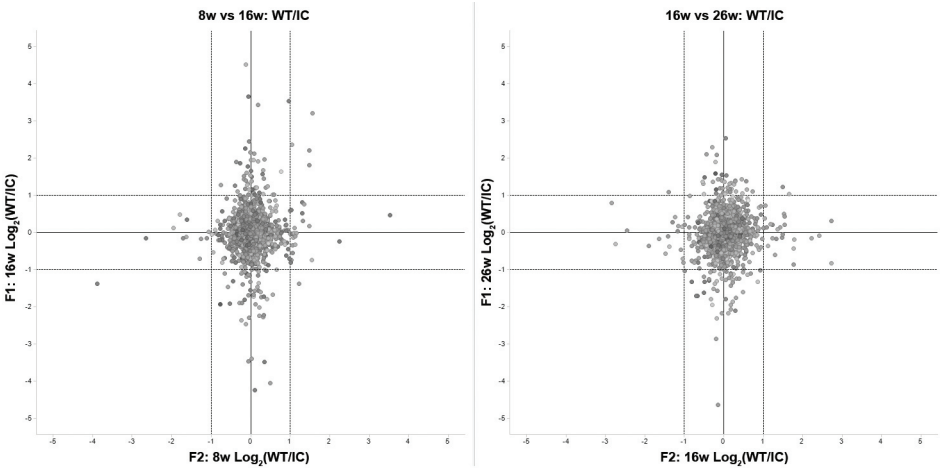

***Figure 2**: Control cerebella comparison over time. Left panel: Protein ratios of 8 week vs. 16 week old control mice. Right panel: Protein ratios of 16 week vs. 26 week old control mice.*

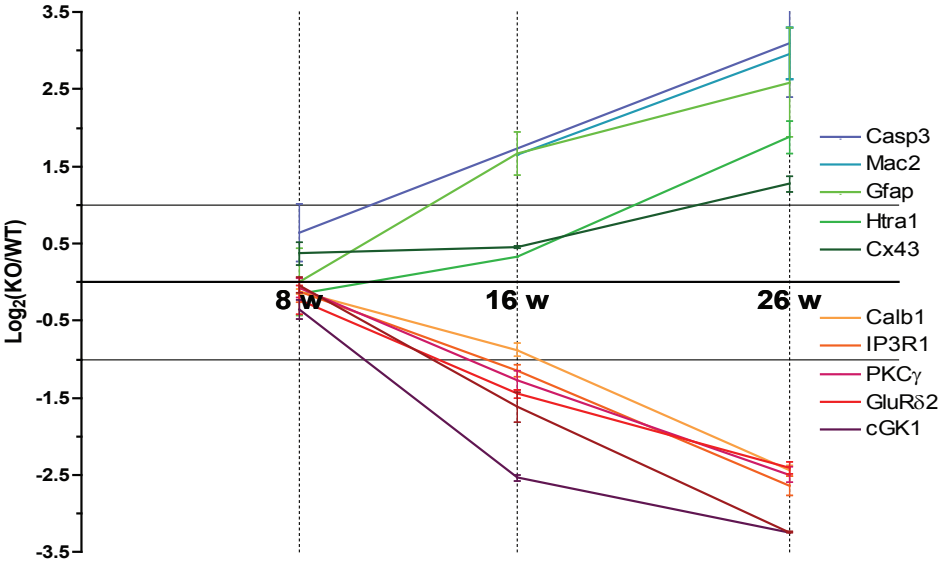

***Figure 3**: KO/Control ratio's over time for proteins selected for IHC. Lines are connected by the average value from the replicates measured by the proteomics screening. Whiskers show the highest and lowest values observed. Caspase 3 was only quantified at 8 and 26 weeks old mice.*

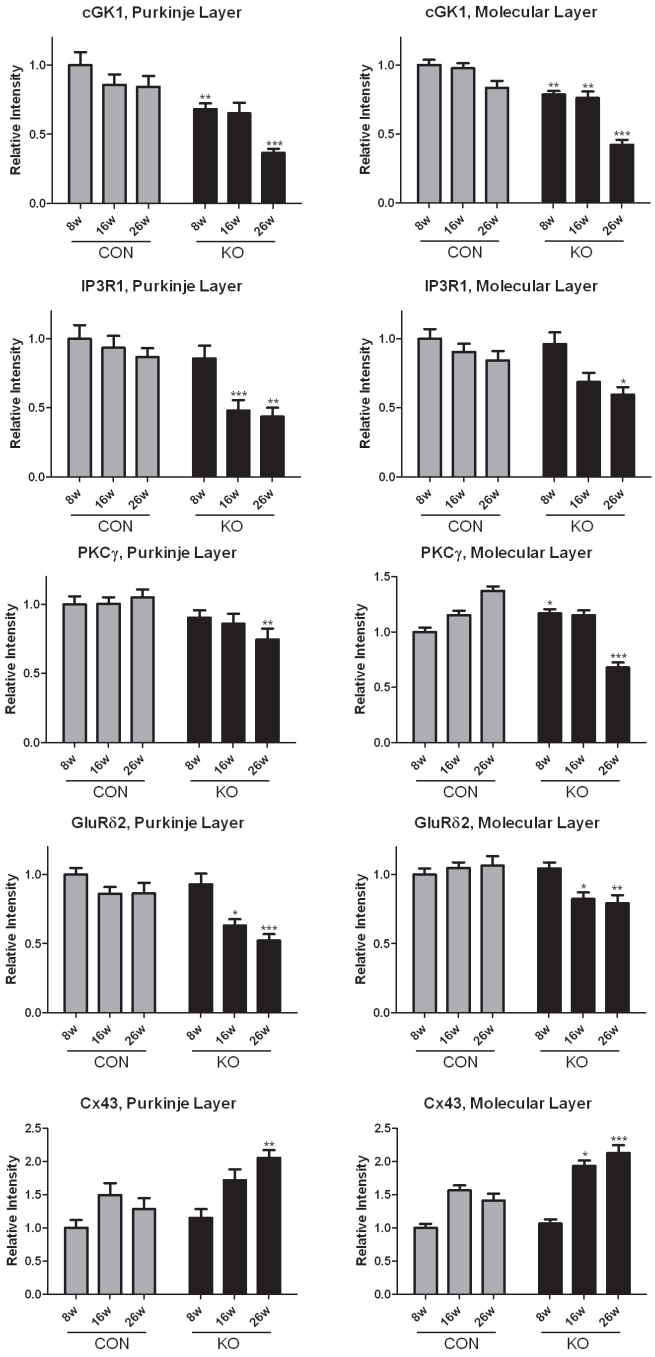

***Figure 4**: Semi-quantitative plots of protein abundance determined by immunohistochemistry. Staining in the molecular and Purkinje layer were quantified for cGK1, IP3R1, PKCγ, GluRδ2 and Cx43. A statistical comparison was made per time point between CON and KO. Significant differences are indicated by single ($p<0.05$), double ($p<0.01$) and triple ($p<0.001$) asterisks.*

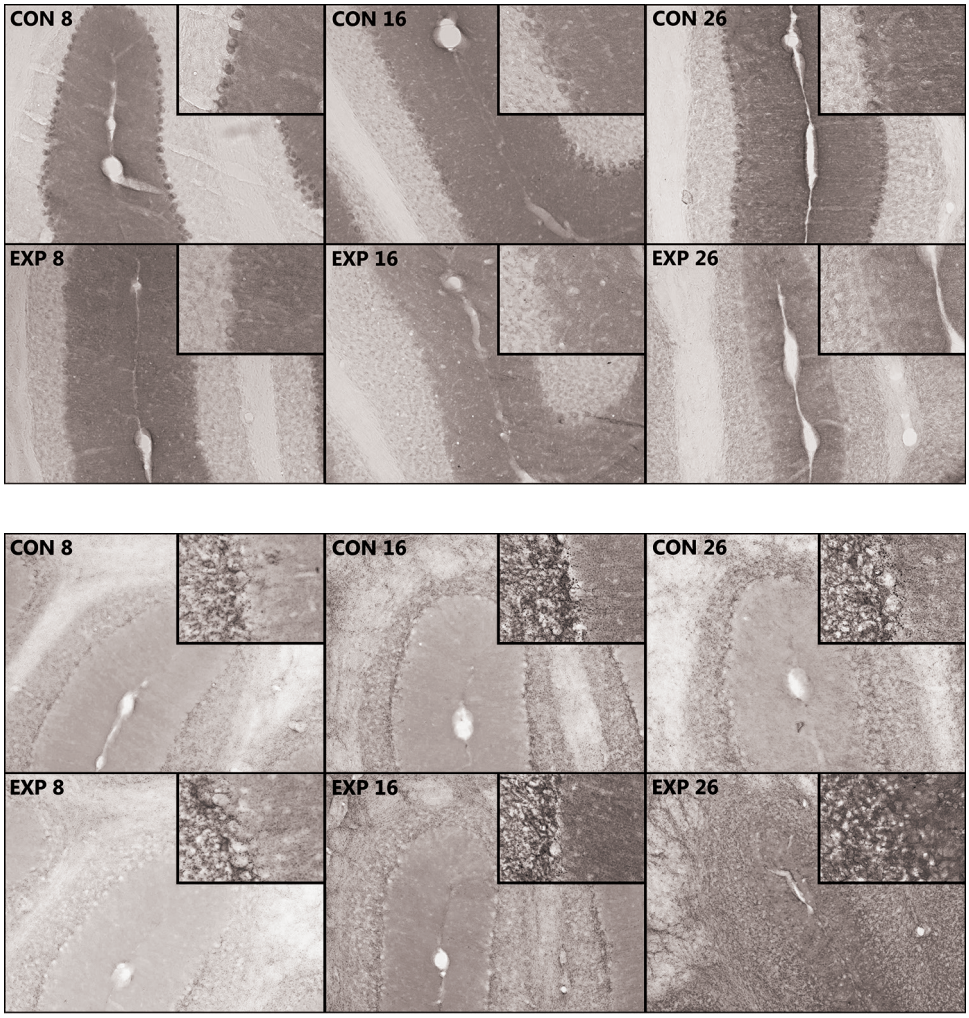

***Figure 5**: Immunohistochemistry photographs of GluRδ2 and Cx43 (Gja1) probed cerebella. Top panel: Control tissues showed a strong localisation of GluRδ2 to the purkinje cell body and molecular layer. Purkinje Ercc1 KO mice show a clear reduction and removal of GluRδ2 protein from the molecular layer and purkinje cell bodies over time. Bottom panel: Control tissues show Cx43 surrounding purkinje cell bodies and its presence in the granular and molecular layer. Purkinje KO mice show a strong increase of Cx43 in the molecular layer as well as in the granular layer.*

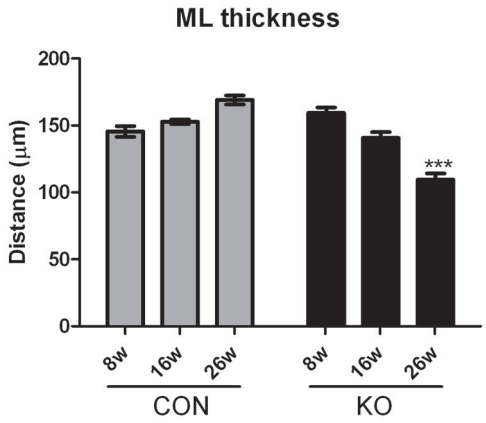

***Figure 6**: Molecular layer thickness measured for all conditions. An increasing trend in thickness is observed in control tissue whereas a strong decrease in purkinje KO is observed. Error bars represent SEM values. A statistical significant difference in ML thickness between 26 week old control and KO mice cerebella is indicated by three asterisks representing a p-value <0.001.*

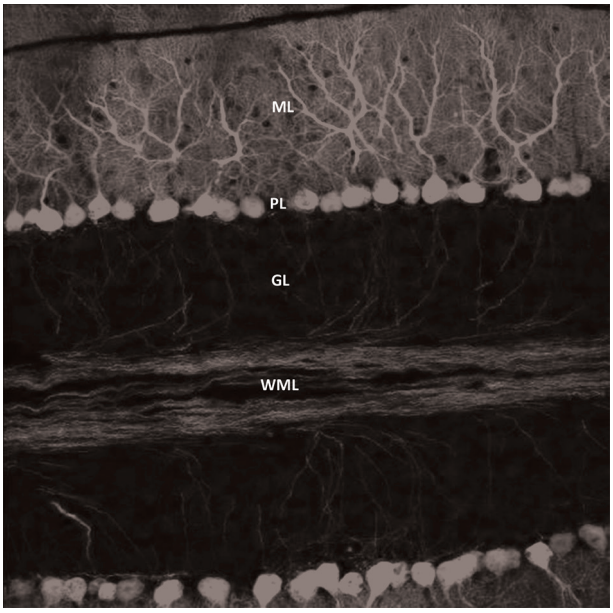

***Figure 7**: Confocal micrograph of cerebellum from transgenic mouse expressing EGFP driven by the L7/pcp2 promotor. Purkinje cell marker indicates the areas of Purkinje cell dendrites in the molecular layer (ML), Purkinje cell bodies in the Purkinje layer (PL) and Purkinje axons in the white matter layer (WML). (adopted from The Gene Expression Nervous System Atlas (GENSAT) Project, NINDS Contracts N01NS02331 & HHSN271200723701C to The Rockefeller University (New York, NY)).*

Supplementary Figures to Chapter 4

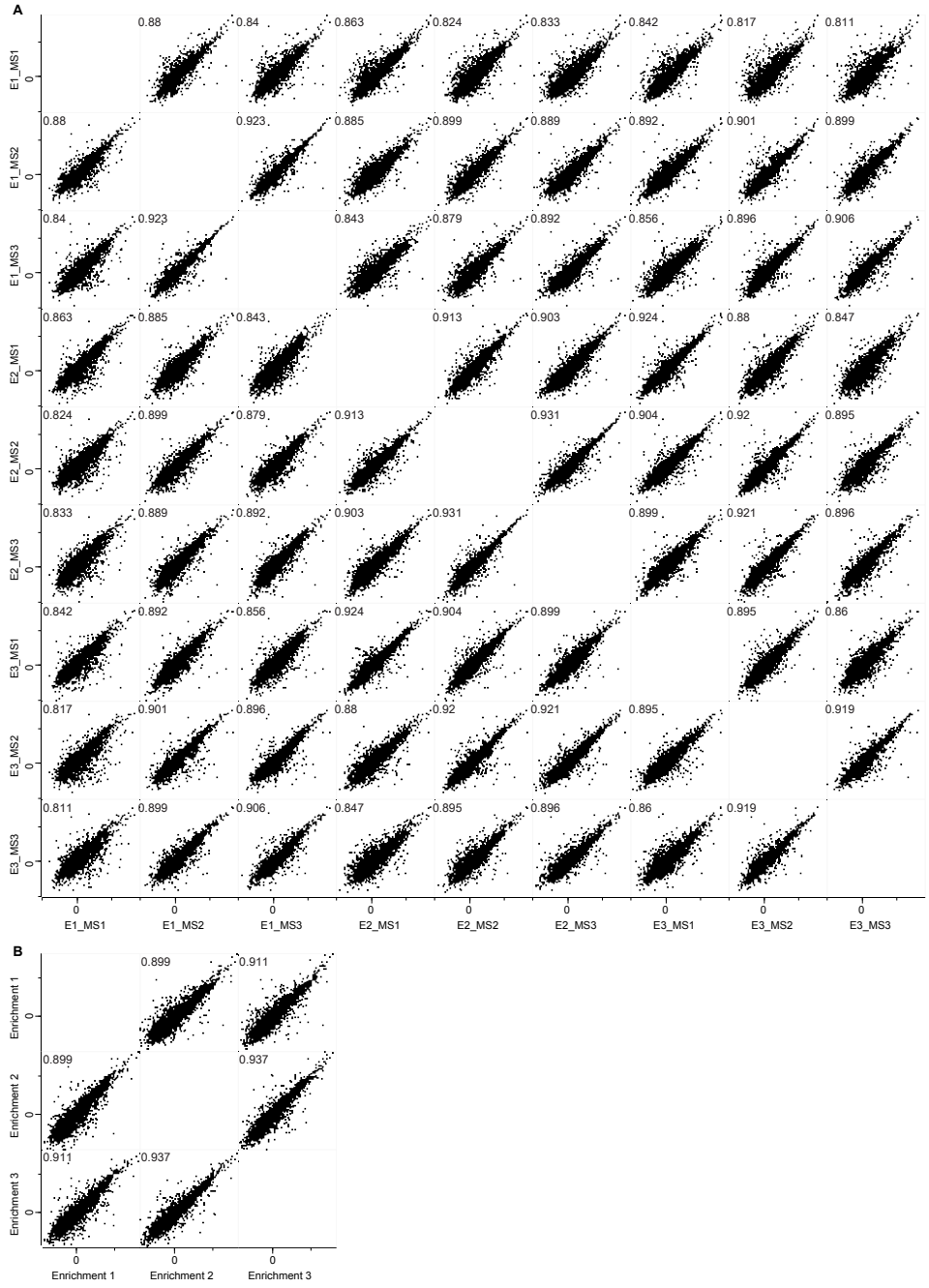

***Figure 1**: A) Intensity correlations for all phosphopeptide quantifications from HeLa. The phosphopeptide log base 2 intensities are plotted for all HeLa enrichments (E1,E2,E3) and their MS replicate analysis (MS1,MS2,MS3). B) Intensity correlation comparisons between each enrichment (average of three MS replicates). The average Pearson correlation was 0.90 and 0.92 for the MS replicates and enrichment replicates, respectively.*

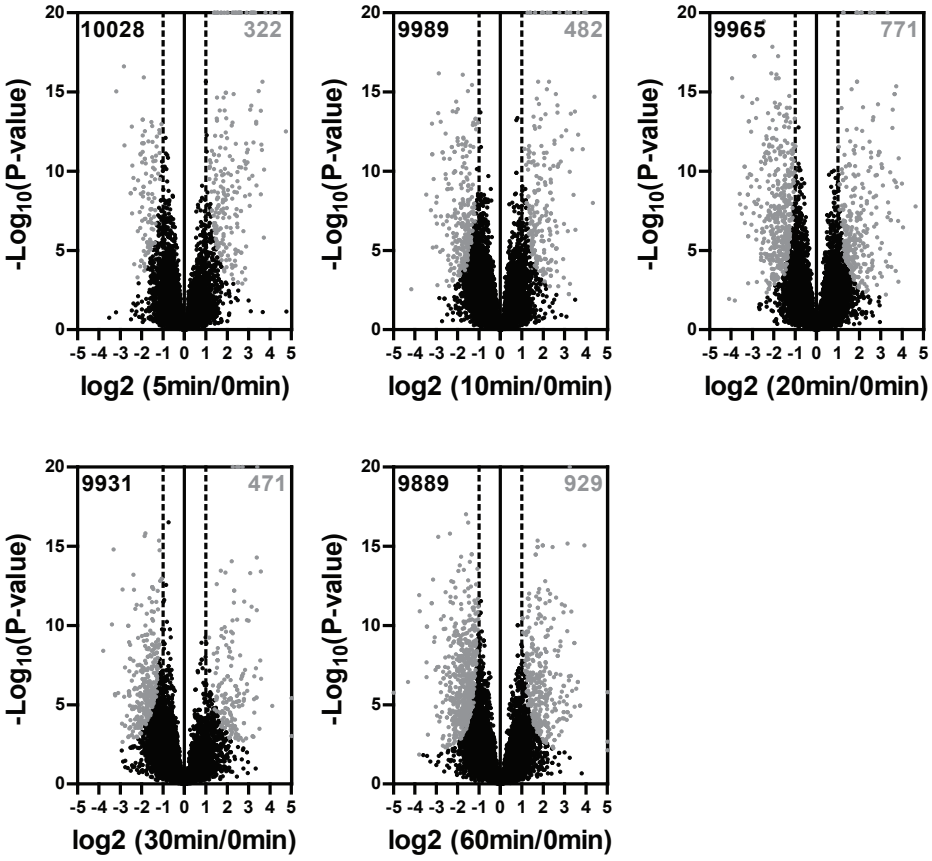

***Figure 2**: Phosphosite ratio versus significance plots of all PGE_2 stimulations versus control comparisons (two sample t-test). P-values (-log base 10) are plotted as a function of the phosphosite ratio (log base 2) for all the PGE_2 stimulations (5, 10, 20, 30, 60 min) versus control (0 min) comparisons. Regulated sites are colored in gray (permutation based FDR=0.005, s0 adjusted for 2-fold regulation).*

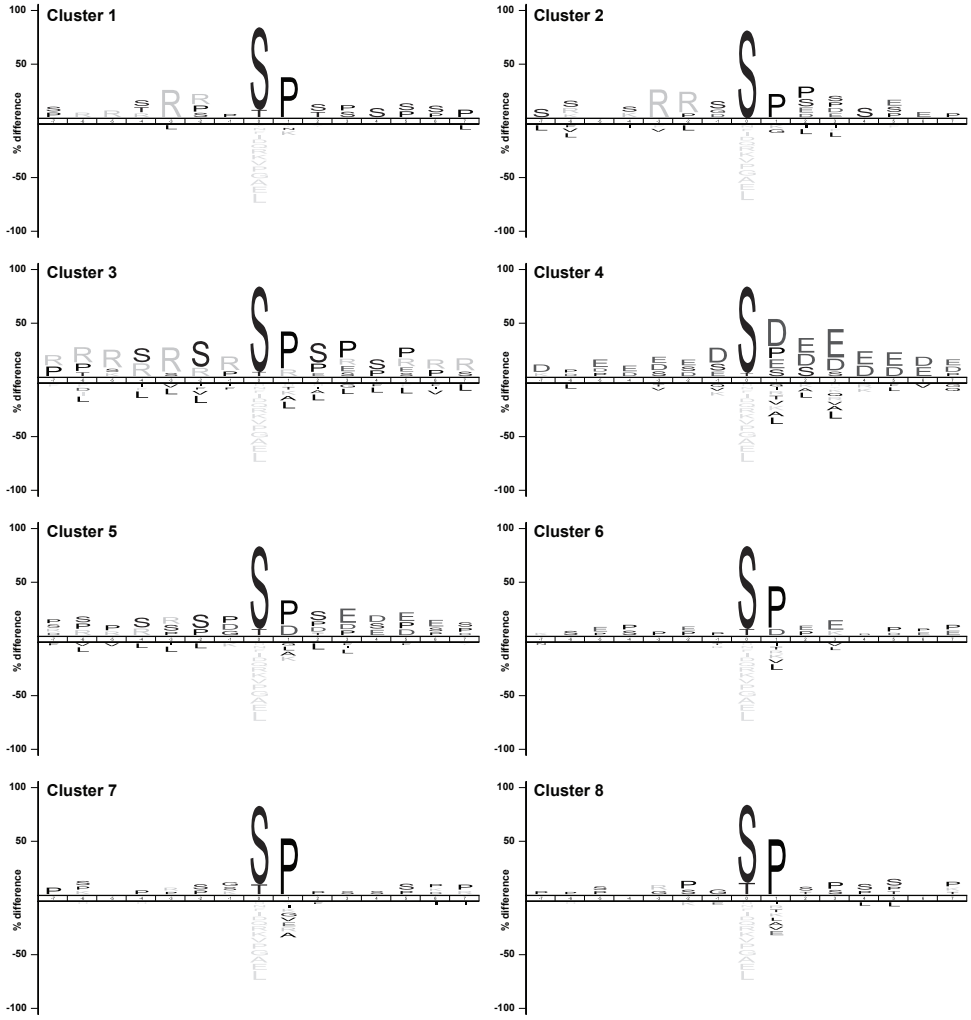

***Figure 3**: Sequence motif IceLogos generated from each cluster showing significantly (P < 0.01) enriched or depleted amino acids (relative amino acid frequencies) flanking the regulated phosphosites when compared to reference background (Swiss-Prot human database).*

Supplementary Figures to Chapter 5

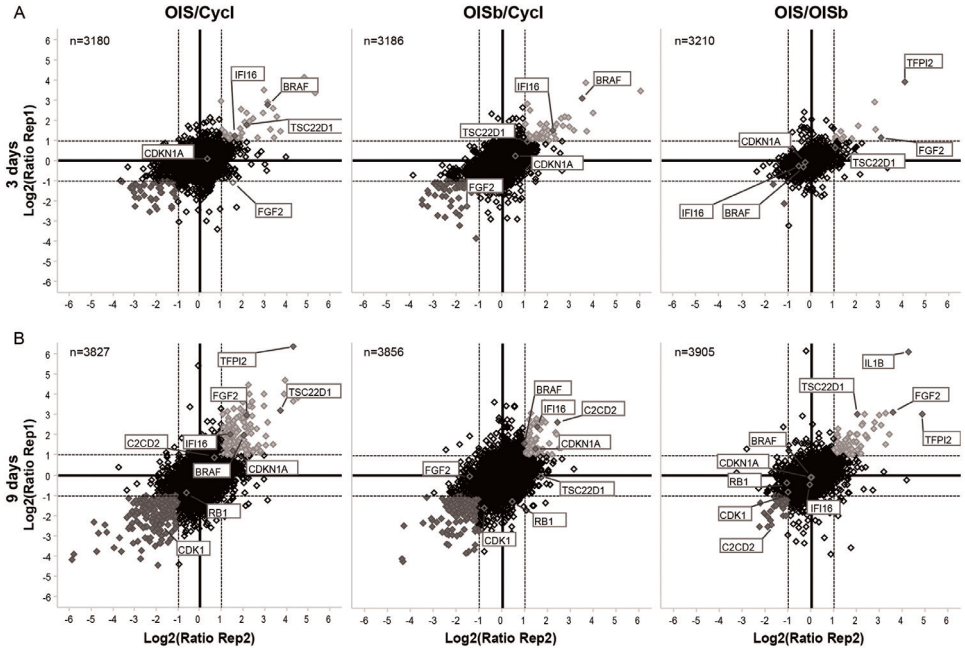

***Figure 1**: Protein expression comparisons for cycling (Cycl), BRAF transfected (OIS) and BRAF/sh-CEBPb transfected (OISb) human diploid fibroblasts after 3 days (A) and 9 days (B) of cell growth. The logarithmic base 2 ratio for two conditions of replicate 1 was plotted against the logarithmic base 2 ratio for the same two conditions of replicate 2 to visualize consistent changes upon replicates. A two-fold cut-off was used to determine up- (light gray) and down- (in dark gray) regulated proteins. Highlighted proteins are displayed in dark green.*

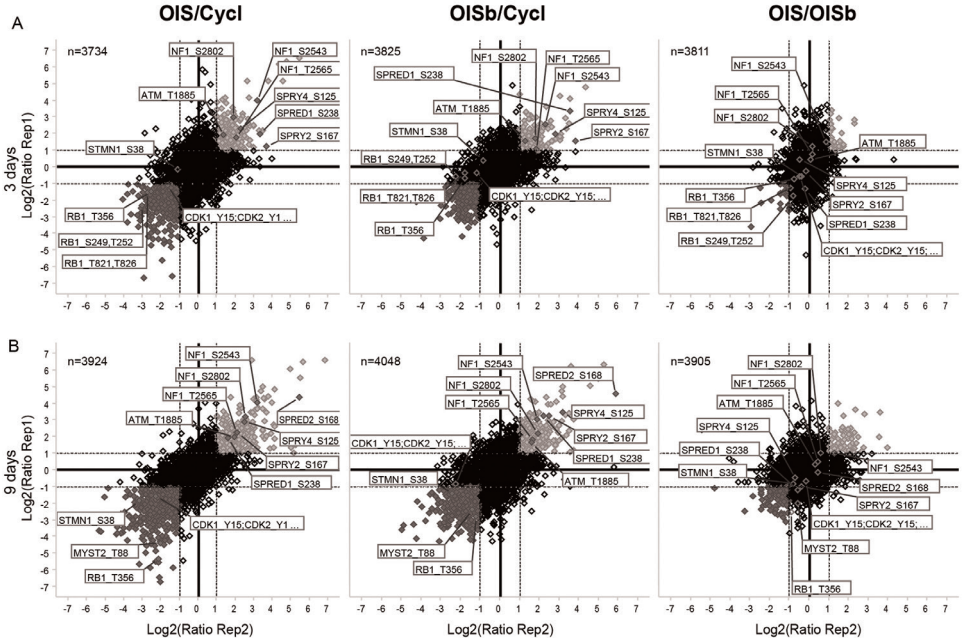

***Figure 2**: Phosphosites regulated specifically in BRAF transfected cells (OIS and OISb). Phosphosite abundance comparisons for cycling (Cycl), BRAF transfected (OIS) and BRAF/sh-CEBPb transfected (OISb) human diploid fibroblasts after 3 days (A) and 9 days (B) of cell growth. The logarithmic base 2 ratio for two conditions of replicate 1 was plotted against the logarithmic base 2 ratio for the same two conditions of replicate 2 to visualize consistent changes upon replicates. A two-fold cut-off was used to determine up- (in light gray) and down- (in dark gray) regulated proteins. Phosphosites regulated in common for OIS and OISb are displayed in dark green.*

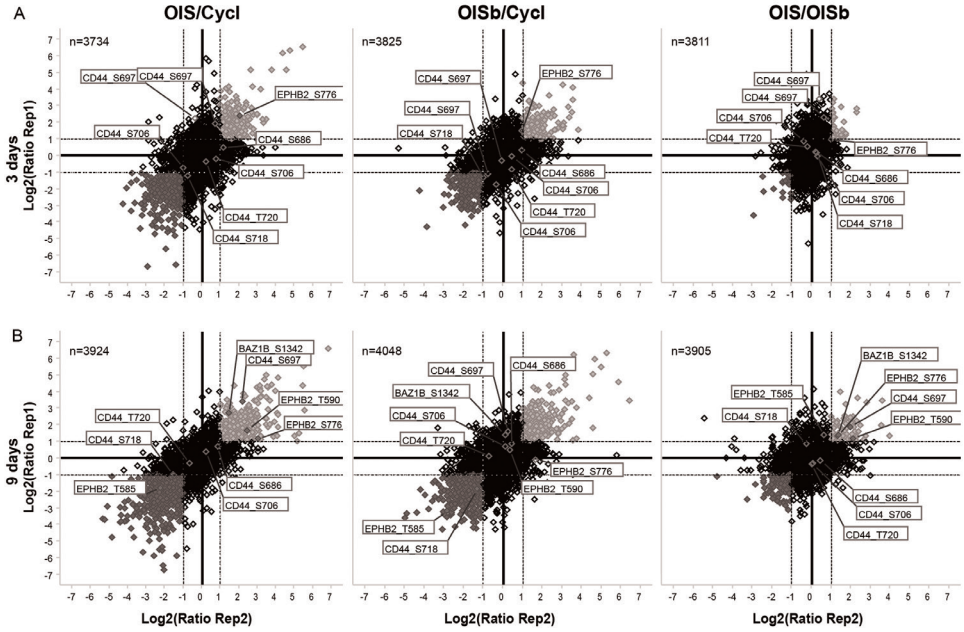

***Figure 3**: Phosphosites regulated specifically in OIS. Phosphosite abundance comparisons for cycling (Cycl), BRAF transfected (OIS) and BRAF/sh-CEBPb transfected (OISb) human diploid fibroblasts after 3 days (A) and 9 days (B) of cell growth. The logarithmic base 2 ratio for two conditions of replicate 1 was plotted against the logarithmic base 2 ratio for the same two conditions of replicate 2 to visualize consistent changes upon replicates. A two-fold cut-off was used to determine up- (in light gray) and down- (in dark gray) regulated proteins. A subset of phosphosites regulated specifically in OIS are displayed in dark green.*