Development and Application of Quantitative Proteomic Strategies to Study the Cell Cycle in Fission Yeast

Dissertation

der Mathematisch-Naturwissenschaftlichen Fakultät

der Eberhard Karls Universität Tübingen

zur Erlangung des Grades eines

Doktors der Naturwissenschaften

(Dr. rer. nat.)

vorgelegt von

Alejandro Carpy Gutierrez Cirlos

aus Naucalpan/Mexiko

Tübingen

2014

Tag der mündlichen Qualifikation: 29.07.2014
Dekan: Prof. Dr. Wolfgang Rosenstiel
1. Berichterstatter: Prof. Dr. Boris Macek
2. Berichterstatter: Prof. Dr. Ralf-Peter Jansen

Table of contents

List of abbreviations

AQUA	Absolute Quantification
ATP	Adenosine Triphosphate
C18	Octadecyl Carbon Chain
CDK	Cyclin Dependent Protein Kinase
CID	Collision Induced Dissociation
DDA	Data-Dependent Analysis
DHB	2,5-Dihydroxybenzoic Acid
eFT	Enhanced Fourier Transform
ESI	Electro Spray Ionization
FDR	False Discovery Rate
FT	Fourier Transform
FTICR	Fourier Transform Ion Cyclotron
HCD	High Energy Collision Dissociation
HILIC	Hydrophilic-Interaction Chromatography
HPLC	High Pressure Liquid Chromatography
HR/AM	High Resolution – Accurate Mass
iBAQ	Intensity Based Absolute Quantification
ICAT	Isotope-Coded Affinity Tags
IMAC	Metal Ion Affinity Chromatography
iTRAQ	Isobaric Tags for Relative and Absolute Quantification
KDAC	Lysine deacetylases
KTAC	Lysine acetyltransferases
LC	Liquid Chromatography
LIT	Linear Ion Trap
LTQ	Linear Trap Quadrupole
MALDI	Matrix Assisted Laser Desorption Ionization
MS	Mass Spectrometry
m/z	Mass-to-Charge Ratio
ORF	Open Reading Frame
PrEST	Protein Epitope Signature Tags
PSAQ	Protein Standard Absolute Quantification
PTMs	Posttranslational Modifications
SAX	Strong Anion Exchange Chromatography
SCX	Strong Cation Exchange Chromatography
SILAC	Stable Isotope-Labeling by Amino Acids in Cell Culture

TiO$_2$	Titanium Dioxide Chromatography
TFA	Trifluoroacetic acid
TOF	Time-Of-Flight
TSQ	Triple Stage Quadrupole

Summary

With the increasing amount of genomic data available, it became obvious that complete genome sequences are not sufficient to elucidate a clear biological function. Proteomics focuses on the gene products (proteins) and their posttranslational modifications (PTMs) in an unbiased and global fashion. Proteomics has become an indispensable tool in the study of molecular and cellular biology, and is becoming increasingly important for systems biology approaches. It highly profits from the development and improvements in technology, especially in mass spectrometry and liquid chromatography. Modern mass spectrometry-based approaches are capable of routinely identifying and quantifying thousands of proteins and post-translational modifications in a single biological experiment.

The aim of this thesis was to develop and optimize quantitative mass spectrometry-based approaches to study protein phosphorylation dynamics, and apply them to study the cell cycle of fission yeast. To this end, I improved a (phospho)proteomic workflow based on in-solution proteome digestion, phosphopeptide enrichment and the use of strong cation exchange and TiO_2 chromatographies. I replaced the commonly used additive 2,5-dihydroxybenzoic acid (DHB) with high concentration of trifluoroacetic acid (TFA), thereby increasing the specificity of the phosphopeptide enrichment while reducing contamination of the mass spectrometer. With this approach it became possible to identify thousands of phosphorylation events with relatively low starting material (1 - 2 mg).

Proteomics can be applied to study complex biological processes such as the cell cycle. The cell cycle involves a complex series of highly regulated and evolutionary conserved events. Aberrations in the cell cycle have severe implications and can cause cancerous growth; therefore, a detailed understanding of the cell cycle and its regulation may identify additional targets for cancer therapy. In this work, I used the model organism *Schizosaccharomyces pombe*, also known as fission yeast, to characterize proteome and phosphoproteome dynamics in the main phases of the cell cycle. *S. pombe* is an extensively exploited model organism used to study cell cycle control, heterochromatin and differentiation. Here I used shot-gun proteomics in combination with a stable isotope labeling by amino acids in cell culture (SILAC) method, termed super-SILAC, and intensity-based absolute quantification (iBAQ). Combining these methods I was able to measure relative and absolute dynamics of the proteome and phosphoproteome during the four main phases of the cell cycle of fission yeast (G1, S, G2 and M). Using highly synchronized cells provided by Silke Hauf´s laboratory, I estimated copy numbers for 3,178 *S. pombe* proteins and 3,682 phosphorylation events and relatively quantified 65 % of the data across all phases. This data was combined with calculated phosphorylation site stoichiometry to estimate the total amount of protein-bound

phosphate and its dynamics across the cell cycle. Quantitative changes during the cell cycle were infrequent and weak at the proteome level, but prominent at the phosphoproteome level. Protein phosphorylation peaked in mitosis, where the median phosphorylation site occupancy was 44%, about two-fold higher than in other phases. Using the iBAQ approach, I measured copy numbers of 3,178 proteins and combined this information with phosphorylation site stoichiometry to estimate the absolute amount of protein-bound phosphate, as well as its change across the cell cycle. The results indicated that 23% of the average intracellular ATP is utilized by protein kinases to phosphorylate their substrates in order to drive regulatory processes during cell division. Accordingly, I observed that phosphate transporters and phosphate-metabolizing enzymes are phosphorylated and therefore likely to be regulated in mitosis. This is one of the first studies to describe global phosphorylation site stoichiometry and the first study to combine this information with absolute protein copy number to calculate ATP dynamics during the cell cycle. This will be a valuable resource for systems biologists and we hope it will encourage researchers to expand this approach to other modifications to help better understand PTM dynamics.

Finally, in collaboration with Ian Hagan´s lab we observed that labeling fission yeast with a frequently used version of arginine ($^{13}C_6{}^{15}N_4$-arginine) lead to severe mislabeling of multiple amino acids and severely compromises the identification of heavy-labeled peptides and therefore of proteins. This inability to use $^{13}C_6{}^{15}N_4$-arginine in triple-SILAC experiments limits the exploitation of SILAC technology in fission yeast. Through experiments using different combinations of SILAC labels we reasoned that the guanidinium group of $^{13}C_6{}^{15}N_4$-arginine is catabolized to $^{15}N_1$-ammonia that is used as a precursor for general amino acid biosynthesis. As a consequence, ^{15}N is randomly incorporated in newly synthetized amino acids and at a later point in newly synthetized proteins. In order to prevent this arginine conversion, we optimized the fission yeast strain previously used for SILAC experiments and growth conditions. We show that disruption of Ni^{2+}-dependent urease activity, through deletion of the sole Ni^{2+} transporter Nic1 in combination with ammonium-supplemented medium, blocks this re-cycling to enable $^{13}C_6{}^{15}N_4$-arginine labeling for SILAC strategies in *S. pombe.* The ability to routinely perform triple-SILAC experiments with doubly- labelled samples (i.e. Arg and Lys) will increase proteome coverage and increase experimental options to understand core principles of molecular cell biology. Finally, our study will point the way towards solving of arginine-related labeling problems in SILAC experiments in other species.

Zusammenfassung

Mit der zunehmenden Verfügbarkeit von Genomdaten wurde deutlich, dass die vollständigen Genomsequenzen nicht ausreichend sind, um eine eindeutige biologische Funktion aufzuklären. Proteomik ist eine unvoreingenommene und umfassende Methode, die sich mit Genprodukten (Proteine) und deren posttranslationalen Modifikationen (PTM) beschäftigt. Sie hat sich als ein unverzichtbares Werkzeug in molekularen und zellbiologischen Studien etabliert und wird immer wichtiger für die Systembiologie. Proteomik profitiert stark vom technologischen Fortschritt, vor allem in der Massenspektrometrie und der Flüssigkeitschromatographie. Mit Hilfe moderner, auf der Massenspektrometrie basierender Versuche, ist man in der Lage, routinemäßig tausende von Proteinen und posttranslationale Modifikationen in einem biologischen Experiment zu identifizieren und zu quantifizieren.

Ziel dieser Doktorarbeit war die Entwicklung und Optimierung quantitativer Massenspektrometrie-basierter Ansätze zur Untersuchung der Phosphorylierungsdynamik von Proteinen und diese anschließend zur Analyse des Zellzyklus der Spalthefe anzuwenden. Zu diesem Zweck verbesserte ich einen Phosphoproteom-Workflow, basierend auf einem Proteinverdau in Lösung, Phosphopeptidanreicherung und der Verwendung von Ionenaustausch- und TiO_2-Chromatographien. Ich ersetzte die üblicherweise verwendete 2,5-Dihydroxybenzoesäure (DHB) gegen eine hohe Konzentration an Trifluoressigsäure (TFA), wodurch die Spezifität der Phosphopeptidanreicherung erhöht und gleichzeitig die Kontamination des Massenspektrometers reduziert wurde. Mit diesem Ansatz ist es möglich, tausende von Phosphorylierungsstellen von Proteinen mit relativ niedrigem Ausgangsmaterial zu identifizieren (1 - 2 mg).

Proteomik kann angewendet werden, um komplexe biologische Prozesse wie den Zellzyklus zu untersuchen. Der Zellzyklus umfasst eine komplexe Reihe von hochregulierten und evolutionär konservierten Vorgängen. Aberrationen im Zellzyklus haben schwere Auswirkungen und können Krebswachstum verursachen. Daher kann ein genaues Verständnis des Zellzyklus und seiner Regulation zur Identifizierung weiterer Targets in der Krebstherapie führen. In dieser Arbeit habe ich den Modellorganismus *Schizosaccharomyces pombe* benutzt, der auch als Spalthefe bekannt ist, um die Dynamik des Proteoms und des Phosphoproteoms in den wichtigsten Phasen des Zellzyklus zu charakterisieren. *S. pombe* ist ein etablierter Modellorganismus, der verwendet wird, um unter anderem Zellzykluskontrolle, Heterochromatin und Differenzierung zu studieren. Hier habe ich Massenspektrometrie-basierte Proteomikmethoden in Kombination mit stabiler Isotopenmarkierung von Aminosäuren in Zellkultur (SILAC), genannt Super-SILAC, und intensitätsbasierter absoluter Quantifizierung (iBAQ) verwendet. Durch Kombination dieser Methoden konnte ich die relative

und absolute Dynamik des Proteoms und Phosphoproteoms während der vier Hauptphasen des Zellzyklus der Spalthefe (G1, S, G2 und M) charakterisieren. An den von Silke Hauf´s Labor zur Verfügung gestellten hoch synchronisierten Zellen berechnete ich die Kopiezahlen für 3.178 *S. pombe* Proteine und 3.682 Phosphorylierungsereignisse und 65% der Daten konnten über alle Phasen relativ quantifiziert werden. Diese Daten wurden mit der Stöchiometrieberechnung von Phosphorylierungsstellen kombiniert, um die Gesamtmenge an proteingebundenem Phosphat und seine Dynamik über den Zellzyklus zu bestimmen. Quantitative Veränderungen während des Zellzyklus waren auf Proteomebene selten und schwach, aber deutlich auf Phosphoproteomebene. Der Level an Proteinphosphorylierung erreichte in der Mitose seinen Höhepunkt, wo der Median der besetzten Phosphorylierungsstellen bei 44% lag, etwa zweifach höher als in anderen Phasen. Mit dem iBAQ Ansatz maß ich Kopiezahlen von 3.178 Proteinen und kombinierte diese Informationen mit jenen aus der Stöchiometrieberechnung der Phosphorylierungsstellen, um die absolute Menge an proteingebundenem Phosphat, sowie dessen Veränderung über den Zellzyklus bestimmen zu können. Die Ergebnisse zeigten, dass 23% des durchschnittlichen intrazellulären ATPs durch Proteinkinasen genutzt wird, um ihre Substrate für regulatorische Prozesse während der Zellteilung zu phosphorylieren. In Übereinstimmung habe ich festgestellt, dass Phosphat-Transporter und phosphatmetabolisierende Enzyme phosphoryliert und daher wahrscheinlich in der Mitose reguliert werden. Dies ist eine der ersten Studien über die globale Stöchiometrie von Phosphorylierungsstellen und die erste Studie, die diese Informationen mit der absoluten Kopiezahl von Proteinen kombiniert, um die ATP-Dynamik während des Zellzyklus zu berechnen. Sie wird eine wertvolle Quelle für Systembiologen sein, und wir hoffen, dass sie Forscher dazu ermutigt, diesen Ansatz auf andere Modifikationen zu erweitern, um die PTM Dynamik besser zu verstehen.

Abschließend wurde in Zusammenarbeit mit Iain Hagan´s Labor beobachtet, dass es bei der Markierung der Spalthefe mit der häufig verwendeten Version von „schwerem" Arginin ($^{13}C_6$$^{15}N_4$-Arginin) zur schwerwiegenden Falschmarkierung zahlreicher Aminosäuren kommt, was die Identifizierung von markierten Peptiden und zugehörigen Proteinen erheblich beeinträchtigt. Die Tatsache, dass $^{13}C_6$$^{15}N_4$-Arginin nicht in Triple-SILAC Experimenten benutzt werden kann, stellt eine erhebliche Einschränkung der SILAC-Technologie in Experimenten in der Spalthefe dar. Durch Versuche mit verschiedenen SILAC-Aminosäuren fanden wir heraus, dass die Guanidinium-Gruppe von $^{13}C_6$$^{15}N_4$-Arginin in $^{15}N_1$-Ammoniak katabolisiert wird, welches weiter als Vorläufer für die allgemeine Aminosäure-Biosynthese dient. Als Folge wird ^{15}N zufällig in neu synthetisierte Aminosäuren und zu einem späteren Zeitpunkt in neu synthetisierte Proteine eingebaut. Um diese Arginin-Umwandlung zu verhindern, optimierten wir den zuvor für SILAC Experimente benutzten Stamm der Spalthefe. Ich konnte zeigen, dass

die Unterbrechung der Ni^{2+}-abhängigen Urease-Aktivität durch den Knockout des einzigen Ni^{2+}-Transporters Nic1 in Verbindung mit einem durch Ammonium ergänzten Medium, die Katabolisierung verhindert und somit eine $^{13}C_6^{15}N_4$-Arginin-Markierung für SILAC Experimente in *S. pombe* ermöglicht. Die Fähigkeit, Triple-SILAC Experimente routinemäßig mit zweifach-markierten Proben (d.h. Arg und Lys) einzusetzen, wird die Tiefe der Proteom Abdeckung erhöhen und erweitert somit die experimentellen Möglichkeiten, grundlegende Prinzipien der molekularen Zellbiologie besser zu verstehen. Die Ergebnisse dieser Studie können auf andere Spezies übertragen werden, um Arginin-bezogene Markierungsprobleme in SILAC Experimenten zu beseitigen.

1. Introduction

1.1   Mass-Spectrometry based proteomics

The central dogma of biology is defined as follows: "…the coded genetic information hard-wired into DNA is transcribed into individual transportable cassettes, composed of messenger RNA (mRNA); each mRNA cassette contains the program for synthesis of a particular protein" (1). Currently we realize that there are many exceptions to this rule and that many of the DNA regions that were supposed to be non-coding (introns) indeed encode various types of functional RNAs (2). With the introduction of shotgun sequencing (3, 4) and next-generation sequencing (5) large amounts of DNA sequences are being accumulated. With this data available, researchers started to notice that having complete genome sequences is not sufficient to elucidate biological function (6), therefore attention shifted towards complementary techniques like proteomics, the global study of proteins (7). With the introduction of the first integrative and comparative studies, it also became clear that the existence of an open reading frame (ORF) in genomic data does not necessarily imply the existence of a expressed gene. Furthermore, despite great efforts in bioinformatics, a complicated task remains to accurately predict genes from genomic data (8). Therefore, verification of a gene product by proteomic methods is an important first step in genome annotation. The link between genomics and proteomics is called transcriptomics which is the global profiling of gene expression at the mRNA level. Transcriptomics is a very useful tool to identify clusters of genes for which the expression is specific for a determined state. These methods are very sensitive but have drawbacks due to correlation inconsistencies with mRNA vs detected protein levels (8). Proteomics is complementary to both transcriptomics and genomics because it focuses on the gene products (proteins) that are the present in the cells and their respective posttranslational modifications (PTMs). Furthermore, modifications of the proteins that are not apparent from the DNA sequence, such as co- and post-translational modifications, can be determined only by proteomic methodologies.

The first approaches used in proteomics were based on two-dimensional (2D) gel electrophoresis. In this approach, proteins separated based on isoelectric point and molecular weight generate different patterns that could be used to study different perturbations in the cell or tissue of interest. The spots (proteins) forming the patterns can be identified using mass spectrometry (MS). This method, although widely used in proteomics, has severe disadvantages; it is wet-lab intensive and time consuming, it is not able to identify very small proteins or proteins with extreme pIs, and it has a limited analytical capacity, i.e. it is generally of low throughput in terms of protein identifications.

Relatively recent developments made in MS instrumentation, liquid chromatography (LC) and bioinformatics approaches resulted in the development of a MS-based technique known as "shotgun" or bottom-up proteomics. The goal of this approach is to identify proteins in a complex mixture using a combination of LC-MS. This widely used approach involves digestion of a whole cell lysate with a protease resulting in peptides that are separated by liquid chromatography before being measured with tandem mass spectrometry and analyzed using downstream bioinformatic workflows.
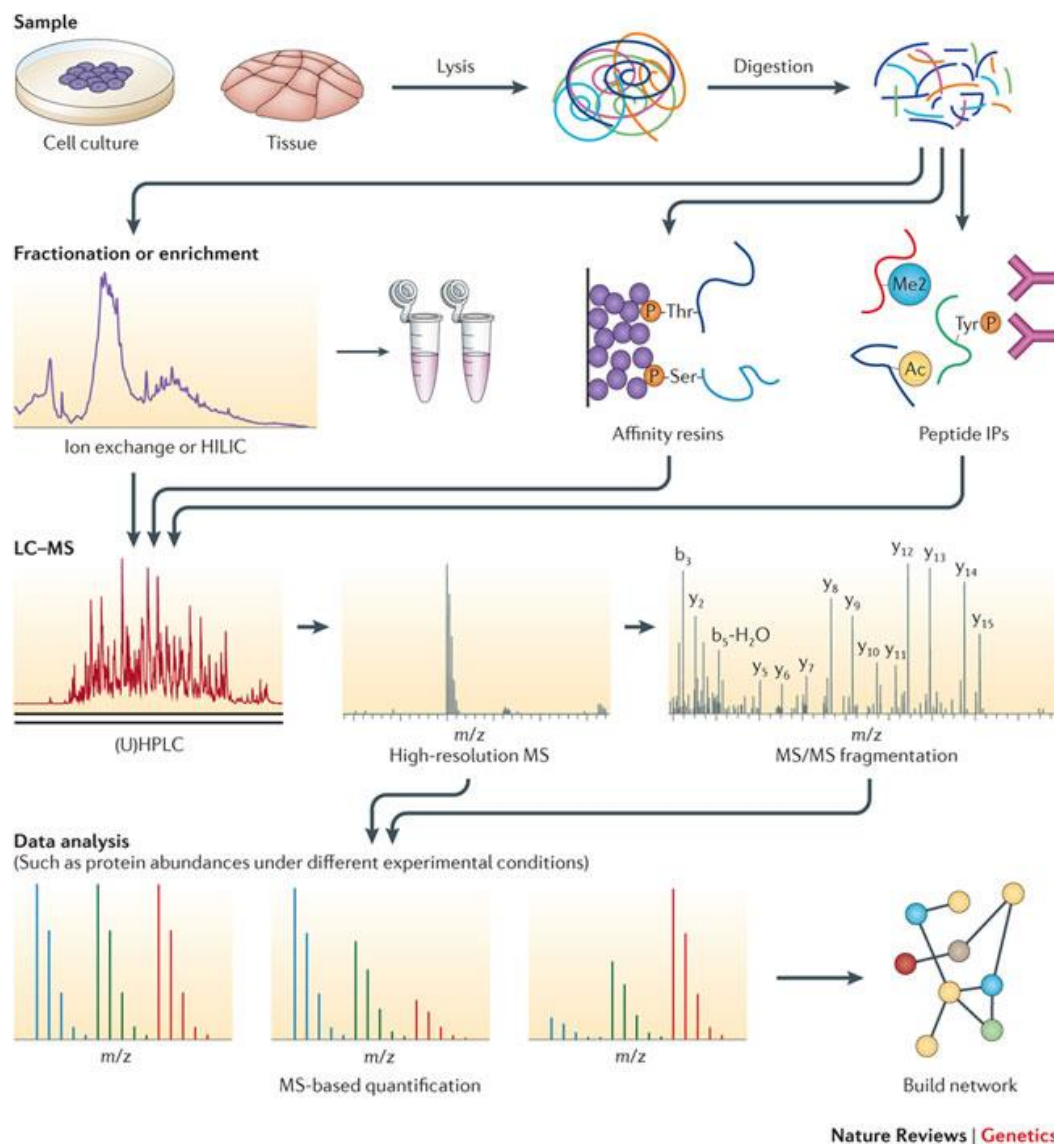
Figure 1. ***Mass Spectrometry based proteomics workflow.*** *In a typical MS experiment cells are lysed. Protein samples are digested using a protease. After digestion, peptides are fractionated or alternatively enriched for PTMs before being measured using LC-MS. The mass of a peptide is recorede in a full scan (MS) and isolated for fragmentation (MS/MS). The information from the MS and MS/MS is used to identify peptides and therefore proteins. The samples are quantified and the information used to build networks. Adapted from (9).*

Workflow

A typical MS-based proteomics experiment **(Figure 1)** consists of several steps. Initially (step 1), proteins are extracted from the cells (cell culture or tissue), generating very complex protein mixtures. Since the MS of whole proteins is less sensitive and does not provide enough information for identification, in step 2 proteins are digested with a protease. Trypsin is the most widely used protease, since it is highly efficient, specific and generates peptides with a suitable size for MS (better "flyability"). To reduce the complexity of this mixture, peptides are fractionated in step 3. Alternatively, for peptide PTM enrichment, affinity resins or modification specific antibodies can be used to target specific PTMs. In step 4 peptides are separated using high pressure liquid chromatography (HPLC or LC) coupled to a MS through electrospray ionization (ESI). In this step, peptides are usually separated using reversed phase chromatography, eluting peptides become ionized and enter the instrument as protonated peptides. In step 5, after entering the instrument a MS spectrum of peptides eluting at this point is recorded in a full scan (mass of the peptide). The most abundant ions (peptides) are then selected for fragmentation with tandem MS or MS/MS (step 6). The combination of precursor m/z (MS) and its fragment ions (MS/MS) is then matched to known peptide sequences from large protein databases using search algorithms in step 7. This results in peptides identified that are matched to proteins. Finally (step 8), data are quantified (relatively or absolutely), and several experimental conditions are compared directly. All this information is then gathered in step 9 and used to build networks of perturbations to the system.

1.2 Mass spectrometry

Proteomics relies heavily on modern analytical instrumentation. One of the most powerful tools for proteomic scientists are mass spectrometers. All mass spectrometers are based on three essential components: ion source, mass analyzer and detector **(Figure 2)**. Molecules from the sample first need to be ionized in a vacuum since the measurements are made in the gas phase (ion source). Ions are then introduced into a mass analyzer, where they are separated based on their mass-to-charge ratio (m/z) in an electric and/or magnetic field, before being recorded in a detector. The m/z can be used to deduce the mass of the analyte (protein, peptide or peptide fragment) with very high precision. There are several types of ion sources, mass analyzers and detectors. The type of ionization method used is determined by the analyte, but the correct combination of the mass analyzer and detector are the key for the quality and reliability of the analysis. The following section will give a brief introduction into the several components of a mass spectrometer as well as shortly discuss the different types of components available relevant to biological applications. Finally, the type of mass spectrometer used in this study will be introduced.
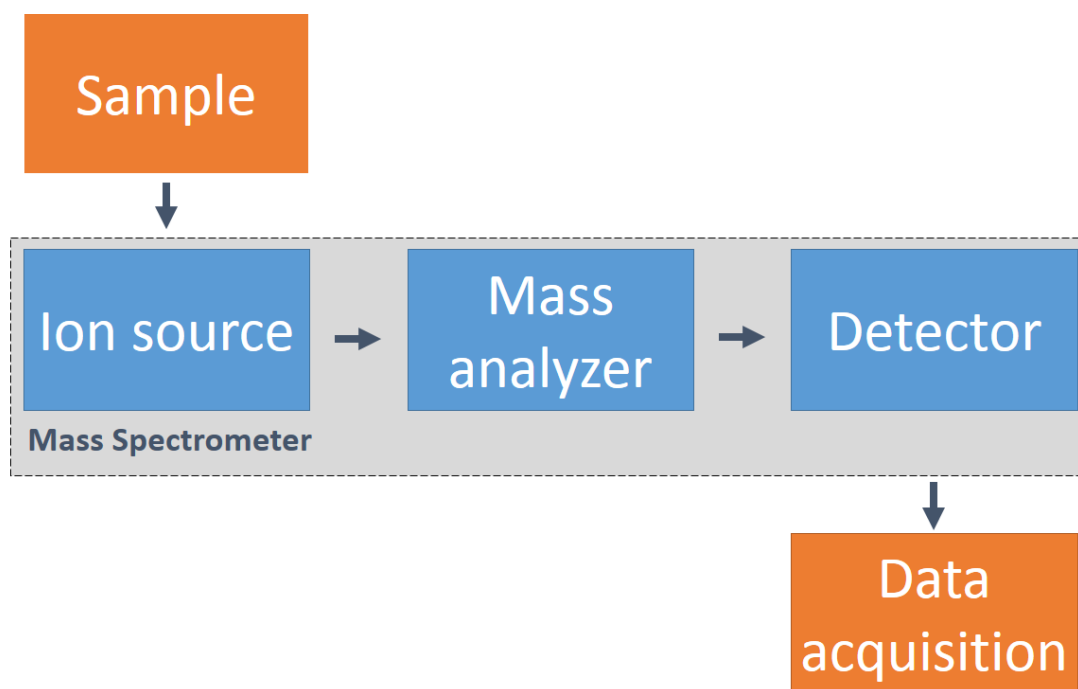
Figure 2. ***Basic components of a mass spectrometer.*** *Molecules from the sample first need to be ionized in a vacuum since the measurements are made in the gas phase (ion source), then ions are introduced into an electric and/or magnetic field (mass analyzer) where ions are separated based on their mass-to-charge ratio (m/z), before being detected (detector). The data is recorded in an acquisition unit.*

Ion source

Currently MS is the method of choice for the analysis of complex protein samples, but for many years it could not be used for the study of biomolecules such as proteins, peptides and nucleic acids due to the lack of proper ionization methods. This was first made possible by the discovery and introduction of two "soft" ionization methods in 1988, which were recognized in 2002 with the Nobel Prize in chemistry to John Fenn, Koichi Tanaka and Kurt Wüthrich **(Figure 3)**.

In the first method termed Matrix-Assisted Laser Desorption/Ionization (MALDI), macro-molecules are mixed with a light absorbing matrix, crystallized, ionized with a short pulse of laser and then desorbed into the vacuum system (10). This method is usually preferred to analyze relatively simple peptide and protein mixtures. In the second method, electrospray ionization or ESI, macromolecules are transferred from solution into the gas phase. In order to do this, the solution containing the analyte is passed through a needle to which a high voltage is applied. The solution is dispersed into a fine spray of charged droplets, these droplets are evaporated and the charge density increases until the Rayleigh limit is reached and a "Coulomb explosion" occurs, reaping the droplets apart. This process is repeated until the surface charge density is strong enough to desorb ions from the droplet into the ambient gas.

13

Protons added during this process give additional charge to the macromolecule (11). ESI can be easily coupled to liquid chromatography and is preferred when analyzing complex peptide samples.
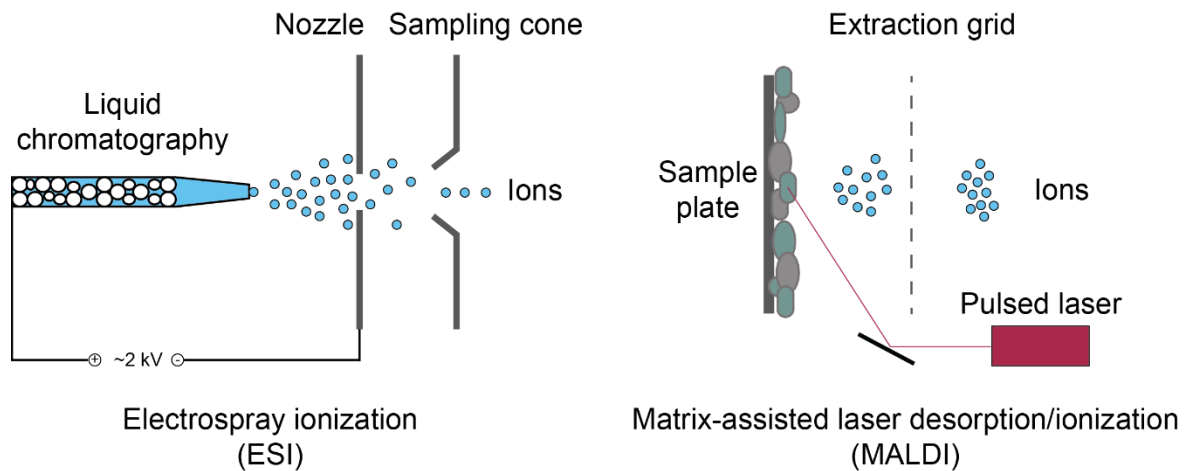
Figure 3. **Biomolecule ionization methods.** *On the left the electrospray ionization and ion introduction method is depicted, on the right matrix-assited laser desorption/ionization.*

Mass analyzer

Many different types of mass analyzers exist and are based on different principles. These include: quadrupole, magnetic sector, ion trap, time-of-flight (TOF), or Fourier transform (FT). These mass analyzers can also be combined to analyze both analytes and their fragments (MS/MS or MS^2). Commonly used mass analyzer combinations (hybrid mass spectrometers) are triple quadrupole, quadrupole/TOF and ion trap/Orbitrap. Analyzers like the ion traps can perform several MS/MS analysis, also called MS^n. Several intrinsic characteristics such as resolution, accuracy, sensitivity and dynamic range can vary, depending on the type of mass analyzer that is used. Each mass spectrometer has a set of parameters that define its performance; here I will briefly introduce them:

Mass Resolution is the ability to distinguish between two points with slightly different mass. Resolution is unit-less and depends on the m/z at which it is defined. For example, an instrument with a resolution of 240,000 @ 400 m/z can distinguish between two peaks with a mass difference of 0.0016 m/z or higher. An instrument with a resolution of 240,000 @ 600 m/z can "only" distinguish between two peaks with a mass difference of 0.0025 m/z or higher. **Mass Accuracy** is the average deviation of the measured m/z to the real m/z and is usually reported in relative units (ppm or parts-per-million). Mass accuracy in ppm is mass dependent. **Sensitivity** is the lowest amount of analyte that provides an MS response (signal). **Dynamic range of detection** is the ratio between the lowest and largest amounts of analyte that can be measured simultaneously in a single MS spectrum or in the whole experiment.

Scan rate is the analysis speed or frequency of performing MS or MS^n scanning under defined conditions.

With the increasing use of MS in biology, the complexity of samples increased, making the use of an instrument with high speed, high resolution and high accuracy indispensable. An overview of the different type of MS analyzers and combinations of these used in proteomics is presented in **Table 1**. In this study an instrument consisting of a combination of linear ion trap/Orbitrap was used. These two types of mass analyzers are described in the next sections.

Table 1. Overview of MS used in proteomics. Table adapted from (12).

| Instrument	Applications	Resolution	Mass accuracy	Sensitivity	Dynamic range	Scan rate
LIT (LTQ)	Bottom-up protein identification in high-complexity, high-throughput analysis, LC-MS^n capabilities	2,000	100 ppm	Femtomole	1e4	Fast
TQ (TSQ)	Bottom-up peptide and protein quantification; medium complexity samples, peptide and protein quantification (SRM, MRM, precursor, product, neutral fragment monitoring)	2,000	100 ppm	Attomole	1e6	Moderate
LTQ-Orbitrap	Protein identification, quantification, PTM identification	240,000	2 ppm	Femtomole	1e4	Moderate
LTQ-TICR, Q-FTICR	Protein identification, quantification, PTM identification, top-down protein identification	>500,000	<2 ppm	Femtomole	1e4	Slow, slow
Q-TOF, IT-TOF	Bottom-up, top-down protein identification, PTM identification	10,000	2–5 ppm	Attomole	1e6	Moderate, fast
Q-LIT	Bottom-up peptide and protein quantification; medium complexity samples, peptide and protein quantification (SRM, MRM, precursor, product, neutral fragment monitoring)	2,000	100 ppm	Attomole	1e6	Moderate, fast

Linear Ion Trap

An ion trap in its simplest definition is a combination of electric and/or magnetic fields to "trap" ions. Several types of ion traps exist, of which the quadrupole ion trap is currently one of the most broadly used analyzers. The invention of an electrodynamic ion trap was recognized in 1989 with the Nobel Prize in Physics to Wolfgang Paul and is therefore also often referred to as the Paul ion trap. Although these 3D traps are sensitive and relatively inexpensive, they have relatively low mass accuracy, due to the limited number of ions that can be accumulated in their point-like center before space-charging effects appear. This problem was circumvented with the introduction of the linear quadrupole ion trap **(Figure 4)**, where ions are stored in a cylindrical volume, which is considerably larger, allowing for increased sensitivity and mass accuracy (13). The ions are stored radially by a two-dimensional radio frequency field, and axially by stopping potentials applied to end electrodes. Apart from relatively low mass accuracy, linear ion trap profit from high sensitivity and very fast scan rates, making them the ideal analyzer to be combined with Orbitraps for proteomic applications.

15
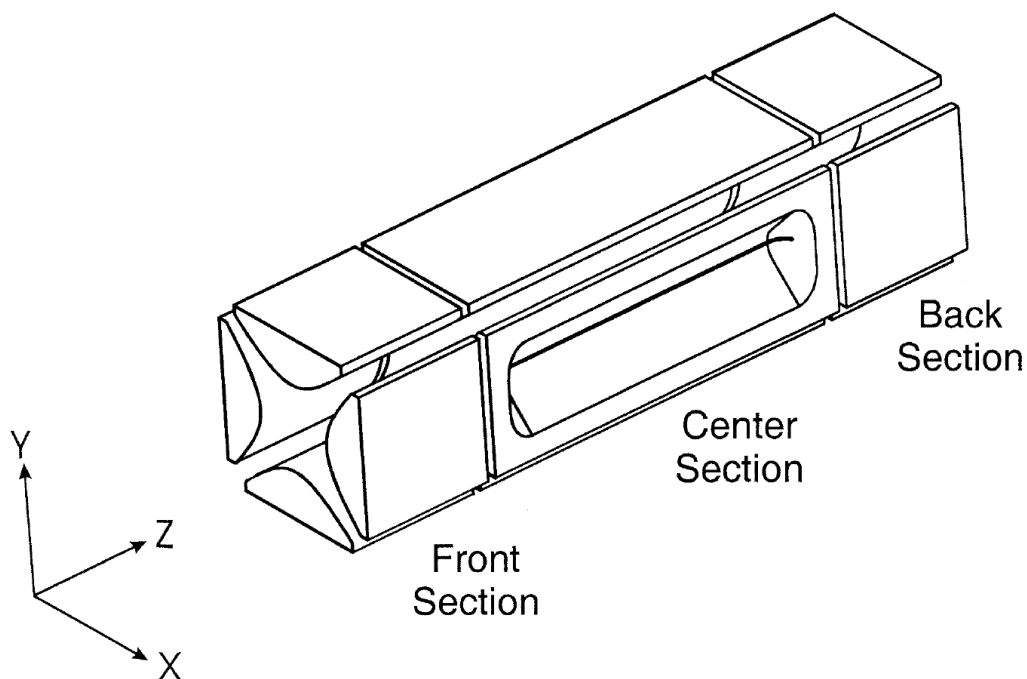
Figure 4. *Two dimensional ion trap.* *Basic design of a two dimensional ion trap. Ions are stored radially by a two-dimensional radio frequency field, and axially by stopping potentials applied to end electrodes. Adapted from (14)*

Orbitrap

Another type of (=electrostatic) ion trap, the Orbitrap was patented in 1999 and was presented shortly after in 2005 in combination with a LTQ linear ion-trap (15). A high-field compact trap version was introduced in 2011 (16). The Orbitrap works in combination with the C-trap, where the ions are stored and injected as an ion packet into the Orbitrap. The Orbitrap operates by trapping ions in its electrostatic fields in which they orbit around a central electrode and oscillate in axial direction. Axial ion frequencies are measured by a non-destructive method in which the time-domain signal is recorded by the outer Orbitrap electrodes. The produced signal is converted using Fourier transform algorithm (FT) into a mass-to-charge spectrum **(Figure 5)**. For a detailed explanation on how the Orbitrap works see (17) and (18). The high-field Orbitrap is a high resolution mass analyzer reaching a resolving power of >240,000 at m/z 400, high mass accuracy of < 2 ppm and a dynamic range greater than 5,000. The potential of the Orbitrap is fully revealed when coupled to an LTQ ion trap, configuration discussed in the next section.
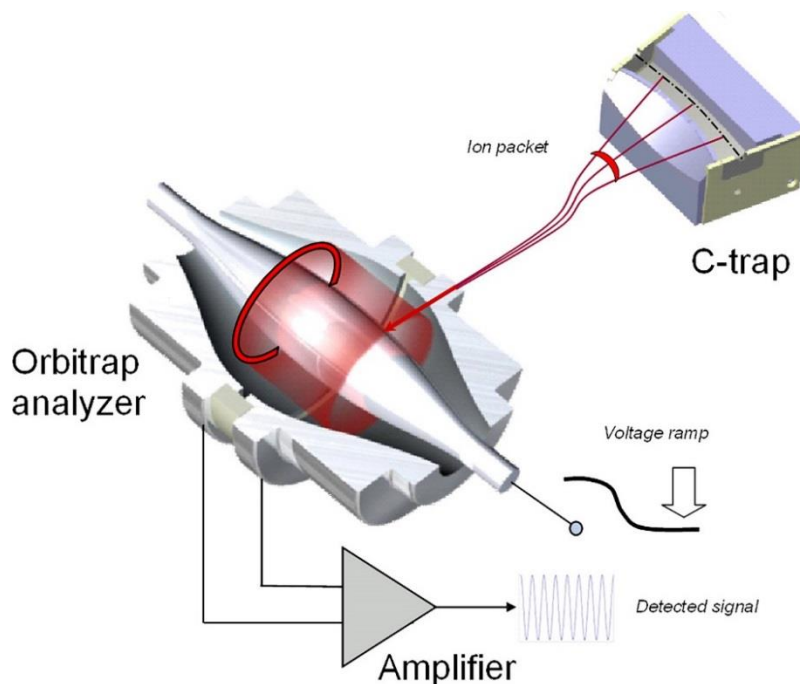
Figure 5. ***Cross-section of the Orbitrap analyzer.*** *Ions are stored in the C-trap and injected as an ion packet into the Orbitrap. Ions are radically trapped in static electrostatic fields. Ions orbit around a central electrode and oscillate in axial direction. Ion frequencies are measured and the produced signal is converted using Fourier transform algorithm (FT) into a mass-to-charge spectrum. Adapted from (15)*

Hybrid ion trap-orbitrap mass spectrometer

The first commercial instrument including an Orbitrap cell was a hybrid instrument (LTQ Orbitrap Classic), which had a linear ion trap in the front-end and the Orbitrap on the back-end. This instrument combines key qualities required in MS-based proteomics, the high resolution – accurate mass (HR/AM) capabilities of the Orbitrap analyzer with the speed and sensitivity of the LTQ ion trap **(Figure 6)**. Since the introduction of the Orbitrap, two major changes have been made to both analyzers; first, a two-stage LTQ ion trap was introduced with the Orbitrap Velos (19), which increased the sequencing speed from 6 to > 12 scans/s and two years later a high-field version of the Orbitrap (Orbitrap "Elite") was presented which doubled resolution from 120,000 to 240,000 combined with an enhanced Fourier Transform algorithm (eFT) which increases the measuring speed by a factor of 2 (16).

Hybrid mass spectrometers like the LTQ Orbitrap can be used to sequence peptides. This is done by combining the strengths of both mass analyzers. A common data acquisition strategy is termed High-Low, for high resolution MS[1] and low resolution MS[2] scan. In order to reduce the search space during data base search a high resolution and high accuracy measurement is required, therefore the peptide mass is measured in the Orbitrap. After ionization peptides are focused into a beam and guided to the C-trap. In the C-trap ions are stored and injected as an ion packet into the Orbitrap. Ions start orbiting around the central electrode; at this point
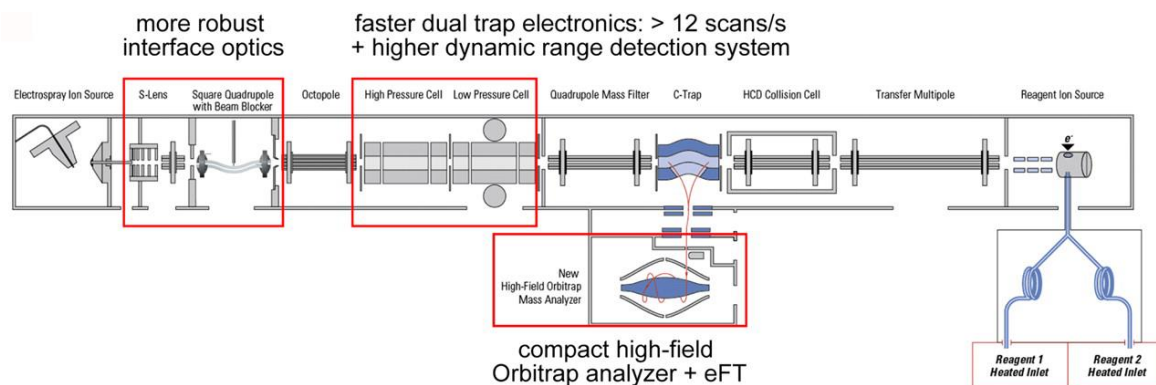
the ion frequencies are recorded. The signal produced is independent of rotational motion and therefore the m/z values of the ions can be calculated.

Based on this MS[1] or full-scan, the peptides with the highest intensity are selected for fragmentation (MS/MS) in a process termed data-dependent analysis (DDA). The chosen ions are consecutively isolated, fragmented and recorded in MS/MS (or MS[2]) spectrum **(Figure 7)**. For MS/MS measurements, high mass accuracy is less critical, but speed and sensitivity are very important. Fragmentation usually occurs in the ion trap in a process called collision induced dissociation (CID). During this process the peptides are fragmented by high-energy collision with a noble gas (e.g. helium). This process fragments peptide ions, with each individual peptide broken only on one place on average. Most commonly cleaved bonds are peptide bonds, giving rise to sequence-specific b- and y- fragment ions. The m/z ratios of all charged fragments are recorded in the LTQ. Another acquisition strategy also frequently used is termed "High-High" (16). In this approach both MS[1] and MS[2] measurements are performed at high accuracy and mass resolution in the Orbitrap analyzer. In this strategy peptide sequencing is done using high energy collision dissociation (HCD). This strategy results in HR/AM measurements of MS[2] spectra, which is very useful for the analysis of PTMs, at the cost of sequencing speed.

LC-MS/MS

The speed of this hybrid Orbitrap instrument makes it the ideal instrument to be coupled to high performance liquid chromatography (HPLC or LC) through ESI. Since ESI is a technique that profits from low solvent amounts it is an advantage to couple MS to nano-LC systems. The technique of choice to separate peptides before ionization is reversed-phase chromatography. In this method peptides are separated based on their hydrophobicity. The
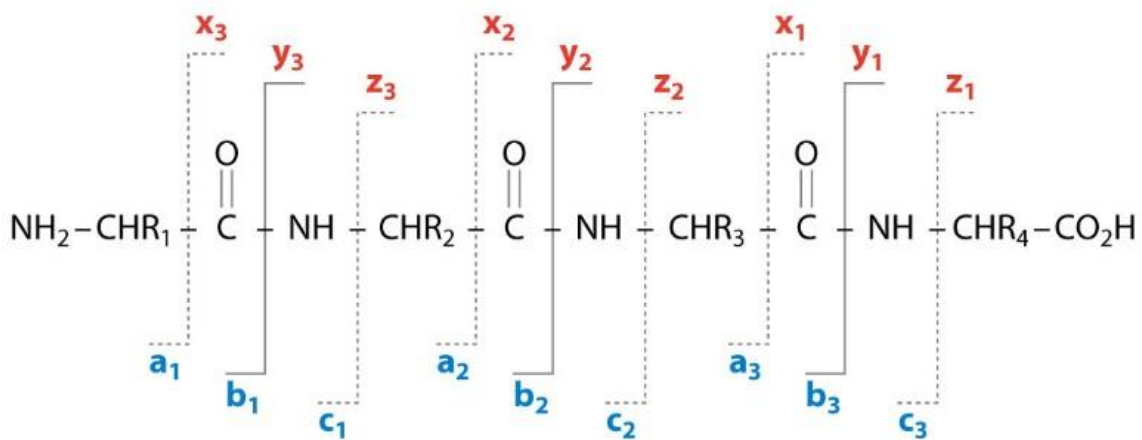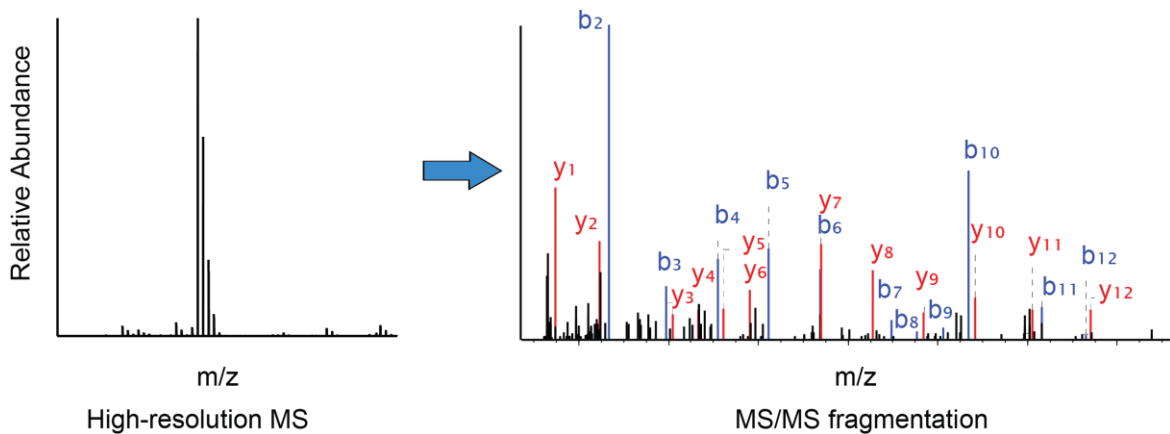
Figure 7. **MS, MS/MS spectra and peptide ion fragmentation.** *On the top panel a typical MS and MS/MS patterns are shown. A peptide is chosen from the full scan (MS) based on its intentity and fragmented. Fragment ions (MS/MS) are recorded in the linear ion trap. In the lower panel the different types of fragmentation along the peptide backbone are presented. The peptide fragments retaining the positive charge on the N-terminal are refered to as a, b and c, and the fragments retaining the positive charge on the C.terminal are refered as x, y, and z. Lower panel adapted from (20)*

capillary chromatographic column is filled with a reversed-phased material, normally alkyl chains bonded to a solid support. The most used alkyl length is octadecyl carbon chain (C18), since it is compatible with general properties of peptides. The peptides bind to this C18 material and are eluted with a hydrophobic mobile phase (acetonitrile) by increasing its concentration over time. A key development in RPLC has been the use of long, narrow capillary RP columns, whose advantages include capacity, sensitivity, dynamic range, peak capacity, sensitivity and reproducibility (21). Although reversed phase has established itself as the backbone in MS-based proteomics, the following types of chromatography can also be used: strong anion exchange (SAX) or strong cation exchange chromatography (SCX), hydrophilic-interaction chromatography (HILIC) and affinity chromatography or a combination of one or more of these types of chromatographies (e.g. MUDPIT (22)).

Data processing and database search

The workflow presented so far generates thousands of MS/MS spectra per hour of analysis making manual validation of each spectrum almost impossible. The recorded data is processed using bioinformatics approaches. The generated spectra are compared against a protein database, making the availability of sequencing data for the organism of interest a prerequisite. The analysis of this data and its assignment to peptides sequences and the determination of its abundance is a complex computational and statistical challenge. Several approaches have been developed over the years to automatically assign peptide sequences to fragment ion spectra. These approaches can be grouped into three categories (23):

(1) Database searching: ion fragments (MS/MS spectra) are correlated to theoretical spectra predicted for each peptide contained in a protein database or by correlating ion fragments to spectral libraries from spectra identified in previous experiments.

(2) *De novo* sequencing: peptide sequences are read directly from ion fragments.

(3) Hybrid approaches: extraction of short sequence tags of 3-5 residues followed by an organism specific database search.

Due to the nature of large scale MS-based proteomics, the most used approach is database searching, and this approach will be discussed in detail in the next section.
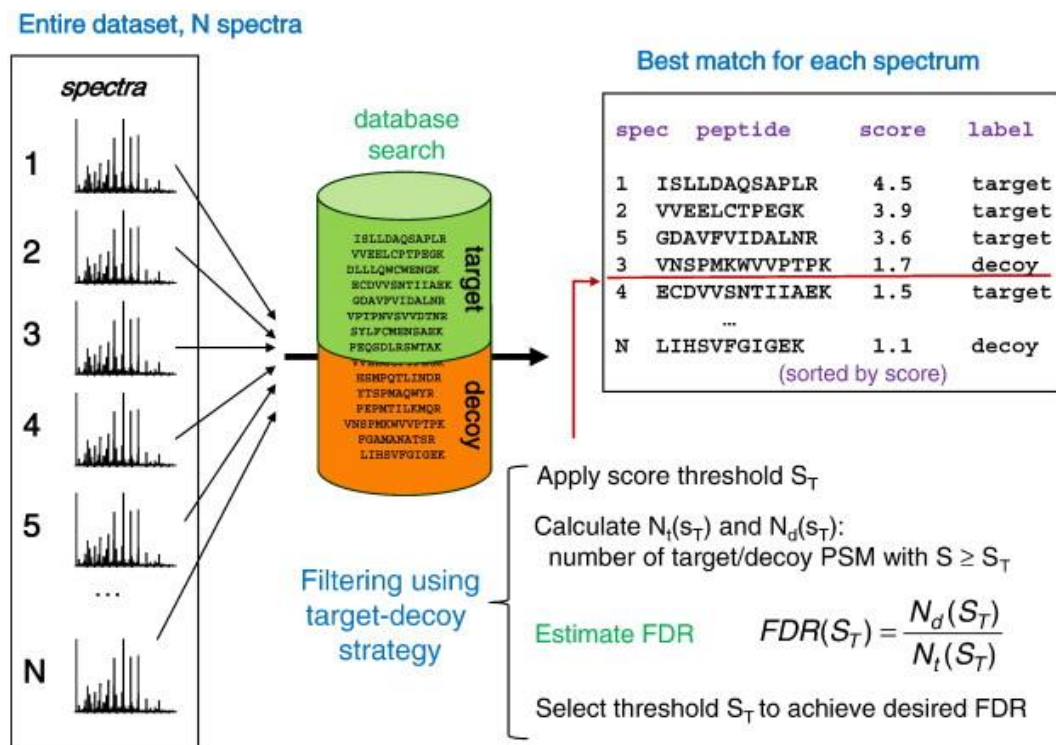
Figure 8. ***Database search: Target-decoy strategy for FDR assessment.*** *In this strategy all MS/MS spectra are searhced against a target/decoy database. The best match for each spectrum is integrated in a list containing all identified peptides and sorted by score. The number of decoy peptide mathces is used to estimate the false discovery rate (FDR) after filtering the data using several score thresholds. Adapted from (24).*

Database searching and statistical analysis

In this approach the fragment ion spectrum of a peptide is scored against theoretical fragmentation patterns from the searched database **(Figure 8)**. The size of this database (database search space) depends, apart from the number of proteins contained in the database, on the user-specified criteria such as mass tolerance, enzyme, missed cleavages and PTMs allowed. This search generates a list of matches ranked according to the search score, the best scoring peptides are then used for the subsequent analysis. The search score reflects how good an experimental spectrum matches a theoretical spectrum. There are many different scoring approaches available such as MASCOT (25), ANDROMEDA (21) , SEQUEST (26) and X! TANDEM (27).

From the results obtained from the search engines it is not possible to deduce what hits are true matches and which ones are false positives. Therefore it is critical to use statistical tools to discriminate true from false positives. To address this issue, proteomics rely on probabilistic approaches that provide a statistical measure of confidences and estimates of error rates (23). The statistical approaches can be grouped into two categories: target-decoy search (28) and empirical Bayes approaches (29). The most established method is the target decoy search, in which all the sequences of the database used for the search are reversed and attached to the database used for the search. This method assumes both matches to the target (forward database) and decoy ("nonsense" database; can be reversed or random) databases follow the same distribution. The matches are then filtered using different cut-offs and the corresponding false discovery rate (FDR) can be calculated by dividing the number of decoy hits by the number of targets above the cut-off. With this method the number of false positives can be calculated as the number of decoy matches.

1.3   Quantitative proteomics

One of the main advantages of MS-based proteomics is the possibility to directly compare two or more perturbations in a biological system. Several approaches have been developed each containing their respective advantages and disadvantages. The protein quantification can be either absolute or relative. An overview of this approaches is presented in **Figure 9**.

Relative quantification

There are several strategies for relative quantification, which can be grouped in two groups: "label-free" and stable isotope labeling methods (12). Label-free approaches are relatively easy to implement, yet they have a drawback that they are very sensitive to experimental variation and require acquisition of multiple replicates. These methods use either spectral counting or peptide signal intensity to estimate abundance of proteins (30). In the second
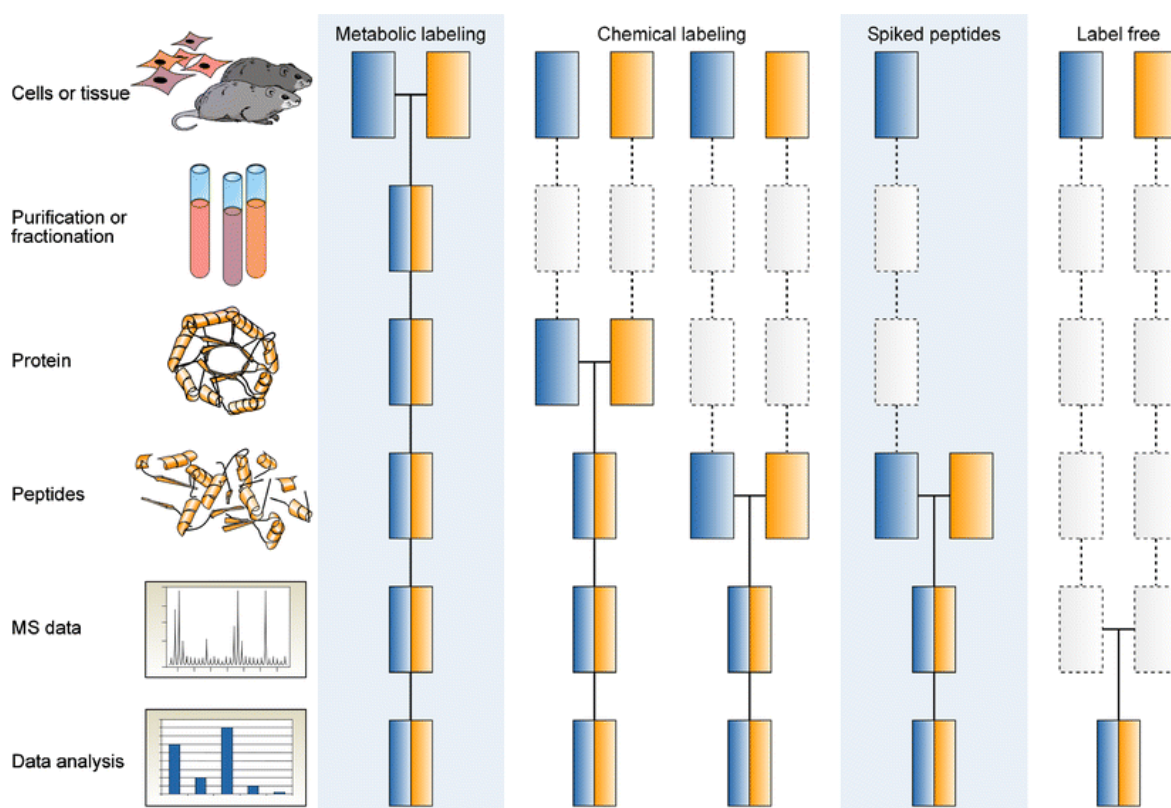
Figure 9. *Typical labling workflows in quantitative proteomics. The scheme shows several isotopic labelling strategies employed in quantitative proteomics. The two colors represent the two different labeling states being compared. The samples are combined with the horizonal lines. The overview outlines from what point on samples are processed together and when in parallel (quantification errors can ocurr). Adapted from (31).*

approach, stable (non-radioactive) isotope labels need to be introduced into one of the samples measured; after labeling and biological treatment, the two protein (or peptide) samples are mixed and each peptide is typically seen as a "doublet" in the mass spectrum (one member of the doublet comes from the "light"- the other from the "heavy"-labeled sample. Stable isotope labeling can be performed metabolically, chemically or enzymatically (32). Enzymatic labelling usually incorporates ^{18}O after or during digestion and it is not very widely used. Chemical introduction of labels can be done at the protein or at the peptide level. Labels can be introduced by e.g. labelling of free cysteines with isotope-coded affinity tags (ICAT) (33), or by labelling free amines with tandem mass tags (TMT or iTRAQ) (34) or with dimethyl labeling (35). In iTRAQ, quantification is performed using reporter ions in the tandem mass spectra, which results in limitations in quantification accuracy (36). Metabolic labeling using stable isotopes reduces the variation that can be introduced during sample handling, as it enables early mixing of differentially labeled samples. In this method the whole proteome can be labeled with a single stable isotope, e.g. ^{15}N (37); alternatively, a defined label can be introduced by addition of stable isotope labeled ("heavy") amino acids in the cell culture. Quantification is performed with labeled peptides at the MS level. The most widely used

method of metabolic labeling is stable isotope-labeling by amino acids in cell culture (SILAC) (38). This method is described in detail in the next section.

SILAC

In this metabolic labeling approach, amino acids containing stable isotopes are supplied in growth media and incorporated into newly synthesized proteins. After several cell doublings (usually 6-8) all the proteins will incorporate the added heavy amino acid. After digestion, ideally every peptide pair will be separated by the mass difference introduced by the label. Arginine and lysine are usually the amino acids of choice because digestion with trypsin will result in all tryptic peptides having at least one isotopically labeled amino acid. Alternatively only lysine can be used when digesting with endoproteinase Lys-C. It is crucial that the amino acid used for labeling is essential, i.e. that it cannot be synthesized by the cell. In some cases this can be bypassed by offering an excess of labeled amino acid to the cell or organism, although a general method is the generation of strains auxotrophic for the amino acid(s) of interest (32).

There are some general limitations when employing SILAC. One of the major challenges faced when applying SILAC is the arginine to proline conversion. It has been observed that arginine can be metabolically converted to proline in some organisms or cell types (39) (see Figure 1 Chapter 4). This can be circumvented for example by employing only heavy lysine for labelling (40), by titration of arginine and proline concentrations (41) or by computational methods (42). Another limitation when using SILAC is that some cell lines are sensitive to changes in the medium composition and may not be suitable for metabolic labelling. Also a maximum of three conditions can be compared in one experiment, although these has been circumvented by the combination of spike-in standards (43) or the use of more than one triple-SILAC experiment with a common standard (44). Because of the early combination of samples, SILAC is the most accurate quantitative MS method in terms of overall experimental process. This makes it especially interesting when assessing changes in levels of proteins or PTMs with low stoichiometry.

Absolute quantification

The methods mentioned before, although important in comparative biological studies, do not provide information on absolute protein quantities or concentrations. In the last years several techniques for absolute quantification that relays on the addition of a stable isotope labeled standard with a known concentration have been developed. The most well-known method is AQUA (Absolute QUAntification), in which peptides of known concentration with an incorporated stable isotope is spiked into the sample of interest to precisely quantify the

absolute amount of the target peptide and therefore of the target protein (45). Other methods such as "protein standard absolute quantification" (PSAQ) (46), FlexiQuant (47), and absolute quantification using protein epitope signature tags (PrEST) (48) have also been described in the literature. A method that has recently gained attention is the intensity based absolute quantification (iBAQ) (49); this method combines both label free quantification with a spiked in standard. The iBAQ value is calculated by summing all of the peptide intensities from a given protein and dividing it by the theoretical amount of MS-observable peptides. Absolute protein amounts are then calculated by correlating these values with iBAQ values from a spiked-in standard of known concentration. For more detailed description of iBAQ and other factors that are important in estimation of absolute protein amounts (e.g. cell count, volume, etc.), please refer to section 3.3.

Proteome analysis

The cell lysate is a very complex mixture of proteins, which becomes even more complex after enzymatic digestion. Low abundant proteins (peptides) are difficult to identify by data-dependent analysis (DDA), since this approach preferentially targets more abundant peptides. Lower abundant peptides will also co-elute with more abundant peptides making its identification more difficult. It has been calculated that from 100,000 peptides eluting during one HeLa lysate measurement only about 16% of them are targeted for MS/MS (50). Therefore several approaches have been successfully applied to reduce the complexity of the samples, such as subcellular (51), protein (GeLC (52) and gel filtration chromatography (53)), or peptide fractionation (Off-Gel (54) and StageTip based fractionation (55)).

Phosphoproteome analysis

Protein phosphorylation is one of the most important and most abundant posttranslational modifications (PTMs) in living cells and it regulates biological process in virtually every known organism. It has been involved in cell cycle control, receptor-mediated signal transduction, differentiation, proliferation, transformation, and general metabolism (20). Phosphorylation is regulated by two types of enzymes with opposite functions – proteins are phosphorylated by kinases and dephosphorylated by phosphatases in a highly conserved mechanism that almost exclusively involves adenosine triphosphate (ATP). Due to modern global proteomics studies, thousands of phosphorylated proteins have been identified, and it has been predicted that more than 500,000 phosphorylation sites are present in the human proteome (56).

Enrichment of STY-phosphorylated peptides

Although it has been proven that phosphorylation site stoichiometry (occupancy) can be very high during specific cellular processes, such as mitosis (44), protein phosphorylation occurs mostly at very low levels (i.e. it is substoichiometric) and its identification is therefore challenging. In order to detect phosphorylation events in a global scale, phosphopeptide enrichment strategies such as phosphopeptide immunoprecipitation (57), immobilized metal ion affinity chromatography (IMAC) (58) or titanium dioxide chromatography (TiO_2) (59) need to be employed. For example, TiO_2 chromatography relies on enrichment through high affinity absorption of phosphate groups (TiO_2-phosphate complex). A strategy commonly used to improve specificity is the addition of DHB, which binds with a higher affinity than acidic amino acid residues but with lower affinity than phosphate groups **(Figure 10)**. These strategies in combination with technological advances (increase of sensitivity and sequencing speed) have resulted in the publication of large phosphoproteomic datasets produced in our laboratory (among others) for human (60), nematodes (61), bacteria (62, 63) and yeast (40).
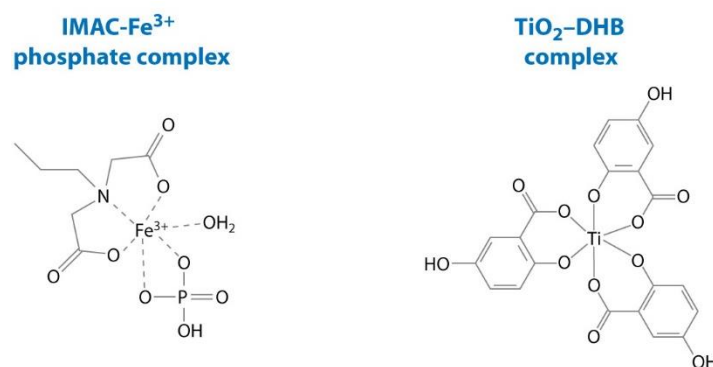

Figure 10. *Complexes of compounds commonly used for enrichment of phosphorylated peptides. On the left coordination of the a phosphate group to a immobilized metal, strategy used with IMAC. On the right TiO_2-DHB complex. DHB binds with a higher affinity than acidic amino acid residues but with a lower affinity than phosphate groups, increasing the specificity of TiO_2. Figure adapted from (20)*

Phosphorylation site stoichiometry (occupancy)

Protein phosphorylation is a dynamic process and not all molecules of a given protein are fully modified at specific residues all times – this phosphorylation site "occupancy" (stoichiometry) can be as low as <1% (i.e. one protein molecule in 100 is phosphorylated), or as high as 100% (all molecules are modified). Phosphorylation site occupancy has severe implications on the signaling potential of a protein and its extent is necessary for systems biologists to better understand regulation of biological process and to distinguish between functional and non-functional phosphorylation. Determining the occupancy of phosphorylation sites is now possible in large scale and two different equally successful approaches have been employed (44, 64). The first approach, also used in this study, calculates stoichiometry using SILAC ratios

of the phosphorylated peptide, the un-phosphorylated version of the same peptide and the protein ratios **(Figure 11)**. The second approach calculates stoichiometry by treating a fraction of the sample with phosphatase and then comparing relative peak intensities from both treated and non-treated fractions.
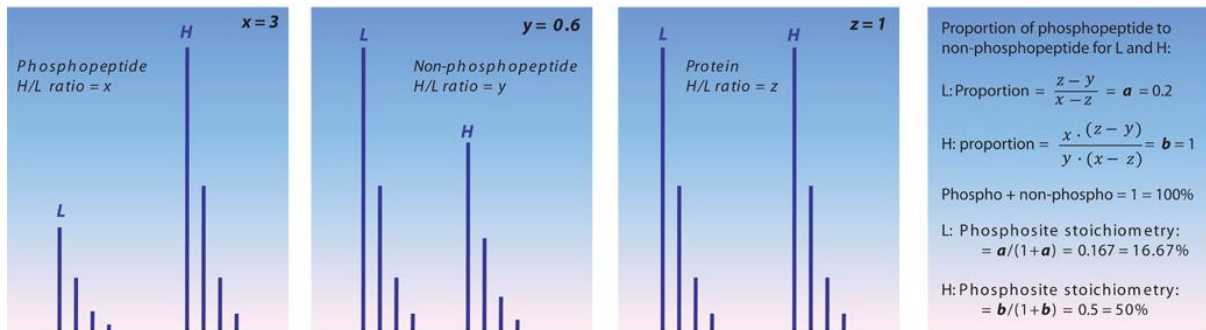
Figure 11. *Phosphorylation site soichiometry calculation. Example for the calculation of phosphorylation site occupancy based on three ratios: phosphopeptide, non-phosphopeptide and protein ratio. The phosphosite stoichiometry is calculated based on the proportion for each labelling state. Figure adapted from (44).*

Within the figure:

$x = 3$

Phosphopeptide H/L ratio = x

$y = 0.6$

Non-phosphopeptide H/L ratio = y

$z = 1$

Protein H/L ratio = z

Proportion of phosphopeptide to non-phosphopeptide for L and H:

L: Proportion $= \dfrac{z - y}{x - z} = a = 0.2$

H: proportion $= \dfrac{x \cdot (z - y)}{y \cdot (x - z)} = b = 1$

Phospho + non-phospho = 1 = 100%

L: Phosphosite stoichiometry:
$= a/(1+a) = 0.167 = 16.67\%$

H: Phosphosite stoichiometry:
$= b/(1+b) = 0.5 = 50\%$

Acetylome analysis

Lysine acetylation is a reversible PTM that has been related to multiple cellular processes such as transcription-related processes (65), metabolic regulation (66), chromatin remodeling, cell cycle, splicing, nuclear transport and actin nucleation (67). Lysine acetyltransferases (KATs) transfer an acetyl group to the ε-amino group of a lysine, therefore neutralizing the positive charge of this amino acid. The acetyl group is removed by lysine deacetylases (KDACs). Its role in gene regulation by histone modification has been deeply characterized and it was where it was originally discovered (68). Apart from histones a large number of proteins have been found to be lysine acetylated. Lysine residues are modified by other PTMs (methylation, ubiquitination, sumoylation, etc.), therefore there is a potential for cross-talk with this modifications. Furthermore, adjacent amino acids can also be modified presenting a potential for interplays between different PTMs. Some of this interplays have been shown to have important functional roles (69, 70). In the first large-scale lysine acetylation, Choudhary and colleagues involved lysine acetylation in diverse cellular processes (67). After this, several large-scale studies in several organisms like mouse (71), budding yeast (72), salmonella (73), *Drosophila* (74) and *Escherichia coli* (75) were published.
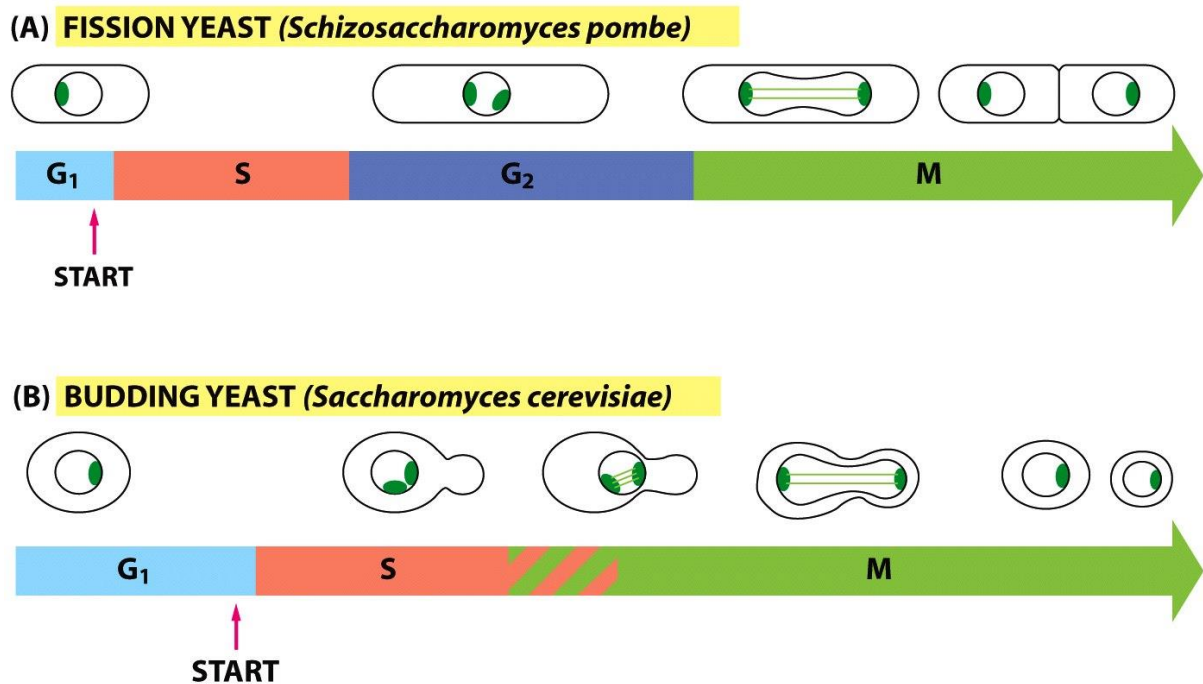
1.4   Fission yeast as a model organism

Schizosaccharomyces pombe, commonly known as fission yeast, is a unicellular eukaryotic organism that is amenable to genetic manipulation and carries many cell cycle features similar to cells from multicellular organisms. Fission yeast is an established model organism to study

processes such as cell cycle and checkpoint controls (76). Fission yeasts grow by elongating at both ends and divides by forming a wall through the middle of the cell. The length of the cell indicates its position in the cell cycle, making it the ideal organism to study cell cycle. Recent global proteomics studies of yeasts and their cell cycle (77-83) have mainly focused on *Saccharomyces cerevisiae* (budding yeast), with only few studies of fission yeast (84, 85), although the fission yeast cell cycle may be more representative of eukaryotic cell cycles (86). Recently, global studies of the fission yeast have been published covering up to 3,542 *S. pombe* proteins (71% of the predicted proteome) and 1,544 phosphorylation events in both, asynchronous and synchronized cell cultures (40, 87-90).

Cell cycle of fission yeast

Cell division requires a delicately orchestrated sequence of complex biochemical events that assures that every daughter cell obtains all the components required for survival. Investigations into the control of cell division in diverse cells have revealed universal regulatory mechanisms. Signaling mechanisms determine whether and when a cell undergoes cell division, and ensure orderly passage through the stages of the cell cycle. When the regulatory mechanisms that limit cell division are defective, the cells undergo unregulated division which could potentially result in cancer. Since a detailed understanding of the cell cycle and its

(A) FISSION YEAST *(Schizosaccharomyces pombe)*

G₁   S   G₂   M

START

(B) BUDDING YEAST *(Saccharomyces cerevisiae)*

G₁   S   M

START

Figure 12. ***The cell cycle phases.*** *Cell cycle of (A) fission yeast and (B) budding yeast. The cell cycle of most organisms is divided in four main phases: M, G1, S, and G2. M phase (mitosis) is usually followed by cell division. S phase (synthesis) is the period during which DNA replication occurs. G1 and G2 (gap) are waiting periods. Adapted from (91)*

regulation may lead to the identification of additional targets for cancer therapy, it is crucial to understand these underlying mechanisms (92-94).

The cell cycle has four major stages. In the S phase (synthesis), DNA is replicated. This is followed by G2 phase (or Gap phase), where proteins are synthesized before entering M phase (mitosis). During M phase the nuclear envelope breaks down, paired chromosomes are pulled to opposite poles of the cell, before being surrounded by a newly formed nuclear envelope, and cytokinesis pinches the cell in half, resulting in two daughter cells. Before starting cell cycle again cells go again though a waiting period called G1 phase (95) **(Figure 12)**. The events that take place during these phases must be coordinated with one another so that they occur in the appropriate order. This coordination of the cell cycle depends on several checkpoints and feedback controls that prevent entry into the next phase until the events of the proceeding phase have been completed (92).

The cell cycle is controlled by protein kinases. By phosphorylating specific proteins at precise time points, these protein kinases organize biochemical events of the cell that result in cell division. The typical cell cycle kinases are heterodimers with a regulatory subunit, cyclin, and a catalytic subunit, cyclin dependent protein kinase (CDK). In the absence of the cyclin, the catalytic subunit is virtually inactive. When cyclin binds, the catalytic site opens up, a residue essential to catalysis becomes accessible and the activity of the catalytic subunit increases 10,000-fold. Animal cells have at least 10 different cyclins and at least 8 CDKs, which act in various combinations at specific points in the cell cycle (95). CDK activities oscillate during the cell cycle as a result of regulation through phosphorylation or dephosphorylation of the CDK, controlled degradation of the cyclin subunit, periodic synthesis of CDKs and cyclins, and action of specific CDK-inhibiting proteins. In general, active CDKs enable a cell to enter a stage of cell division. The list of proteins that CDKs target continues to expand and the phosphorylation of key proteins at specific points in the cell cycle is key in modulating its activities. In the 1970s Paul Nurse identified a gene (cdc2) that when mutated would cause the growth of cells at elevated temperature to stop at certain points in the cell cycle (96). This identification was later recognized in 2001 with the Nobel Prize in Physiology or Medicine. In fission yeast cdc2 is responsible for both transition points at G1 and G2. After the first transition point in late G1, the cell commits to replicate its DNA. This pass requires activation of Cdc2 by G1 cyclins (levels rise during G1). This activation leads to replication at sites where pre-replication complexes had previously assembled (97).

The G2/M passage requires activation by mitotic cyclins. Cdks containing a mitotic cyclin phosphorylate substrates that are required for the cell to enter mitosis. The substrates include

proteins required for changes in organization of the chromosomes and cytoskeleton. Exit from mitosis and entry into G1 depend on a rapid decrease in Cdk activity that results from a drop in concentration of the mitotic cyclins. When the cell reaches a critical size, the cdc25 phosphatase is activated and removes the inhibitory phosphate on Tyr 15 residue. The resulting activation of cdc2 kinase drives the cell into mitosis (97). The mutation of the cdc25 gene results in cells that do not divide but continue to grow. A temperature sensitive mutation of this gene was used to arrest cells in one study presented later (98).

MS-based proteomics is the ideal tool to study protein dynamics and phosphorylation dynamics during the cell cycle. In a very comprehensive study Olsen et al. (44) measured the dynamics of thousands of proteins and phosphorylation events across cell cycle phases of HeLa cells, providing insights into the underlying regulatory mechanisms. In this study they found out that many phosphorylation sites have very high occupancy (>90%) and that in general there is an increase in occupancy during M phase. In another study, Pagliuca et al. (99) investigated proteins that interact with cyclins in HeLa cells. In this study they identified mechanisms by which cyclin-cdks control the cell cycle and characterized the interaction between DNA replication and mitosis giving insight into the cell cycle machinery.

SILAC labelling in fission yeast

Quantitative proteomics based on metabolic labeling require full incorporation of labeled amino acids. As discussed above, some organisms have the ability to convert arginine to proline, which results in tryptic peptides containing heavy proline (100). To circumvent this problem several experimental and bioinformatics approaches have been developed (41, 101-103). The degree of conversion varies from organism to organism. *S. pombe* is especially efficient and researchers need to deal with this problem. This can be by-passed by either employing stable isotope-marked lysine only (40) or by using strains that only very inefficiently convert arginine to proline (104). These approaches although successful have certain limitations. These limitations are thoroughly discussed in paper 2 as well as an experimental strategy to circumvent these issues.

1.5 Aims

The overall goal of this thesis is to develop mass spectrometry based strategies for the study of fission yeast and its cell cycle. The strategy followed a three step process: 1) optimization of quantitative proteomics and phosphoproteomics methods, 2) study of the absolute proteome and phosphoproteome dynamics during the cell cycle of the fission yeast and 3) solving the "Arg10 labelling" problem in fission yeast and. The specific aims are:

1. Optimization of quantitative proteomics and phosphorylation and acetylation enrichment methods
    a. Eliminate the use of DHB in order to decrease chemical noise in the sample.
    b. Improve phosphopeptide enrichment using lower protein amounts.
    c. Establish a lysine acetylation enrichment method.
2. Study of the absolute proteome and phosphoproteome dynamics during the cell cycle of the fission yeast
    a. Increase the understanding of the cell cycle and its regulation through characterization at both proteome and phosphoproteome level.
    b. Measure relative and absolute dynamics of the proteome and phosphoproteome during the cell cycle of fission yeast.
    c. Estimate copy numbers of proteins.
    d. Calculate phosphorylation site stoichiometry.
    e. Estimate total amount of protein-bound phosphate, its dynamics across the cell cycle and implications for overall ATP consumption during the cell cycle.
3. Solving the "Arg10 labelling" problem in fission yeast
    a. Improve the standard growth media and adapt it for SILAC experiments.
    b. Enable the use of heavy arginine.

1.6 References

1. Lodish, H. (2008) *Molecular cell biology*, Macmillan

2. Birney, E., et al. (2007) Identification and analysis of functional elements in 1% of the human genome by the ENCODE pilot project. *Nature* 447, 799-816

3. Staden, R. (1979) A strategy of DNA sequencing employing computer programs. *Nucleic acids research* 6, 2601-2610

4. Anderson, S. (1981) Shotgun DNA sequencing using cloned DNase I-generated fragments. *Nucleic acids research* 9, 3015-3027

5. Schuster, S. C. (2008) Next-generation sequencing transforms today's biology. *Nature methods* 5, 16-18

6. Pandey, A., and Mann, M. (2000) Proteomics to study genes and genomes. *Nature* 405, 837-846

7. Aebersold, R., and Mann, M. (2003) Mass spectrometry-based proteomics. *Nature* 422, 198-207

8. Eisenberg, D., Marcotte, E. M., Xenarios, I., and Yeates, T. O. (2000) Protein function in the post-genomic era. *Nature* 405, 823-826

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

10. Tanaka, K., Waki, H., Ido, Y., Akita, S., Yoshida, Y., Yoshida, T., and Matsuo, T. (1988) Protein and polymer analyses up to m/z 100 000 by laser ionization time-of-flight mass spectrometry. *Rapid Communications in Mass Spectrometry* 2, 151-153

11. 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

12. Yates, J. R., Ruse, C. I., and Nakorchevsky, A. (2009) Proteomics by mass spectrometry: approaches, advances, and applications. *Annual review of biomedical engineering* 11, 49-79

13. Hager, J. W. (2002) A new linear ion trap mass spectrometer. *Rapid Communications in Mass Spectrometry* 16, 512-526

14. Schwartz, J. C., Senko, M. W., and Syka, J. E. (2002) A two-dimensional quadrupole ion trap mass spectrometer. *Journal of the American Society for Mass Spectrometry* 13, 659-669

15. Zubarev, R. A., and Makarov, A. (2013) Orbitrap Mass Spectrometry. *Analytical Chemistry* 85, 5288-5296

16. Michalski, A., Damoc, E., Lange, O., Denisov, E., Nolting, D., Müller, 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* 11

17. Hu, Q., Noll, R. J., Li, H., Makarov, A., Hardman, M., and Graham Cooks, R. (2005) The Orbitrap: a new mass spectrometer. *Journal of mass spectrometry : JMS* 40, 430-443

18. Makarov, A. (2000) Electrostatic Axially Harmonic Orbital Trapping: A High-Performance Technique of Mass Analysis. *Analytical Chemistry* 72, 1156-1162

19. 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. *Molecular & Cellular Proteomics* 8, 2759-2769

20. Macek, B., Mann, M., and Olsen, J. V. (2009) Global and Site-Specific Quantitative Phosphoproteomics: Principles and Applications. *Annual Review of Pharmacology and Toxicology* 49, 199-221

21. Thakur, S. S., Geiger, T., Chatterjee, B., Bandilla, P., Frohlich, F., Cox, J., and Mann, M. (2011) Deep and highly sensitive proteome coverage by LC-MS/MS without prefractionation. *Molecular & cellular proteomics : MCP* 10, M110.003699

22. Washburn, M. P., Wolters, D., and Yates, J. R. (2001) Large-scale analysis of the yeast proteome by multidimensional protein identification technology. *Nat Biotech* 19, 242-247

23. Nesvizhskii, A. I., Vitek, O., and Aebersold, R. (2007) Analysis and validation of proteomic data generated by tandem mass spectrometry. *Nat Meth* 4, 787-797

24. Nesvizhskii, A. I. (2010) A survey of computational methods and error rate estimation procedures for peptide and protein identification in shotgun proteomics. *Journal of Proteomics* 73, 2092-2123

25. 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

26. Eng, J. K., McCormack, A. L., and Yates, J. R. (1994) An approach to correlate tandem mass spectral data of peptides with amino acid sequences in a protein database. *Journal of the American Society for Mass Spectrometry* 5, 976-989

27. Craig, R., and Beavis, R. C. (2004) TANDEM: matching proteins with tandem mass spectra. *Bioinformatics (Oxford, England)* 20, 1466-1467

28. Elias, J. E., and Gygi, S. P. (2007) Target-decoy search strategy for increased confidence in large-scale protein identifications by mass spectrometry. *Nature methods* 4, 207-214

29. Storey, J. D., and Tibshirani, R. (2003) Statistical significance for genomewide studies. *Proceedings of the National Academy of Sciences of the United States of America* 100, 9440-9445

30. Domon, B., and Aebersold, R. (2006) Mass Spectrometry and Protein Analysis. *Science* 312, 212-217

31. Bantscheff, M., Schirle, M., Sweetman, G., Rick, J., and Kuster, B. (2007) Quantitative mass spectrometry in proteomics: a critical review. *Anal Bioanal Chem* 389, 1017-1031

32. Ong, S. E., and Mann, M. (2005) Mass spectrometry-based proteomics turns quantitative. *Nature chemical biology* 1, 252-262

33. 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

34. 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

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

36. Bantscheff, M., Boesche, M., Eberhard, D., Matthieson, T., Sweetman, G., and Kuster, B. (2008) Robust and Sensitive iTRAQ Quantification on an LTQ Orbitrap Mass Spectrometer. *Molecular & Cellular Proteomics* 7, 1702-1713

37. 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

38. 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* 1, 376-386

39. Ong, S.-E., Kratchmarova, I., and Mann, M. (2002) Properties of 13C-Substituted Arginine in Stable Isotope Labeling by Amino Acids in Cell Culture (SILAC). *Journal of Proteome Research* 2, 173-181

40. Koch, A., Krug, K., Pengelley, S., Macek, B., and Hauf, S. (2011) Mitotic substrates of the kinase aurora with roles in chromatin regulation identified through quantitative phosphoproteomics of fission yeast. *Science signaling* 4, rs6

41. Bendall, S. C., Hughes, C., Stewart, M. H., Doble, B., Bhatia, M., and Lajoie, G. A. (2008) Prevention of amino acid conversion in SILAC experiments with embryonic stem cells. *Molecular & cellular proteomics : MCP* 7, 1587-1597

42. Park, S. K., Liao, L., Kim, J. Y., and Yates, J. R., 3rd (2009) A computational approach to correct arginine-to-proline conversion in quantitative proteomics. *Nature methods* 6, 184-185

43. Geiger, T., Cox, J., Ostasiewicz, P., Wisniewski, J. R., and Mann, M. (2010) Super-SILAC mix for quantitative proteomics of human tumor tissue. *Nat Meth* 7, 383-385

44. Olsen, J. V., Vermeulen, M., Santamaria, A., Kumar, C., Miller, M. L., Jensen, L. J., Gnad, F., Cox, J., Jensen, T. S., Nigg, E. A., Brunak, S., and Mann, M. (2010) Quantitative Phosphoproteomics Reveals Widespread Full Phosphorylation Site Occupancy During Mitosis. *Sci. Signal.* 3, ra3-

45. 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. *Proceedings of the National Academy of Sciences of the United States of America* 100, 6940-6945

46. 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* 6, 2139-2149

47. Singh, S., Springer, M., Steen, J., Kirschner, M. W., and Steen, H. (2009) FLEXIQuant: A Novel Tool for the Absolute Quantification of Proteins, and the Simultaneous Identification and Quantification of Potentially Modified Peptides. *Journal of Proteome Research* 8, 2201-2210

48. Zeiler, M., Straube, W. L., Lundberg, E., Uhlen, M., and Mann, M. (2012) A Protein Epitope Signature Tag (PrEST) Library Allows SILAC-based Absolute Quantification and Multiplexed Determination of Protein Copy Numbers in Cell Lines. *Molecular & Cellular Proteomics* 11

49. Schwanhausser, B., Busse, D., Li, N., Dittmar, G., Schuchhardt, J., Wolf, J., Chen, W., and Selbach, M. (2011) Global quantification of mammalian gene expression control. *Nature* 473, 337-342

50. Michalski, A., Cox, J., and Mann, M. (2011) More than 100,000 Detectable Peptide Species Elute in Single Shotgun Proteomics Runs but the Majority is Inaccessible to Data-Dependent LC−MS/MS. *Journal of Proteome Research* 10, 1785-1793

51. Andersen, J. S., and Mann, M. (2006) Organellar proteomics: turning inventories into insights. *EMBO reports* 7, 874-879

52. Shevchenko, A., Wilm, M., Vorm, O., and Mann, M. (1996) Mass Spectrometric Sequencing of Proteins from Silver-Stained Polyacrylamide Gels. *Analytical Chemistry* 68, 850-858

53. Nagaraj, N., Wisniewski, J. R., Geiger, T., Cox, J., Kircher, M., Kelso, J., Paabo, S., and Mann, M. (2011) Deep proteome and transcriptome mapping of a human cancer cell line. *Molecular systems biology* 7, 548

54. Hubner, N. C., Ren, S., and Mann, M. (2008) Peptide separation with immobilized pI strips is an attractive alternative to in-gel protein digestion for proteome analysis. *Proteomics* 8, 4862-4872

55. Wiśniewski, J. R., Zougman, A., and Mann, M. (2009) Combination of FASP and StageTip-Based Fractionation Allows In-Depth Analysis of the Hippocampal Membrane Proteome. *Journal of Proteome Research* 8, 5674-5678

56. Lemeer, S., and Heck, A. J. (2009) The phosphoproteomics data explosion. *Current opinion in chemical biology* 13, 414-420

57. Rush, J., Moritz, A., Lee, K. A., Guo, A., Goss, V. L., Spek, E. J., Zhang, H., Zha, X.-M., Polakiewicz, R. D., and Comb, M. J. (2005) Immunoaffinity profiling of tyrosine phosphorylation in cancer cells. *Nat Biotech* 23, 94-101

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

59. Pinkse, M. W. H., Uitto, P. M., Hilhorst, M. J., Ooms, B., and Heck, A. J. R. (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

60. Franz-Wachtel, M., Eisler, S. A., Krug, K., Wahl, S., Carpy, A., Nordheim, A., Pfizenmaier, K., Hausser, A., and Macek, B. (2012) Global detection of protein kinase D-dependent phosphorylation events in nocodazole-treated human cells. *Molecular & cellular proteomics : MCP* 11, 160-170

61. Borchert, N., Krug, K., Gnad, F., Sinha, A., Sommer, R. J., and Macek, B. (2012) Phosphoproteome of Pristionchus pacificus provides insights into architecture of signaling networks in nematode models. *Molecular & cellular proteomics : MCP* 11, 1631-1639

62. Ravikumar, V., Shi, L., Krug, K., Derouiche, A., Jers, C., Cousin, C., Kobir, A., Mijakovic, I., and Macek, B. (2014) Quantitative phosphoproteome analysis of Bacillus subtilis reveals novel substrates of the kinase PrkC and phosphatase PrpC. *Molecular & cellular proteomics : MCP*

63. Soares, N. C., Spat, P., Krug, K., and Macek, B. (2013) Global dynamics of the Escherichia coli proteome and phosphoproteome during growth in minimal medium. *J Proteome Res* 12, 2611-2621

64. Wu, R., Haas, W., Dephoure, N., Huttlin, E. L., Zhai, B., Sowa, M. E., and Gygi, S. P. (2011) A large-scale method to measure absolute protein phosphorylation stoichiometries. *Nat Meth* 8, 677-683

65. Lee, K. K., and Workman, J. L. (2007) Histone acetyltransferase complexes: one size doesn't fit all. *Nat Rev Mol Cell Biol* 8, 284-295

66. Zhao, S., Xu, W., Jiang, W., Yu, W., Lin, Y., Zhang, T., Yao, J., Zhou, L., Zeng, Y., Li, H., Li, Y., Shi, J., An, W., Hancock, S. M., He, F., Qin, L., Chin, J., Yang, P., Chen, X., Lei, Q., Xiong, Y., and Guan, K.-L. (2010) Regulation of Cellular Metabolism by Protein Lysine Acetylation. *Science* 327, 1000-1004

67. Choudhary, C., Kumar, C., Gnad, F., Nielsen, M. L., Rehman, M., Walther, T. C., Olsen, J. V., and Mann, M. (2009) Lysine acetylation targets protein complexes and co-regulates major cellular functions. *Science* 325, 834-840

68. Strahl, B. D., and Allis, C. D. (2000) The language of covalent histone modifications. *Nature* 403, 41-45

69. van Noort, V., Seebacher, J., Bader, S., Mohammed, S., Vonkova, I., Betts, M. J., Kuhner, S., Kumar, R., Maier, T., O'Flaherty, M., Rybin, V., Schmeisky, A., Yus, E., Stulke, J., Serrano, L., Russell, R. B., Heck, A. J., Bork, P., and Gavin, A. C. (2012) Cross-talk between phosphorylation and lysine acetylation in a genome-reduced bacterium. *Molecular systems biology* 8, 571

70. Sims, R. J., and Reinberg, D. (2008) Is there a code embedded in proteins that is based on post-translational modifications? *Nat Rev Mol Cell Biol* 9, 815-820

71. Lundby, A., Lage, K., Weinert, Brian T., Bekker-Jensen, Dorte B., Secher, A., Skovgaard, T., Kelstrup, Christian D., Dmytriyev, A., Choudhary, C., Lundby, C., and Olsen, Jesper V. (2012) Proteomic Analysis of Lysine Acetylation Sites in Rat Tissues Reveals Organ Specificity and Subcellular Patterns. *Cell Reports* 2, 419-431

72. Henriksen, P., Wagner, S. A., Weinert, B. T., Sharma, S., Bacinskaja, G., Rehman, M., Juffer, A. H., Walther, T. C., Lisby, M., and Choudhary, C. (2012) Proteome-wide analysis of lysine acetylation suggests its broad regulatory scope in Saccharomyces cerevisiae. *Molecular & cellular proteomics : MCP* 11, 1510-1522

73. Wang, Q., Zhang, Y., Yang, C., Xiong, H., Lin, Y., Yao, J., Li, H., Xie, L., Zhao, W., Yao, Y., Ning, Z. B., Zeng, R., Xiong, Y., Guan, K. L., Zhao, S., and Zhao, G. P. (2010) Acetylation of metabolic enzymes coordinates carbon source utilization and metabolic flux. *Science* 327, 1004-1007

74. Weinert, B. T., Wagner, S. A., Horn, H., Henriksen, P., Liu, W. R., Olsen, J. V., Jensen, L. J., and Choudhary, C. (2011) Proteome-Wide Mapping of the Drosophila Acetylome Demonstrates a High Degree of Conservation of Lysine Acetylation. *Sci. Signal.* 4, ra48-

75. Weinert, B. T., Iesmantavicius, V., Wagner, S. A., Scholz, C., Gummesson, B., Beli, P., Nystrom, T., and Choudhary, C. (2013) Acetyl-phosphate is a critical determinant of lysine acetylation in E. coli. *Mol Cell* 51, 265-272

76. Forsburg, S. L., and Nurse, P. (1991) Cell Cycle Regulation in the Yeasts Saccharomyces Cerevisiae and Schizosaccharomyces Pombe. *Annu. Rev. Cell Biol.* 7, 227-256

77. Ghaemmaghami, S., Huh, W., Bower, K., Howson, R. W., Belle, A., Dephoure, N., O'Shea, E. K., and Weissman, J. S. (2003) Global analysis of protein expression in yeast. *Nature* 425, 737-741

78. Ross, P. L., Huang, Y. L. 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. *Mol. Cell. Proteomics* 3, 1154-1169

79. 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. *Nat. Biotechnol.* 17, 994-999

80. Ho, Y., et al. (2002) Systematic identification of protein complexes in Saccharomyces cerevisiae by mass spectrometry. *Nature* 415, 180-183

81. de Godoy, L. M. F., Olsen, J. V., Cox, J., Nielsen, M. L., Hubner, N. C., Froehlich, F., Walther, T. C., and Mann, M. (2008) Comprehensive mass-spectrometry-based proteome quantification of haploid versus diploid yeast. *Nature* 455, 1251-U1260

82. Flory, M. R., Lee, H., Bonneau, R., Mallick, P., Serikawa, K., Morris, D. R., and Aebersold, R. (2006) Quantitative proteomic analysis of the budding yeast cell cycle using acid-cleavable isotope-coded affinity tag reagents. *Proteomics* 6, 6146-6157

83. Breitkreutz, A., Choi, H., Sharom, J. R., Boucher, L., Neduva, V., Larsen, B., Lin, Z. Y., Breitkreutz, B. J., Stark, C., Liu, G. M., Ahn, J., Dewar-Darch, D., Reguly, T., Tang, X. J., Almeida, R., Qin, Z. S., Pawson, T., Gingras, A. C., Nesvizhskii, A. I., and Tyers, M. (2010) A Global Protein Kinase and Phosphatase Interaction Network in Yeast. *Science* 328, 1043-1046

84. Schmidt, M. W., Houseman, A., Ivanov, A. R., and Wolf, D. A. (2007) Comparative proteomic and transcriptomic profiling of the fission yeast Schizosaccharomyces pombe. *Molecular systems biology* 3, 79

85. Matsuyama, A., Arai, R., Yashiroda, Y., Shirai, A., Kamata, A., Sekido, S., Kobayashi, Y., Hashimoto, A., Hamamoto, M., Hiraoka, Y., Horinouchi, S., and Yoshida, M. (2006) ORFeome cloning and global analysis of protein localization in the fission yeast Schizosaccharomyces pombe. *Nat Biotech* 24, 841-847

86. Nurse, P. (1985) Cell cycle control genes in yeast. *Trends in Genetics* 1, 51-55

87. Marguerat, S., Schmidt, A., Codlin, S., Chen, W., Aebersold, R., and Bähler, J. (2012) Quantitative Analysis of Fission Yeast Transcriptomes and Proteomes in Proliferating and Quiescent Cells. *Cell* 151, 671-683

88. Gunaratne, J., Schmidt, A., Quandt, A., Neo, S. P., Sarac, O. S., Gracia, T., Loguercio, S., Ahrne, E., Xia, R. L., Tan, K. H., Lossner, C., Bahler, J., Beyer, A., Blackstock, W., and Aebersold, R. (2013) Extensive mass spectrometry-based analysis of the fission yeast proteome: the Schizosaccharomyces pombe PeptideAtlas. *Molecular & cellular proteomics : MCP* 12, 1741-1751

89. Wilson-Grady, J. T., Villen, J., and Gygi, S. P. (2008) Phosphoproteome analysis of fission yeast. *J. Proteome Res.* 7, 1088-1097

90. Beltrao, P., Trinidad, J. C., Fiedler, D., Roguev, A., Lim, W. A., Shokat, K. M., Burlingame, A. L., and Krogan, N. J. (2009) Evolution of phosphoregulation: comparison of phosphorylation patterns across yeast species. *PLoS Biol.* 7, e1000134

91. Alberts, B., Johnson, A., Lewis, J., Raff, M., Roberts, K., and Walter, P. (2008) Molecular biology of the cell. *New York: Garland Science* 1

92. Kastan, M. B., and Bartek, J. (2004) Cell-cycle checkpoints and cancer. *Nature* 432, 316-323

93. Malumbres, M., and Barbacid, M. (2009) Cell cycle, CDKs and cancer: a changing paradigm. *Nat. Rev. Cancer* 9, 153-166

94. Lapenna, S., and Giordano, A. (2009) Cell cycle kinases as therapeutic targets for cancer. *Nat. Rev. Drug Discov.* 8, 547-566

95. Lehninger, A., Nelson, D., and Cox, M. (2008) *Lehninger Principles of Biochemistry*, W. H. Freeman

96. Nurse, P. (1975) Genetic control of cell size at cell division in yeast. *Nature* 256, 547-551

97. Coleman, T. R., and Dunphy, W. G. (1994) Cdc2 regulatory factors. *Current Opinion in Cell Biology* 6, 877-882

98. Carpy, A., Krug, K., Graf, S., Koch, A., Popic, S., Hauf, S., and Macek, B. (2014) Absolute proteome and phosphoproteome dynamics during the cell cycle of fission yeast. *Molecular & cellular proteomics : MCP*

99. Pagliuca, Felicia W., Collins, Mark O., Lichawska, A., Zegerman, P., Choudhary, Jyoti S., and Pines, J. (2011) Quantitative Proteomics Reveals the Basis for the Biochemical Specificity of the Cell-Cycle Machinery. *Molecular cell* 43, 406-417

100. Ong, S. E., Kratchmarova, I., and Mann, M. (2003) Properties of 13C-substituted arginine in stable isotope labeling by amino acids in cell culture (SILAC). *J Proteome Res* 2, 173-181

101. Van Hoof, D., Pinkse, M. W., Oostwaard, D. W., Mummery, C. L., Heck, A. J., and Krijgsveld, J. (2007) An experimental correction for arginine-to-proline conversion artifacts in SILAC-based quantitative proteomics. *Nature methods* 4, 677-678

102. Park, S. K., Venable, J. D., Xu, T., and Yates, J. R., 3rd (2008) A quantitative analysis software tool for mass spectrometry-based proteomics. *Nature methods* 5, 319-322

103.   Lossner, C., Warnken, U., Pscherer, A., and Schnolzer, M. (2011) Preventing arginine-to-proline conversion in a cell-line-independent manner during cell cultivation under stable isotope labeling by amino acids in cell culture (SILAC) conditions. *Analytical biochemistry* 412, 123-125

104.   Bicho, C. C., de Lima Alves, F., Chen, Z. A., Rappsilber, J., and Sawin, K. E. (2010) A genetic engineering solution to the "arginine conversion problem" in stable isotope labeling by amino acids in cell culture (SILAC). *Molecular & cellular proteomics : MCP* 9, 1567-1577

2. SILAC-based quantitative proteomics and phosphoproteomics in fission yeast (Manuscript I)

Alejandro Carpy, André Koch, Claudia C. Bicho, Weronika E. Borek, Silke Hauf, Kenneth E. Sawin, Boris Macek

This chapter presents manuscripts that have been accepted for publication as:

2.1 Abstract

Shot-gun proteomics combined with Stable Isotope Labeling by Amino acids in Cell culture (SILAC) is a powerful approach to quantify proteins and post-translational modifications across the entire proteome. Use of SILAC technology in *S. pombe* needs to cope with the "arginine-conversion problem", in which isotope-labeled arginine is converted to other amino acids. This can be circumvented by either employing stable isotope-marked lysine only (as opposed to the more standard lysine/arginine double labeling), or by using the power of yeast genetics to create strains that only very inefficiently convert arginine to proline. Both strategies have been employed successfully in *S. pombe* large-scale (phospho)proteomics projects. Here we present a (phospho)proteomic workflow based on in-solution proteome digestion of SILAC-labeled samples, phosphopeptide enrichment using strong cation exchange and TiO_2 chromatographies, followed by high accuracy mass spectrometry measurement on an Orbitrap mass spectrometer and subsequent bioinformatic processing using MaxQuant software.

2.2 Introduction

Shot-gun proteomics has emerged as state-of-the-art technology to globally identify and quantify proteins and post-translational modifications using liquid chromatography combined with mass spectrometry (LC-MS). This technique, in which the complete proteome is digested into peptides for analysis, is by itself semi-quantitative and needs to be combined with other chemical, biological or bioinformatic techniques to yield quantitative data (1-4)*.*

Although many shot-gun proteomics studies have focused on *Saccharomyces cerevisiae* (5-11), attention of the proteomics community towards *S. pombe* is increasing. Several

comprehensive proteomic and phosphoproteomic studies have recently been published (12-19). So far these studies have covered up to 3,542 *S. pombe* proteins (71 % of the predicted proteome) (19) and 1,544 phosphoproteins (17) and have examined the *S. pombe* proteome under various physiological conditions.

SILAC (Stable Isotope Labeling by Amino acids in Cell culture) is an elegant method for relative quantification of peptides and proteins from different samples (3, 20). SILAC involves metabolic labeling of cells with exogenous amino acids containing defined numbers of stable, non-radioactive, "heavy" isotopes (e.g. ^2H, ^{13}C and ^{15}N), thereby introducing a defined isotope label into newly synthesized cellular proteins. In parallel, a separate cell culture is grown in presence of the corresponding natural "light" amino acids. "Heavy" and "light" cultures can either correspond to two different treatments or growth conditions of the same strain, or two different strains altogether. Upon completion of the experiment, cells are lysed, and the "heavy" and "light" lysates are mixed in equal amounts according to cell count or protein content. The combined protein extract is then digested with an appropriate protease, and the resulting peptide mixture is measured in a mass spectrometer. In a SILAC MS measurement, peptide ions are detected as doublets, and the total signal ratio between the members of a doublet corresponds to the relative difference in the amount of the relevant protein or peptide (e.g. for a post-translational modification) between the two analyzed states (**Figure 1**). Because most of the amino acids used for labeling allow for synthesis of two different "heavy" forms, SILAC can be extended to three samples at once, often called "heavy", "medium-heavy" and "light"; in this case, peptide ions are detected as triplets.

For labeling cells, an ideal SILAC growth medium will contain both "heavy" arginine and "heavy" lysine. The reason is that trypsin, which is the standard protease for digestion of proteins (because it typically creates peptide sizes suitable for MS), cleaves polypeptide chains after arginine and lysine residues. Labeling with "heavy" arginine and lysine therefore ensures that all peptides contain at least one "heavy" amino acid after digestion and therefore are quantifiable (due to incomplete digestion, some peptides may contain more than one "heavy" amino acid). Because wild-type *S. pombe* can synthesize all amino acids, labeling cells with "heavy" amino acids requires the use of auxotrophic strains. Labeling fission yeast lysine auxotrophic cells with "heavy" lysine is straightforward and does not present any problems (17, 21). However, arginine is catabolized and converted into other amino acids, including but not limited to proline. As a result, "heavy" label can be incorporated into proteins at amino-acid residues other than arginine, leading to the appearance of additional peaks in MS and MS/MS spectra of a given peptide. This so-called "arginine-conversion problem" can seriously
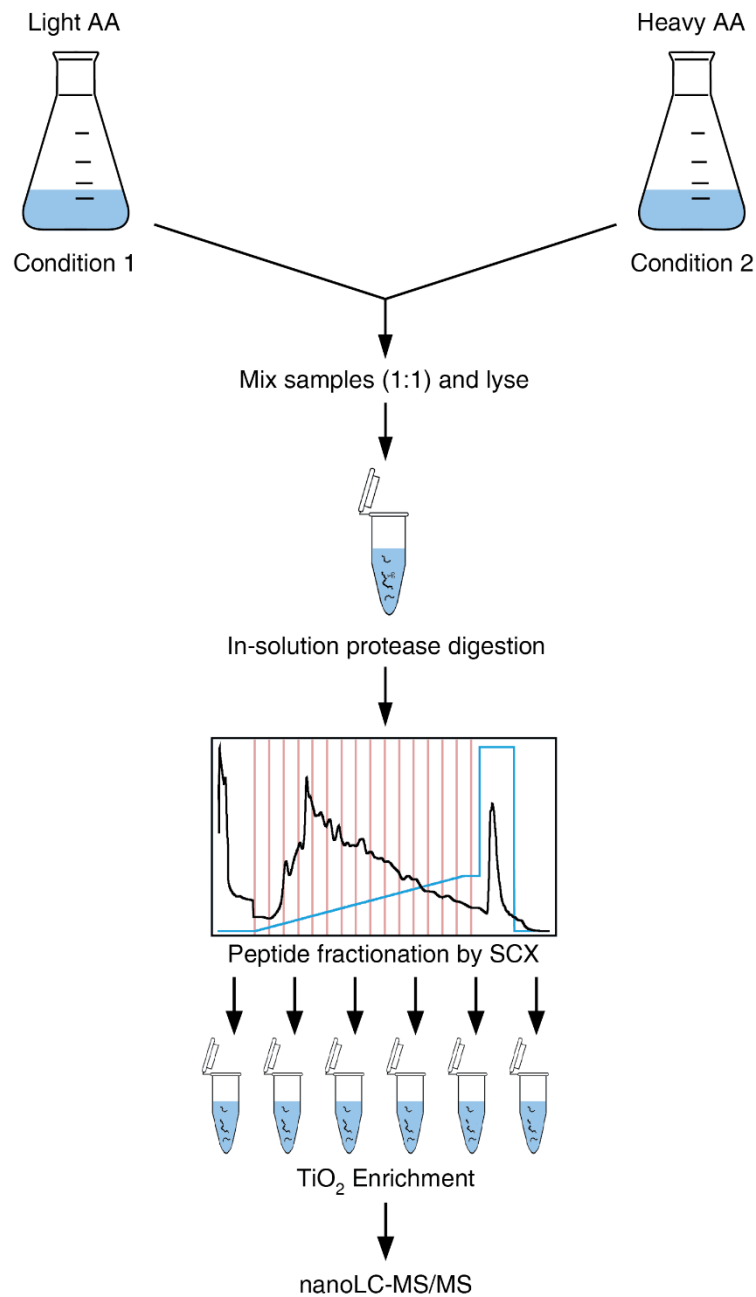
Figure 1. *SILAC-based phosphoproteomics workflow described in this protocol.*

confound peptide identification and quantification. In addition, certain auxotrophs in fission yeast, including lysine/arginine double auxotrophs, can show impaired growth even in the presence of the required amino acids. Therefore, SILAC experiments in fission yeast face two important issues: 1) avoiding catabolic conversion of labeled arginine to other amino acids; and 2) obtaining robust growth of auxotrophic strains used in labeling experiments.

In mammalian cells and budding yeast, arginine-conversion is minor, and several approaches have been developed to minimize the problem, including titration of arginine and proline concentrations and computational methods (20, 22-28). In fission yeast, arginine-conversion is much more extensive (16) and thus must be dealt with more decisively. Two general approaches are possible, and both have been used successfully (16, 17). The first approach (17) avoids "heavy" arginine altogether. Instead, cells are labeled only with "heavy" lysine, and, in place of trypsin, proteolytic digestion into peptides is performed using lysyl endopeptidase Lys-C, which cleaves after lysine residues. While this approach is technically straightforward, a potential disadvantage is that peptides produced by Lys-C digestion will be significantly longer than those produced by trypsin. This can affect "flyability" of peptides and thus identification in the mass spectrometer. The second approach (16) takes advantage of the power of yeast genetics and deletes genes that encode enzymes involved in arginine catabolism, thereby preventing or at least decreasing arginine-conversion. Deletion of the predicted ornithine transaminase gene *car2+* or simultaneous double-deletion of the redundant predicted arginase genes *car1+* and *aru1+* has been shown to decrease conversion to acceptably low levels (16). Because constructing single-deletion strains is easier, deletion of the *car2+* gene has generally been the preferred method.

Here we describe a protocol for constructing fission yeast strains suitable for SILAC MS, as well as conditions for robust growth, harvesting, and processing of cells (16, 17). We further describe a protocol for enrichment and analysis of protein Ser/Thr/Tyr phosphorylation in fission yeast using strong cation exchange (SCX) and TiO_2 chromatography, followed by LC-MS measurement. This strategy has been used in several large-scale phosphoproteomics studies in both eukaryotes (29, 30) and prokaryotes (31, 32). The protocol follows methods outlined in (33) **(Figure 1)** and presents a revised version of a previously published protocol (34). Parts of this protocol can be applied to enrichment and analysis of other post-translational modifications: for example, SCX can be used to enrich for N-terminal acetylation (35), and TiO_2 can be used to enrich for sugars containing negatively charged residues, such as sialic acid (36). In this chapter we focus on enrichment of phosphopeptides and note how the protocol can be adapted for proteome analysis. Other methods can be used if deeper proteome coverage is required (37, 38).

2.3   Method

In-solution protein digestion

1. Dissolve the protein sample in denaturation buffer at RT to a final concentration between 2 and 5 mg/mL.

This protocol is optimized for protein amounts between 2 and 10 mg. We recommend using at least 5 mg of protein to achieve good coverage of the phosphoproteome. Do not heat-up samples in the presence of urea; high concentration of urea will lead to carbamylation of free amino groups. Keep the salt concentration at a minimum during all steps, salt concentrations above 10 mM may interfere with SCX chromatography.

2. Add reduction buffer to a final concentration of 1 mM DTT and incubate under mild shaking conditions for 1 h at RT.

3. Add alkylation buffer to a final concentration of 5.5 mM IAA and incubate under mild shaking conditions for 1 h at RT in the dark.

4. Check pH using a pH indicator strip and adjust to 8.0 if necessary.
   Adjust the pH with a very low volume of 1M Tris-HCl, pH 8.0.

5. Add 1 µg of lysyl endopeptidase LysC per 100 µg of protein and incubate under mild shaking conditions for 3 h at RT.

6. Dilute the sample 1:5 with Milli-Q water.

7. Add 1 µg of trypsin per 100 µg of protein.
   If samples are only labeled with lysine, add another 1 µg of lysyl endopeptidase LysC per 100 µg of protein instead of trypsin.

8. Check pH and adjust to 8.0 if necessary. Incubate overnight (ca. 18 h) under mild shaking conditions at RT.
   The pH of the sample may change after digestion is completed, furthermore trypsin is stored in acidic conditions and the pH of the sample may change after trypsin has been added. Adjust the pH with a very low volume of 1M Tris-HCl, pH 8.0. At this point the protein digest can be directly measured by LC-MS (proteome measurement). In this case skip the "SCX Chromatography" and "TiO2 Chromatography" protocols and continue with "Liquid Chromatography".

SCX Chromatography

9. Quench the digestion reaction by adding trifluoroacetic acid (TFA) to a final concentration of 0.1% (vol/vol). Incubate for 10 min at RT while shaking.
   Make sure that the pH is <2.7. If necessary, add more TFA, but keep the TFA concentration as low as possible. Usually a maximum of 0.3% TFA is required.

10. Centrifuge the sample for 5 min at 3,000 g to remove precipitate. Transfer supernatant to a new tube and discard the precipitate.
    The precipitate contains mainly urea.

11. Load the protein digest onto an equilibrated SCX column with SCX solvent A at a flow rate of 1 mL/min.

If salt concentration is higher than 10 mM dilute sample with Milli-Q water. Conductivity during sample loading should be under 4 mS/cm since higher conductivity can interfere with peptide binding; if higher conductivity is observed flow-through should be diluted with Milli-Q water and reloaded onto the column.

12. Collect the flow-through.

Collection and subsequent phosphopeptide enrichment and analysis of the flow-through is particularly important since multiply phosphorylated peptides will not bind to the SCX column.

13. Elute the bound peptides with a linear gradient of 0 – 30% SCX solvent B over 30 minutes at a flow rate of 1 mL/min, collect the eluent in 2 mL fractions.

TiO$_2$ Chromatography

14. Each SCX fraction, including the flow-through, is incubated with separate batches of TiO$_2$ beads and will be referred as "FT" (Flow-Through) or "Fraction" in the text.

15. Desalt peptides from FT using reversed-phase Sep-Pak C18 cartridges.

The FT fraction contains a high amount of Urea. Each cartridge has a binding capacity of around 18 mg of protein.

16. Lyophilize the FT and each fraction separately. Dried peptides can be stored at −80 °C for several weeks.

Drying samples for phosphopeptide enrichment for longer periods in the vacuum centrifuge is not recommended. Vacuum centrifuging usually involves increase in temperature, which could affect some labile phosphorylation events. In addition, drying to completeness usually leads to low recovery of (phospho)peptides from the walls of centrifuge tubes.

17. Re-dissolve the FT in 10 mL of loading solution and the fractions in 2 mL of loading solution.

Alternatively the fractions can be adjusted to a final concentration of 80% ACN and 6% trifluoroacetic acid by adding ACN and TFA directly to the FT and the fractions. The fractions (not the FT) contain 30% ACN so a final concentration of 80% ACN and 6% TFA can be achieved by adding 7400 µL ACN and 600 µL TFA per fraction.

18. Weigh TiO$_2$ beads into a dedicated tube; for each round of enrichment weigh 5 mg of beads.

We recommend performing five successive rounds of enrichment for the FT and three for the first 5 SCX fractions. These fractions usually contain the highest amount of phosphopeptides and performing serial enrichment of these samples will increase the amount of phosphopeptides identified.

19. Wash TiO$_2$ beads by adding 50 µL of washing solution per 5 mg of TiO$_2$ beads, centrifuge and discard supernatant.

20. Resuspend beads with 50 µL of washing solution per 5 mg of TiO$_2$ beads.

21. Spike 50 µL of TiO$_2$ beads in washing solution separately into each sample.

22. Mix 30 min at RT using an orbital shaker.

23. Centrifuge at 10,000 g for 2 min then remove and keep the supernatant.

24. For successive rounds of enrichment repeat steps 7 – 9 with the supernatant as required.

    Each successive round of enrichment will generate a new bead sample and should be treated separately.

25. Wash beads twice with 1 mL wash solution, incubate with shaking in a Thermomixer at 850 rpm at RT for 10 min. Centrifuge at 10,000 g for 2 min and discard supernatant.

26. Prepare one C8 microcolumn ("stage-tip") for each sample by placing a ~1 mm^2 piece of Empore C$_8$ material into a 200 µl (yellow) pipette tip, as described previously (39).

    C8 microcolumns can also be commercially obtained (SP301, Thermo Fisher Scientific).

27. Resuspend beads in 50 µL of wash solution and transfer to a C8 microcolumn.

28. Pack column at 2,000 g for 3 min and discard flow-through.

29. Elute each sample with 3 x 100 µL of elution solution.

30. Dry the eluates to a final volume of 100 µL in a vacuum centrifuge.

31. Add formic acid to a final concentration of 1%.

32. Check pH, should be <2.7.

    Samples can be further dried to a final volume of 5 µL and measured directly, which may rescue very hydrophobic phosphopeptides.

33. Desalt peptides as described previously (39).

Liquid chromatography

34. Pack the nano-emitter with 15 cm of reversed-phase C18 with methanol at a constant helium pressure of 80 bar using a pressure injection cell.

    Alternatively pre-packed columns can be acquired commercially, e.g. PicoFrit column 360 µm OD x 75 µm ID, 8 µm Tip ID, 15 cm length packed with Reprosil-PUR C18 AQ, 3 µm, 120 Å (PF7508-150H354, New Objective).

35. Set the system with a liquid junction, with no pre-column and no split by connecting the column (Packed Emitter) to a PEEK MicroCross. Attach to the MicroCoss the "waste in" and "column out" lines as well as a platine electrode to supply the spray voltage.

    For a more detailed protocol please refer to Schieltz et al. (40).

36. Elute samples from StageTips with 50 µL HPLC solvent B and dry in the vacuum centrifuge to a final volume of 5 µL.

    If sample is <5 µL adjust volume with HPLC solvent A. Do not let the sample dry and do not use temperature while drying.

37. Add 1 µL of HPLC loading solvent to the sample.

38. Load 5 µL of the phosphopeptide sample onto the column at a fixed flow rate of 0.7 µL/min, limiting pressure to 280 bar using the Intelliflow option.

    This method was developed with an HPLC. A UHPLC can also be used and the parameters should be adjusted accordingly.

39. Equilibrate the analytical column.

40. Elute the peptides using a 120 min linear gradient from 5 – 30 % solvent B at a flow of 200 nL/min.

41. Wash-out hydrophobic peptides by linearly increasing the concentration of solvent B to 50 % over 10 min.

42. Wash the column with 90 % solvent B for 5 min.

43. Equilibrate column with 5 % solvent B before next sample is measured.

Mass spectrometry

44. Samples are analyzed using liquid chromatography tandem mass spectrometry (LC-MS/MS). This protocol has been performed using an Easy-nLC II connected to an LTQ Orbitrap Elite (Thermo Fisher Scientific) using a nano-electrospray source.

    To perform this type of analysis, any nanoflow LC connected to a high resolution mass spectrometer can be used.

45. Operate the mass spectrometer using data-dependent acquisition mode. In this mode a defined number of precursor ions from the survey scan or full MS are selected for further MS/MS analysis.

46. Analyze the samples using the parameters described in table 1.

    Two data-dependent methods are presented here. For high-accuracy and full mass range measurements of fragment ions the HCD method is recommended. The CID MSA method yields a fast-scanning and high-sensitivity but low resolution measurements of fragment ions. We prefer using the HCD method which usually yields better spectra for modified peptides and facilitates localization of phosphorylation sites due to high-accuracy. We recommend to use the CID method in older versions of the Orbitrap instruments (e.g. Orbitrap Classic and Orbitrap XL) since HCD sequencing is considerably slower than CID in these instruments.

Data processing and Analysis

47. We routinely use the MaxQuant software suite (41-44) to process phosphoproteomics data. Download and install the latest MaxQuant version ensuring that the software and hardware requirements are fulfilled.

Alternatively there are several freely or commercially available software packages that can manage phosphoproteomics data (e.g. Mascot, Proteome Discoverer, Scaffold, Trans-Proteomic Pipeline, etc.).

48. Load all the files into MaxQuant. Use predefined parameters unless stated otherwise (see below).

    Make sure that all the files are located in one folder. If many files are going to be loaded the load file option can be used.

49. Create an experimental template after loading the files by pressing the "Write Template" icon.

    A file named experimentalDesignTemplate.txt will be created in a subfolder called "combined" in the folder containing your raw files.

50. Open experimentalDesignTemplate.txt using any txt reader (eg. Excel) and define your experiment accordingly.

    As many experiments as required can be defined. Our recommendation is to group all the samples corresponding to one biological replicate into one experiment and to define replicates as separate experiments (eg. BiologicalReplicate1 and BiologicalReplicate2). If proteome measurements were performed in parallel define these also separately.

51. Load the experimental template using the "Read from" file icon.

52. In the "Group-specific parameters" Tab define multiplicity and the used labels, add Phospho(STY) as variable modification and specify the used enzyme.

    Multiplicity will be depending on the number of channels used in the experiment, in a double-SILAC experiment and triple-SILAC experiment the multiplicity will be 2 and 3, respectively. If proteome only samples were measured an extra experimental group can be defined in the "Raw files" tab – a new group tab will appear under "group specific parameters" – search this files without Phospho(STY) as variable modification, all other parameters will be the same.

53. Under the "Global parameters" tab add the fission yeast fasta file for the database search.

    The fasta file can be downloaded from online resources such as PomBase (www.pombase.org) or UniProt (www.uniprot.org). The exact version and date of download should be recorded - this information will be required when the data is published.

54. Define the number of threads and start the processing.

    The number of threads should be equal or lower than the number of cores available, although at least 2 GB per thread are recommended.

55. Once the processing is completed, a folder named "txt" will be created in the "combined" folder. This folder contains the result files in txt format which can be opened and analyzed using Excel.

The "txt" folder also contains a file named tables.pdf which contains an extensive description of the files generated in this folder as well as the columns contained in each table.

2.4   Contributions

This protocol is part of a book chapter in a laboratory manual on fission yeast published by Cold Spring Harbor Laboratory Press in 2014. André Koch, Claudia C. Bicho, Weronika E. Borek, Silke Hauf, and Kenneth E. Sawin were mostly involved in the other components of the chapter and participated with comments, suggestions and helped writing the introduction. Boris Macek developed the original protocol and together with me wrote the protocol. I optimized and established the protocol presented here. My contribution to this manuscript was about 70%.

2.5 References

1. Luber, C. A., Cox, J., Lauterbach, H., Fancke, B., Selbach, M., Tschopp, J., Akira, S., Wiegand, M., Hochrein, H., O'Keeffe, M., and Mann, M. (2010) Quantitative Proteomics Reveals Subset-Specific Viral Recognition in Dendritic Cells. *Immunity* 32, 279-289

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

3. 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

4. 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. *Anal. Chem.* 75, 1895-1904

5. 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. *Nat. Biotechnol.* 17, 994-999

6. Ho, Y., et al. (2002) Systematic identification of protein complexes in Saccharomyces cerevisiae by mass spectrometry. *Nature* 415, 180-183

7. Ghaemmaghami, S., Huh, W., Bower, K., Howson, R. W., Belle, A., Dephoure, N., O'Shea, E. K., and Weissman, J. S. (2003) Global analysis of protein expression in yeast. *Nature* 425, 737-741

8. Ross, P. L., Huang, Y. L. 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. *Mol. Cell. Proteomics* 3, 1154-1169

9. de Godoy, L. M. F., Olsen, J. V., Cox, J., Nielsen, M. L., Hubner, N. C., Froehlich, F., Walther, T. C., and Mann, M. (2008) Comprehensive mass-spectrometry-based proteome quantification of haploid versus diploid yeast. *Nature* 455, 1251-U1260

10. Flory, M. R., Lee, H., Bonneau, R., Mallick, P., Serikawa, K., Morris, D. R., and Aebersold, R. (2006) Quantitative proteomic analysis of the budding yeast cell cycle using acid-cleavable isotope-coded affinity tag reagents. *Proteomics* 6, 6146-6157

11. Breitkreutz, A., Choi, H., Sharom, J. R., Boucher, L., Neduva, V., Larsen, B., Lin, Z. Y., Breitkreutz, B. J., Stark, C., Liu, G. M., Ahn, J., Dewar-Darch, D., Reguly, T., Tang, X. J., Almeida, R., Qin, Z. S., Pawson, T., Gingras, A. C., Nesvizhskii, A. I., and Tyers, M. (2010)

A Global Protein Kinase and Phosphatase Interaction Network in Yeast. *Science* 328, 1043-1046

12. Matsuyama, A., Arai, R., Yashiroda, Y., Shirai, A., Kamata, A., Sekido, S., Kobayashi, Y., Hashimoto, A., Hamamoto, M., Hiraoka, Y., Horinouchi, S., and Yoshida, M. (2006) ORFeome cloning and global analysis of protein localization in the fission yeast Schizosaccharomyces pombe. *Nat Biotech* 24, 841-847

13. Schmidt, M. W., Houseman, A., Ivanov, A. R., and Wolf, D. A. (2007) Comparative proteomic and transcriptomic profiling of the fission yeast Schizosaccharomyces pombe. *Mol. Syst. Biol.* 3, 79

14. Wilson-Grady, J. T., Villen, J., and Gygi, S. P. (2008) Phosphoproteome analysis of fission yeast. *J. Proteome Res.* 7, 1088-1097

15. Beltrao, P., Trinidad, J. C., Fiedler, D., Roguev, A., Lim, W. A., Shokat, K. M., Burlingame, A. L., and Krogan, N. J. (2009) Evolution of phosphoregulation: comparison of phosphorylation patterns across yeast species. *PLoS Biol.* 7, e1000134

16. Bicho, C. C., de Lima Alves, F., Chen, Z. A., Rappsilber, J., and Sawin, K. E. (2010) A Genetic Engineering Solution to the "Arginine Conversion Problem" in Stable Isotope Labeling by Amino Acids in Cell Culture (SILAC). *Mol. Cell. Proteomics* 9, 1567-1577

17. Koch, A., Krug, K., Pengelley, S., Macek, B., and Hauf, S. (2011) Mitotic Substrates of the Kinase Aurora with Roles in Chromatin Regulation Identified Through Quantitative Phosphoproteomics of Fission Yeast. *Sci. Signal.* 4, rs6

18. Marguerat, S., Schmidt, A., Codlin, S., Chen, W., Aebersold, R., and Bähler, J. (2012) Quantitative Analysis of Fission Yeast Transcriptomes and Proteomes in Proliferating and Quiescent Cells. *Cell* 151, 671-683

19. Gunaratne, J., Schmidt, A., Quandt, A., Neo, S. P., Sarac, O. S., Gracia, T., Loguercio, S., Ahrne, E., Li Hai Xia, R., Tan, K. H., Loessner, C., Bahler, J., Beyer, A., Blackstock, W., and Aebersold, R. (2013) Extensive Mass Spectrometry-Based Analysis of the Fission Yeast Proteome: The S. pombe PeptideAtlas. *Mol. Cell. Proteomics* 12, 1741-1751

20. Ong, S. E., and Mann, M. (2006) A practical recipe for stable isotope labeling by amino acids in cell culture (SILAC). *Nat. Protoc.* 1, 2650-2660

21. Wang, H., Tipthara, P., Zhu, L., Poon, S. Y., Tang, K., and Liu, J. (2012) Chromatin-Associated Proteins Revealed by SILAC-Proteomic Analysis Exhibit a High Likelihood of Requirement for Growth Fitness under DNA Damage Stress. *Int J Proteomics* 2012, 630409

22. Gruhler, A., Olsen, J. V., Mohammed, S., Mortensen, P., Faergeman, N. J., Mann, M., and Jensen, O. N. (2005) Quantitative phosphoproteomics applied to the yeast pheromone signaling pathway. *Mol. Cell. Proteomics* 4, 310-327

23. Mousson, F., Kolkman, A., Pijnappel, W. W., Timmers, H. T., and Heck, A. J. (2008) Quantitative proteomics reveals regulation of dynamic components within TATA-binding protein (TBP) transcription complexes. *Mol. Cell. Proteomics* 7, 845-852

24. Park, S. K., Liao, L., Kim, J. Y., and Yates, J. R., 3rd (2009) A computational approach to correct arginine-to-proline conversion in quantitative proteomics. *Nature methods* 6, 184-185

25. Blagoev, B., and Mann, M. (2006) Quantitative proteomics to study mitogen-activated protein kinases. *Methods* 40, 243-250

26. Graumann, J., Hubner, N. C., Kim, J. B., Ko, K., Moser, M., Kumar, C., Cox, J., Scholer, H., and Mann, M. (2008) Stable isotope labeling by amino acids in cell culture (SILAC) and proteome quantitation of mouse embryonic stem cells to a depth of 5,111 proteins. *Mol. Cell. Proteomics* 7, 672-683

27. Prokhorova, T. A., Rigbolt, K. T., Johansen, P. T., Henningsen, J., Kratchmarova, I., Kassem, M., and Blagoev, B. (2009) Stable isotope labeling by amino acids in cell culture (SILAC) and quantitative comparison of the membrane proteomes of self-renewing and differentiating human embryonic stem cells. *Mol. Cell. Proteomics* 8, 959-970

28. Bendall, S. C., Hughes, C., Stewart, M. H., Doble, B., Bhatia, M., and Lajoie, G. A. (2008) Prevention of amino acid conversion in SILAC experiments with embryonic stem cells. *Molecular & cellular proteomics : MCP* 7, 1587-1597

29. Olsen, J. V., Blagoev, B., Gnad, F., Macek, B., Kumar, C., Mortensen, P., and Mann, M. (2006) Global, in vivo, and site-specific phosphorylation dynamics in signaling networks. *Cell* 127, 635-648

30. 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

31. Macek, B., Gnad, F., Soufi, B., Kumar, C., Olsen, J. V., Mijakovic, I., and Mann, M. (2008) Phosphoproteome analysis of E. coli reveals evolutionary conservation of bacterial Ser/Thr/Tyr phosphorylation. *Mol. Cell. Proteomics* 7, 299-307

32. Macek, B., Mijakovic, I., Olsen, J. V., Gnad, F., Kumar, C., Jensen, P. R., and Mann, M. (2007) The serine/threonine/tyrosine phosphoproteome of the model bacterium Bacillus subtilis. *Mol. Cell. Proteomics* 6, 697-707

33. Macek, B., Mann, M., and Olsen, J. V. (2009) Global and site-specific quantitative phosphoproteomics: principles and applications. *Annu. Rev. Pharmacol. Toxicol.* 49, 199-221

34. Olsen, J., and Macek, B. (2009) High Accuracy Mass Spectrometry in Large-Scale Analysis of Protein Phosphorylation. In: Lipton, M., and Paša-Tolic, L., eds. *Mass Spectrometry of Proteins and Peptides*, pp. 131-142, Humana Press

35. Helbig, A. O., Gauci, S., Raijmakers, R., van Breukelen, B., Slijper, M., Mohammed, S., and Heck, A. J. (2010) Profiling of N-acetylated protein termini provides in-depth insights into the N-terminal nature of the proteome. *Mol. Cell. Proteomics* 9, 928-939

36. Larsen, M. R., Jensen, S. S., Jakobsen, L. A., and Heegaard, N. H. (2007) Exploring the sialiome using titanium dioxide chromatography and mass spectrometry. *Mol. Cell. Proteomics* 6, 1778-1787

37. Shevchenko, A., Wilm, M., Vorm, O., and Mann, M. (1996) Mass spectrometric sequencing of proteins silver-stained polyacrylamide gels. *Anal. Chem.* 68, 850-858

38. Hubner, N. C., Ren, S., and Mann, M. (2008) Peptide separation with immobilized pI strips is an attractive alternative to in-gel protein digestion for proteome analysis. *Proteomics* 8, 4862-4872

39. Rappsilber, J., Mann, M., and Ishihama, Y. (2007) Protocol for micro-purification, enrichment, pre-fractionation and storage of peptides for proteomics using StageTips. *Nat. Protocols* 2, 1896-1906

40. Schieltz, D. M., Washburn, M. P., and Hays, L. G. (2006) Analysis of Complex Protein Mixtures Using Nano-LC Coupled to MS/MS. *Cold Spring Harb. Protoc.* 2006, pdb.prot4553

41. 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 Biotech* 26, 1367-1372

42. Cox, J. r., Neuhauser, N., Michalski, A., Scheltema, R. A., Olsen, J. V., and Mann, M. (2011) Andromeda: A Peptide Search Engine Integrated into the MaxQuant Environment. *Journal of Proteome Research* 10, 1794-1805

43. Cox, J., Matic, I., Hilger, M., Nagaraj, N., Selbach, M., Olsen, J. V., and Mann, M. (2009) A practical guide to the MaxQuant computational platform for SILAC-based quantitative proteomics. *Nat. Protoc.* 4, 698-705

44. Kirchner, M., and Selbach, M. (2012) In vivo quantitative proteome profiling: planning and evaluation of SILAC experiments. *Methods Mol. Biol.* 893, 175-199

3. Absolute proteome and phosphoproteome dynamics during the cell cycle of fission yeast (Manuscript II)

Alejandro Carpy, Karsten Krug, Sabine Graf, André Koch, Sasa Popic, Silke Hauf, Boris Macek

3.1 Abstract

To quantify cell cycle-dependent fluctuations on a proteome-wide scale, we performed integrative analysis of the proteome and phosphoproteome during the four major phases of the cell cycle in *S. pombe*. In highly synchronized cells we identified 3,753 proteins and 3,682 phosphorylation events and relatively quantified 65 % of the data across all phases. Quantitative changes during the cell cycle were infrequent and weak in the proteome, but prominent in the phosphoproteome. Protein phosphorylation peaked in mitosis, where the median phosphorylation site occupancy was 44 %, about two-fold higher than in other phases. We measured copy numbers of 3,178 proteins, which together with phosphorylation site stoichiometry enabled us to estimate the absolute amount of protein-bound phosphate, as well as its change across the cell cycle. Our results indicate that 23 % of the average intracellular ATP is utilized by protein kinases to phosphorylate their substrates in order to drive regulatory processes during cell division. Accordingly, we observe that phosphate transporters and phosphate-metabolizing enzymes are phosphorylated and therefore likely to be regulated in mitosis.

3.2 Introduction

Cell replication involves a complex series of highly regulated and evolutionary conserved events, called the 'cell cycle'. Aberrations in the cell cycle have severe implications and can cause cancerous growth. A detailed understanding of the cell cycle and its regulation may identify additional targets for cancer therapy (1-3). The cell cycle has been the subject of previous proteomics studies. Olsen et al. (4) measured the dynamics of thousands of proteins and phosphorylation events across cell cycle phases of HeLa cells, providing insights into the underlying regulatory mechanisms and pointing to a general increase in phosphorylation site occupancy during M phase. In a targeted study, Pagliuca et al. (5) investigated interactors of cyclins E1, A2 and B1 in HeLa cells, revealing key mechanistic links between DNA replication and mitosis.

Schizosaccharomyces pombe (fission yeast) is a unicellular organism, which can be easily genetically manipulated and carries many cell cycle features similar to metazoan cells. It is an important model organism to study the cell cycle and its checkpoint controls (6). Recent global proteomics studies of yeasts and their cell cycle (7-13) have mainly focused on *Saccharomyces cerevisiae* (budding yeast), with only few studies of fission yeast (14, 15), although the fission yeast cell cycle may be more representative of eukaryotic cell cycles (16). However, attention of the proteomics community towards *S. pombe* is increasing. Recent proteomics studies covered up to 4,087 *S. pombe* proteins (71 % of the predicted proteome) and 1,544 phosphoproteins in both asynchronous and synchronized cell cultures (17-22); however, a comprehensive analysis of the *S. pombe* cell cycle is so far missing.

Here we use high-resolution mass spectrometry in combination with a stable isotope labeling by amino acids in cell culture (SILAC) method, termed super-SILAC (23), and intensity-based absolute quantification (iBAQ) (24) to measure relative and absolute dynamics of the proteome and phosphoproteome during the cell cycle of fission yeast. We estimate copy numbers for 3,178 *S. pombe* proteins and combine these data with calculated phosphorylation site stoichiometry to estimate the total amount of protein-bound phosphate and its dynamics across the cell cycle. Providing the global absolute dynamics and stoichiometry of proteins and their modifications will be a valuable resource for classical and systems biologists alike.

3.3   Materials and methods

Overview of the experimental design

SILAC-labeled cells were synchronized using the temperature sensitive *cdc25-22* allele (25). Briefly, cells were grown at 25°C (permissive temperature) to a cell density of 6-7 x 10^6 cells/mL and arrested at late G2 phase by shifting to 36°C (restrictive temperature) for 5 hrs. Cells were released into the cell cycle by shifting the temperature back to 25°C. The Lys0-labeled cells were harvested at 17 min (M phase) and 32 min (G1 phase), whereas the Lys8-labeled cells were harvested at 0 min (G2 phase; without shifting back to 25 °C) and 50 min (S phase). To perform relative quantification of the four cell cycle phases we generated a super-SILAC standard (26) by mixing an equal number of cells from the G2, M, G1 and S phases labeled with medium-heavy lysine (Lys4). We added this standard to equal numbers of cells from M (Lys0) and S (Lys8); and, separately, from G1 (Lys0) and G2 (Lys8), resulting in two triple-SILAC samples **(Figure 1)**. Each sample was digested with endoproteinase Lys-C followed by fractionation via isoelectric focusing (Off-Gel) and strong anion exchange chromatography (SAX). Phosphopeptides were enriched using SCX followed by TiO_2 chromatography. The resulting fractions were measured on Orbitrap Elite and Orbitrap XL mass spectrometers and quantified by MaxQuant software at the proteome and phosphoproteome level.
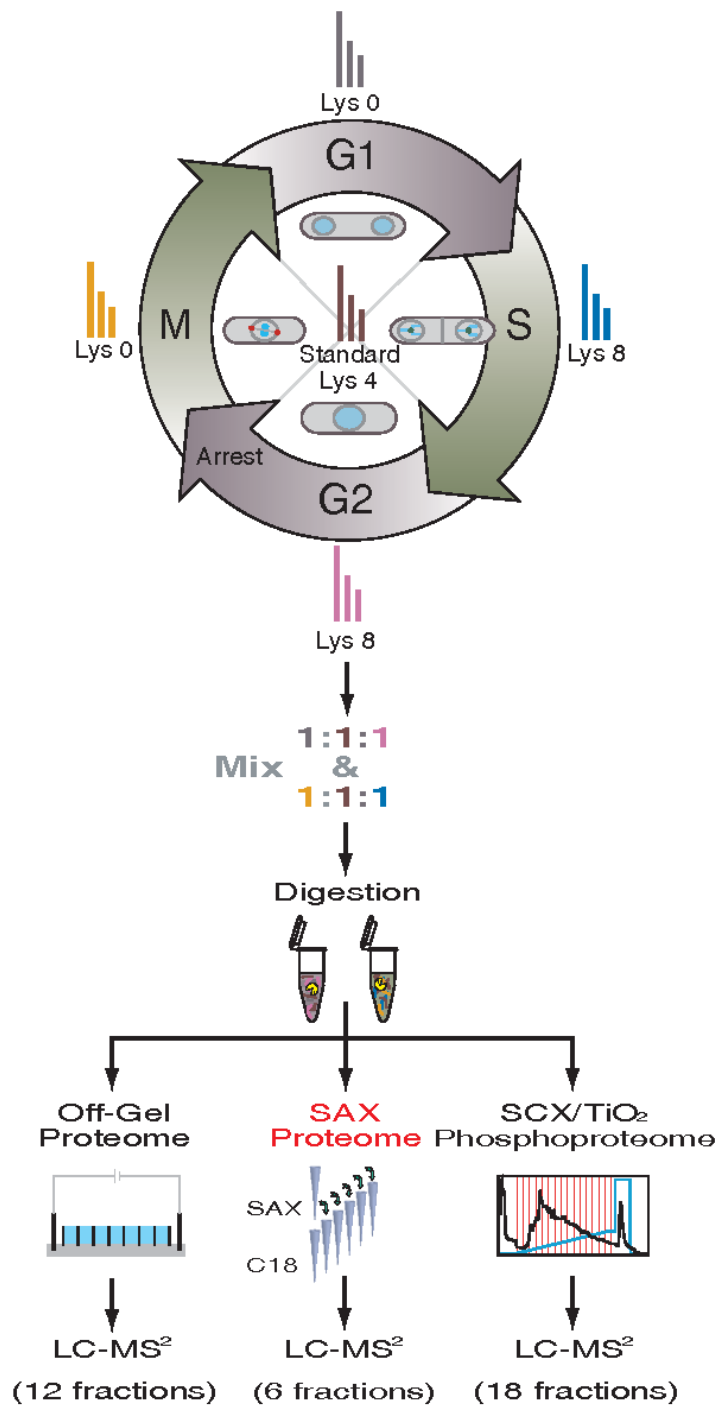
Figure 1. ***Relative and absolute dynamics of the proteome and phosphoproteome during four phases of the fission yeast cell cycle.*** *Four separate cell cultures of S. pombe representing the cell cycle phases and four separate cell cultures for the super-SILAC standard were labeled using SILAC media containing either Lys0, Lys4 or Lys8. The cultures were arrested in late G2 phase using the temperature-sensitive cdc25-22 allele. Reduction to permissive temperature allowed the cells to synchronously enter mitosis and synchronously continue the cell cycle. The G2 phase sample was collected by harvesting the cells after 5 hr of incubation at 36°C. Samples of M, G1 and S phase were collected 17 min, 32 min and 50 min after release to 25 ºC, respectively. M phase (Lys0) and S phase (Lys8) were mixed with the super-SILAC (Lys4) experiment before protein digestion, as were G1 phase (Lys0) and G2 phase (Lys8). Cells were mixed in a 1:1:1 ratio based on cell count. The two experiments were digested with Lys-C and either fractionated at the peptide level (Off-Gel or SAX) or enriched for phosphopeptides using SCX followed by TiO₂ enrichment. All samples were measured using LC-MS/MS.*

56

S. pombe strains

We used SM205 h+ lys1 leu1-32::ura4+ plo1+-mCherry<<natR ura4::EGFP-pcn1+ cdc25-22 and SM222 h+ uds1-GFP:kanMX6 leu1 plo1+-mCherry<<natR cdc25-22.

SILAC

The amino acid derivatives 4,4,5,5-D4-L-lysine (Lys4, "medium-heavy lysine", K4) and $^{13}C_6^{15}N_2$-L-lysine (Lys8, "heavy lysine", K8) (both: Cambridge Isotope Laboratories) were used. Cells were cultured in K4 or K8 containing minimal medium for approximately 8-10 cell cycles in order to ensure sufficient incorporation of the labeled amino acids into proteins.

Cell harvesting

We harvested 1.26×10^9 cells at each cell cycle phase. Cells were pelleted by centrifugation (3000 g, 2 min, centrifuge at 40°C for G2 samples, or 4 °C for samples from other phases). All tubes and reagents were kept at 37 °C for G2 samples and at 4 °C for samples from other phases. After washing with H_2O and decanting of the supernatant, a small amount of residual liquid remained, in which cells were resuspended. The cell suspension was dripped into liquid nitrogen (lN_2) so that cells were shock-frozen immediately. Residuals of the cell suspension were taken up in 1 mL cold methanol (-20°C) for microscopic analysis. The cell pellets resulting from freezing in lN_2 were stored at -80°C until protein extraction was performed. At the same time as harvesting, cells from 10 mL cell culture were pelleted to prepare a TCA (trichloroacetic acid) protein extract that was analyzed by immunoblotting.

Cell disruption and protein extraction

For large-scale preparation of fission yeast protein extracts, cell disruption was performed using a ball mill grinder (Mixer Mill MM 400, Retsch) under cryogenic conditions. Frozen cell droplets of light (K0) G1 sample and heavy (K8) labeled G2 sample, as well as cell droplets from light (K0) M sample and heavy (K8) labeled S phase sample were mixed according to their cell number in a 1:1 ratio. For the SILAC standard, cell droplets from medium (K4) labeled G2, M, G1 and S phase samples were mixed according to their cell number in a 1:1:1:1 ratio. Droplets were filled into a lN_2-cooled 50 mL metal bucket with 1 metal ball (Ø 25 mm) and subjected to 5 ball mill runs (3 min at frequency 30/sec) with cooling in lN_2 between runs. The resulting powder was stored at -80°C. The cell powder was resuspended in denaturation buffer (6 M urea, 2 M thiourea, 1 % (w/v) n-octylglucoside (Roche), 10 mM Tris pH 8.0) to a final protein concentration of approximately 10 mg/mL. The suspension was incubated for about 1 hr at 4°C while shaking before insoluble material was removed twice by centrifugation (7197 g, 10 min, 4°C). The supernatant was aliquoted and stored at -80°C. This procedure yielded

approximately 130 mg and 60 mg of protein for the SILAC standard and M/S – G1/G2 samples respectively.

Microscopy

Methanol-fixed cells were washed with PEM/methanol ((100 mM PIPES, 1 mM EGTA, 1 mM MgSO$_4$; pH 6.8 with KOH)/50 % methanol) and with PEM. For DNA staining, cells were resuspended in 100 μL PEM containing 2 µg/mL DAPI (Sigma) and incubated at room temperature for at least 10 min. Cells were pelleted and 4 μL of the pellet were mounted. Images were acquired on an Axio Imager.M1 microscope (Zeiss) with a Plan-Apochromat 63x/1.4 oil objective (Zeiss) driven by a Piezo Z motor. Images were recorded with a charged-coupled device camera (CoolSnap EZ, Roper) and processed using the software MetaMorph 7.6.0.0 (Molecular Devices). 14 individual planes spaced by 0.3 μm were acquired. In order to create color-combined pictures, out-of focus planes were removed, 10 sections of each the GFP and mCherry channels were projected (maximum intensity projection), and signal intensity was adjusted for all pictures in a similar way. For the DAPI channel, only one of the in-focus planes was chosen for the color combined picture.

Immunoblots

Proteins were separated by SDS-PAGE and transferred onto PVDF membranes (Immobilon-P, Millipore) using a semi-dry transfer system (Amersham). Primary antibodies are listed in **Supplementary Table 1**. Horseradish peroxidase (HRP)-coupled secondary antibodies were detected by chemiluminescence.

Trichloroacetic acid extraction

4×10^7 cells were harvested by centrifugation (3000 g, 1 min, RT) and resuspended in 1 mL cold 20 % (v/v) TCA. Cells were pelleted by centrifugation (775 g, 2 min, RT). The supernatant was removed and cells were washed with 1 mL 1 M Tris and pelleted again. The cell pellet was resuspended in 100 – 200 μL 2x SDS sample buffer (125 mM Tris pH 6.8, 4 % (w/v) SDS, 20 % glycerol, 200 mM DTT, 0.02 % (w/v) bromophenol blue) and boiled at 95°C for 10 min. 600 μL glass beads (500 μm) were added and the samples were vortexed strongly (FastPrep™120 BIO101, Thermo Electron Corporation; 3 x 40 sec, level 5, with a 3 min break in between the individual runs). A hole was made in the bottom of the tube with a needle and the cell extract was eluted in a new reaction tube by centrifugation (60 sec, 956 g, 4°C) thereby separating the extract from the glass beads. The resulting samples were boiled again at 95°C for 10 min and centrifuged (16000 g, 10 min, RT) in order to pellet the cell debris.

Intracellular ATP quantification

Cells were disrupted using a ball mill grinder (Mixer Mill MM 400, Retsch). ATP was extracted using 5 % (v/v) TCA and measured using the ENLITEN Kit (Promega) according to the manufacturer's instructions. After ATP extraction, samples were neutralized with 20 mM sodium acetate buffer and diluted 50-fold and 1000-fold with ATP-free water. The luciferase reagent was added and luminescence was recorded for 10 s after a 2.5 s wait. A standard curve, using ATP included in the kit, was used to transform relative luminescence units (RLU) into ATP concentration. The amount of cells measured in each sample was used to calculate ATP molecules per cell. Because the *cdc25-22*-synchronized cells have a cell volume that is 3.51-times larger than unsynchronized cells (27), the resulting value was divided by 3.51.

In-solution digestion

The SILAC standard was added to the G1/G2 and M/S samples before digestion. For absolute quantification, the Proteomics Dynamic Range Standard set (UPS2, Sigma-Aldrich) was spiked in before protein in-solution digestion (24). 8 mg of protein extract dissolved in denaturation buffer (6 M urea, 2 M thiourea, 1 % (w/v) n-octylglucoside (Roche)) were digested in-solution (28). Briefly, proteins were reduced with 1 mM dithiothreitol, alkylated with 5.5 mM iodoacetamide, digested at RT for 3 hr with 1:100 (w/w) endoproteinase Lys-C (high urea concentration). Sample was diluted with 4 volumes of water before digestion over night at RT with 1:100 (w/w) endoproteinase Lys-C or 1:100 (w/w) Trypsin (low urea concentration). The protein digest was acidified using TFA to a final concentration of 0.1 % (v/v) to stop the protein digestion. In the case of samples fractionated using the Off-Gel fractionation no TFA was added.

Lysine Acetylation Enrichment

Lysine-acetylated peptides were enriched by immunoprecipitation using anti-AcK antibodies as described before with some modifications (29). Peptides were desalted using solid phase extraction before being dissolved in IP buffer (50 mM MOPS pH 7.2, 10 mM Sodium phosphate and 50 mM Sodium chloride). Peptides were incubated with agarose conjugated anti-acetyllysine antibody (Immunechem) overnight at 4°C on a rotation wheel. The immunoprecipitates were washed four times with IP buffer and two times with water before being eluted three times with 0.1% TFA in water. The eluted peptides were purified using Stage-Tips (30).

Phosphopeptide enrichment

After protein digestion, 7.5 mg of the samples was enriched for phosphopeptides using strong cation exchange chromatography (SCX) followed by TiO$_2$ chromatography as described before

(31) with minor modifications. Samples were injected using SCX Buffer A (5 mM KH_2PO_4, 30% acetonitrile (ACN), pH 2.7) and separated with a gradient from 0 to 35 % of SCX Buffer B (350 mM KCl, 5 mM KH_2PO_4, 30% ACN, pH 2.7). The resulting 15 fractions and phosphopeptides from the flow-through (FT) were further enriched by TiO_2 chromatography. Three consecutive rounds of enrichment were performed with the FT. The resulting 18 samples were eluted three times with 40 % ammonium hydroxide solution in ACN with pH 10.5. After elution, samples were evaporated to a final volume of 5 µL and injected into the LC-MS system.

Off-Gel fractionation

100 µg of protein digest were fractionated by isoelectric focusing using the OFFGEL 3100 Fractionator (Agilent) according to the manufacturer's instructions (32). Separation was performed on a 13 cm Immobiline DryStrip (GE Healthcare) at a maximum of 50 µA for 20 kVh. The peptide fractions were purified using Stage-Tips (33).

Strong anion exchange chromatography (SAX)

50 µg of protein digest were fractionated using strong anion exchange chromatography (SAX) (34). Briefly, digested proteins were desalted using solid phase extraction (SPE), mixed 1:1 with Britton & Robinson Universal Buffer at pH 11 (0.02 M CH_3COOH, 0.02 M H_3PO_4 and 0.02 M H_3BO_3) and loaded onto a Stage-Tip based anion exchange column containing six layers of Empore Anion Exchange (Varian). The flow through was collected and desalted using a Stage-Tip. Peptides were then eluted with Britton & Robinson Universal Buffer at pH 8, 6, 5, 4 and 3 and desalted using Stage-Tips.

NanoLC-MS/MS

An EASY-nLC II system (Proxeon Biosystems) coupled to an LTQ Orbitrap XL, an LTQ Orbitrap Velos or an LTQ Orbitrap Elite (Thermo) was used for the LC-MS/MS measurements. Peptides were separated on a 15 cm PicoTip fused silica emitter ID 75 µm (New Objective) packed in-house using reversed-phase beads ReproSil-Pur C18-AQ 3 µm (Dr. Maisch GmbH). Peptides were injected with Solvent A (0.5 % acetic acid) at a maximum pressure of 280 bar using the IntelliFlow option and separated at 200 nL/min. Separation was performed using a linear 106 min gradient of 5-33 % solvent B (80 % ACN in 0.5 % acetic acid) for phosphopeptide samples and either an 81 or a 221 min segmented 5-50 % solvent B gradient for all other samples.

For proteome samples, an LTQ Orbitrap Velos and an LTQ Orbitrap Elite were operated in positive mode. Precursor ions were recorded in the orbitrap analyzer at a resolution of 60,000 (LTQ Orbitrap Velos) or 120,000 (LTQ Orbitrap Elite) followed by MS/MS spectra acquired in

the LTQ mass analyzer. Data-dependent analysis was used and either the 15 (LTQ Orbitrap Velos) or the 20 (LTQ Orbitrap Elite) most abundant precursor ions from the full scan were fragmented using collision-induced dissociation (CID). Fragmented masses were excluded for 90 s. The target values were 1E6 charges for the MS scans in the Orbitrap and 5,000 charges for the MS/MS scans in the LTQ with a maximum fill time of 100 ms and 25 ms respectively.

Phosphoproteome samples were measured using an LTQ Orbitrap XL with the same parameters as for the proteome samples with the following changes: precursor ions were recorded at 60,000; the 5 most abundant precursor ions were fragmented by activation of the neutral loss ions at -98, -49 and -32.6 Thompson relative to the precursor ion (35); the maximum fill time was 150 ms for MS and 1000 ms for MS/MS scans.

Data processing and analysis

MS raw files were processed using MaxQuant software (v.1.2.2.9) (36). Multiplicity was set to three, matching the number of SILAC labels used ('light', 'medium', 'heavy') in each experiment; Lys4 and Lys8 were specified as 'medium' and 'heavy' labels, respectively. Database search was performed after peptide quantification using the Andromeda search engine (37), which is part of MaxQuant. Full enzyme specificity was required and up to two missed cleavages were allowed. For proteome and phosphoproteome measurements, Lys-C was specified as the enzyme used, for lysine acetylation-enriched samples, trypsin was specified as the enzyme used. The MS/MS spectra were searched against a target database consisting of 5,075 entries from the *S. pombe* proteome database (http://www.pombase.org/, Protein Dataset in FASTA format, downloaded on April 6th, 2011), 50 entries from the Proteomics Dynamic Range Standard (UPS2, Sigma), the mCherry protein sequence, the EGFP protein sequence, and 248 frequently seen laboratory contaminants. To control the false discovery rate of database search results, the software appends a corresponding decoy version of the target database resulting in a total of 10,750 database entries. Carbamidomethylation on cysteines was defined as fixed modification, oxidation of methionine and protein N-terminal acetylation were set as variable modifications; initial mass tolerance for the precursor ion was set to 6 ppm and fragment ion mass tolerance for CID fragmentation data was set to 0.5 Da; for HCD fragmentation data fragment ion mass tolerance was set to 20 ppm. In addition, we included phosphorylation of serine, threonine and tyrosine (STY) and acetylation of lysine as variable modifications for the data derived from samples enriched for phosphopeptides and lysine-acetylated peptides, respectively. Database search results were further processed for statistical validation and identified MS/MS spectra were assembled into peptides, protein groups and modification sites at a default false discovery rate (FDR) of 1 % at all levels. The estimated FDR after database search was 1 % at the protein group level, 0.97

% at the phosphorylation site level and 0.54 % at the peptide level (*S. pombe* and UPS proteins). A list containing *S. pombe* peptide and UPS sequences detected at an estimated FDR of 0.54 % without reverse hits and without contaminants can be found in **Supplementary Table 2**. For protein quantification, a minimum of two quantitation events was required. Peptide evidences of single-peptide-based protein identifications are listed in **Supplementary Table 3**.

Relative quantification

The two triple SILAC experiments were combined to get a ratio profile for each protein and phosphorylation event across the four cell cycle phases relative to the Lys4-labeled super-SILAC standard. Ratios of proteins and phosphorylation events were transformed to $\log_2$ scale and only proteins and phosphorylation events (P-events) that were quantified in all four phases were considered for further quantitative analysis. The magnitude of fluctuation was expressed by calculating the standard deviation of the log-transformed ratios across the four phases. The quartiles of the distribution resulting from all quantified features were determined and used to separately bin proteins and P-events according to the extent of fluctuation across the cell cycle phases. Those belonging to the quartile with the highest standard deviation (75-100 %) were defined as "fluctuating" or dynamic, whereas those belonging to the quartile with the lowest standard deviation (0-25 %) were defined as "non-fluctuating" or static.

Absolute protein quantification

We used the iBAQ approach for absolute protein quantification, which calculated absolute protein amounts proportional to the molar amounts of a spiked-in protein standard consisting of 48 proteins (UPS2, Sigma). The method is natively supported by the MaxQuant software and was operated using standard parameters (log-fit option). Resulting iBAQ values were subsequently adjusted according to the amount of UPS2 protein that was spiked in to one of the measured samples (M and S cell cycle phases, Off-Gel proteome measurement). The total number of cells in the sample (N_{cells}) was used to calculate the absolute protein copy numbers ($CN_{protein}$) by applying the following formula:

$$CN_{protein} = \frac{N_A\left(iBAQ_{protein} * 10^{-15}\right)}{N_{cells}}$$

with N_A being the Avogadro constant. Because the *cdc25-22*-synchronized cells have a cell volume that is 3.51-times larger than unsynchronized cells (27), the resulting values were divided by 3.51 to obtain the copy number in a typical unsynchronized cell. This assumes that protein concentration is similar between *cdc25+* and *cdc25-22* cells, which is not formally

known. The absolute protein amounts calculated for the super-SILAC standard together with corresponding SILAC ratios were used to calculate the absolute amount for each protein in the four cell cycle phases.

Absolute phosphorylation site occupancy

The proportion between the phosphorylated and nonphosphorylated version of the same peptide defines the occupancy of the corresponding phosphorylation site. Olsen et al. (4) described a model to calculate the occupancy of singly phosphorylated peptides by using the SILAC ratios of (1) the phosphorylated peptide, (2) the nonphosphorylated peptide, and (3) the corresponding protein, assuming that all three ratios are measured accurately, that the ratios change between the different states, and that there are only two versions of the peptide (nonphosphorylated and phosphorylated on a single site). For details on the calculation see Olsen et al. (4). We calculated occupancy values separately for the two replicates of our dataset using an in-house developed perl-script. The SILAC ratios of phosphorylated peptides were derived from the corresponding SCX fractions, whereas for nonphosphorylated peptides and proteins we used the SILAC ratios measured in the proteome measurements (Off-Gel).

Absolute phosphate stoichiometry

We calculated the absolute phosphate stoichiometry as product of site occupancy and protein copy number of the protein carrying the phosphorylation site. This calculation was done only when both values for a specific phase were available (101 sites in G2 phase, 71 in sites M phase, 81 in G1 phase, 67 in S phase). In order to estimate the global level of protein bound phosphate, we first inferred estimates of median occupancy and standard deviation by performing 10,000 bootstrap iterations on the available occupancy values in each cell cycle phase. We then imputed missing occupancy values by sampling from a normal distribution using the bootstrap estimates as parameters for mean and standard deviation and calculated absolute phosphate stoichiometry as described above. We calculated bootstrap confidence intervals (CI) for the median occupancy using the 0.025 and 0.975 percentiles of the bootstrap distribution resulting in a 95 % CI.

Peak time index calculation

We performed the analysis on expression peak time indices of proteins and phosphorylation sites as described in (4). Briefly, the SILAC ratios for each item (i.e. protein or P-event) were normalized to the maximal change across the four cell cycle phases. For each item we calculated the weighted mean of the expression ratio at the phase with maximal change with respect to the normalized ratios of adjacent phases. To assign the resulting 'peak time index' of every item to a specific cell cycle phase, we applied hierarchical clustering on these values

using the Euclidian distance and a defined cluster number corresponding to the four cell cycle phases addressed in this study.

Functional enrichment analysis

We retrieved Gene Ontology (GO) annotation of *S. pombe* from GeneDB (downloaded on December 6th, 2011). To test whether specific annotation terms are enriched or depleted within a set of proteins of interest we applied Fisher's exact test using the theoretical *S. pombe* proteome as background. Derived p-values were further adjusted to address multiple hypothesis testing using the method proposed by (38).

3.4   Results

We measured relative and absolute dynamics of the proteome and phosphoproteome during the four major phases of the cell cycle – G2, M, G1 and S – using highly synchronized *S. pombe* cultures and SILAC **(Figure 1; Supplementary Figure 1)**. To relatively quantify all four cell cycle phases we generated a super-SILAC standard (26), consisting of a mixture of cells from all four phases. The standard was combined with two of the cell cycle phase-specific samples (either G1 and G2, or M and S) to yield a triple-SILAC sample. The quantitative information from both triple-SILAC measurements was integrated through this common standard to obtain the relative dynamics of the proteome and phosphoproteome across the cell cycle **(Figure 2A)**. Fluorescence microscopy confirmed more than 90 % synchronicity for all analyzed cell cycle phases, and characteristic changes were verified by immunoblotting **(Figure 2B and C; Supplementary Figure 1)**. In 137 LC-MS runs, we recorded 2,471,241 MS/MS spectra that identified a total of 32,016 non-redundant peptides from 3,753 *S. pombe* proteins at estimated false discovery rates (FDR) of 0.53 % at the peptide and 0.99 % at the protein level. The identified proteins accounted for 74 % of the predicted *S. pombe* proteome (5,075 proteins). In the phosphoproteome measurement we identified 3,682 phosphorylation events (on 1,554 proteins) of which 2,465 were localized to a specific Ser/Thr/Tyr residue **(Supplementary Table 4)**. The identified and quantified proteins as well as phosphorylation events are presented in **Supplementary Table 5 and 6**. Additionally, we detected 328 acetylation events on 184 proteins (including histones, cohesin, SAGA subunits, ribosomal proteins and metabolic enzymes; **Supplementary Table 7**), which were not analyzed further due to technical bias towards highly abundant proteins as a result of low enrichment efficiency (not shown).
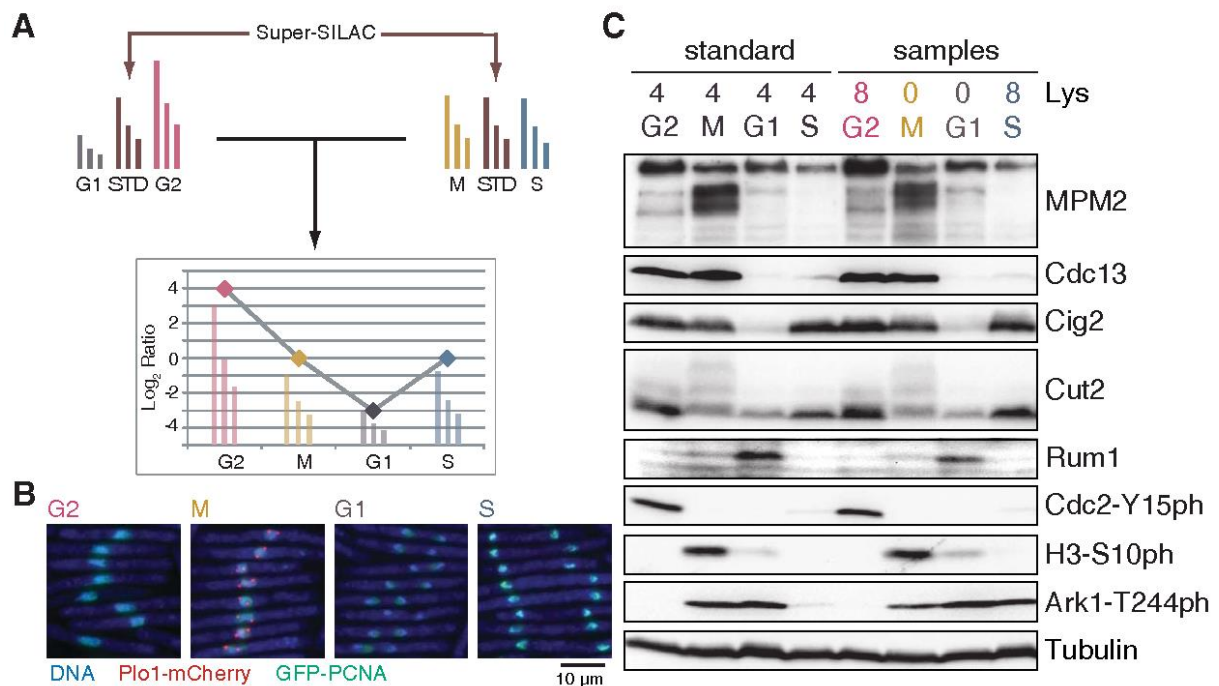
Figure 2. **Data integration and synchronicity for all analyzed cell cycle phases.** *(A) The samples are compared to each other using a super-SILAC standard (STD) containing a mixture of G2, M, G1 and S phases. By comparing the information of these two experiments through the common standard, a profile for each quantified protein and phosphorylation event is obtained. (B) Synchronization of the cells was confirmed by fluorescence microscopy using Plo1-mCherry (M phase) and EGFP-PCNA (S phase) as markers. A synchronicity of more than 90 % was achieved in the samples used for G1/G2 and the M/S samples (Supplementary Fig. 1). (C) Immunoblot analysis of known cell cycle-dependent proteins or phosphorylation events in the standard and experimental samples. MPM2 is an antibody generated against mitosis-specific phosphoepitopes (Supplementary Table 1).*

Relative quantification of proteome and phosphoproteome dynamics

Application of the super-SILAC approach enabled relative quantification of 2,416 proteins and 1,963 phosphorylation events across all analyzed phases of the cell cycle. We achieved high reproducibility between the spiked-in super-SILAC standard of each triple SILAC sample ($0.9 \leq r \leq 0.93$) **(Supplementary Figure 2)**. Changes on the protein level were minimal **(Figure 3A)**, which lead to a relatively low correlation of protein ratios ($0.26 \leq r \leq 0.62$) **(Supplementary Figure 3A)**. In contrast, changes at the phosphoproteome level were considerably more pronounced **(Figure 3B)** and showed high reproducibility between both replicates ($0.7 \leq r \leq 0.82$) **(Supplementary Figure 3B)**. Normalization of phosphorylation site abundance to the protein level resulted in negligible differences **(Supplementary Figure 3C-D)**, so that we analyzed phosphorylation data without protein normalization. We validated our data by performing immunoblots and fluorescence microscopy to measure dynamics of selected positive controls, which were overall in agreement with the MS data **(Supplementary Figures 4 and 5).**
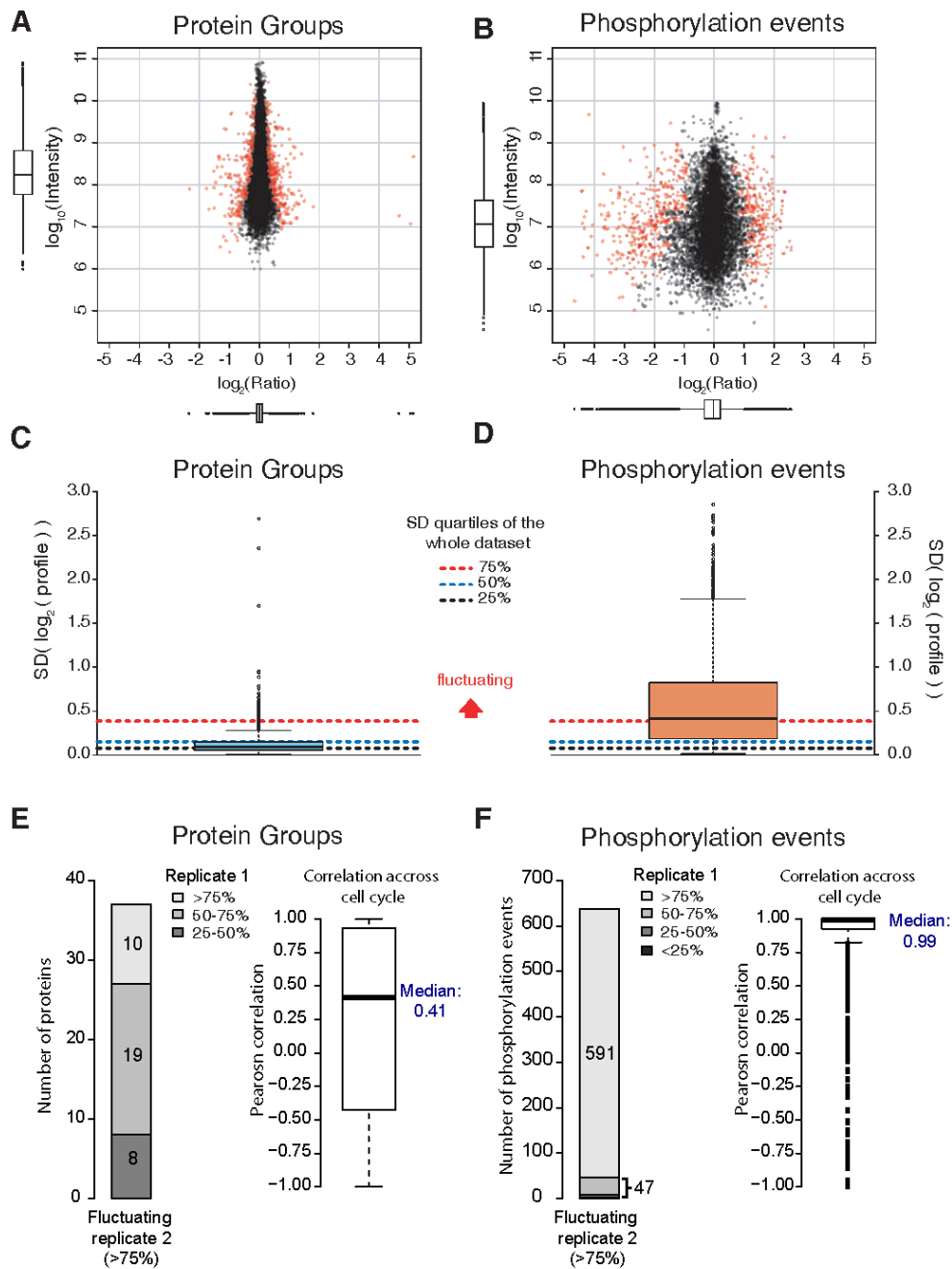
Figure 3. *Dynamics of proteins and phosphorylation events across the cell cycle phases. Diagrams for each quantified (A) protein and (B) phosphorylation event relative to the super-SILAC standard plotted against its intensity recorded in the mass spectrometer. Red dots indicate statistically significant (p < 0.01) outliers from the distribution. The boxplots below the figures depict the distribution of SILAC ratios described above. Boxplot diagrams of standard deviations (SD) calculated from SILAC ratios across the four phases of the cell cycle for (C) proteins and (D) phosphorylation events; dashed lines represent the quartiles of SDs based on the whole dataset; these were used to group proteins and phosphorylation events according to their extent of fluctuation across the cell cycle; the upper quartile (>75%) was used to define fluctuating proteins and phosphorylation events (see methods). Overlap of (E) fluctuating proteins or (F) fluctuating phosphorylation events between the replicates and boxplot of correlation coefficients between the profiles of fluctuating proteins or phosphorylation events.*

To assess the extent of fluctuation of proteins and phosphorylation events across the cell cycle, we used the standard deviation of log-transformed SILAC ratios across the different cell cycle phases and binned the proteins and phosphorylation events into quartiles of the resulting distribution **(Figure 3C-D)**. Proteins or phosphorylation events with the highest standard deviation were defined as fluctuating. This analysis confirmed that strong fluctuation predominantly occurred at the phosphoproteome level (1,032 fluctuating events, **Supplementary Table 8**), whereas protein abundance during the cell cycle was much more static (only 54 fluctuating proteins were detected, **Supplementary Table 9**). Of these 1,032 fluctuating phosphorylation events and 54 proteins, 638 events and 37 proteins were quantified across all four analyzed cell cycle phases in the replicate analysis. From this overlap, 591 phosphorylation events were fluctuating in both replicates and showed high correlation in the relative expression across the four cell cycle phases (median correlation coefficient: 0.99). A lower number of proteins (10) was fluctuating in both replicates and had lower correlation (median correlation coefficient: 0.41), likely as a consequence of the lower extent of protein fluctuation **(Figure 3E-F)**. Relative expression of fluctuating proteins and phosphorylation events can be found in **Supplementary Data File I and II**. Due to high correlation, especially of fluctuating phosphorylation events, we concluded that each of the replicates is representative of the whole dataset and therefore we proceeded with one replicate for further analysis.

Functional classes of fluctuating proteins and phosphorylation events

The strongest fluctuating protein that was identified in all four phases was the cyclin B Cdc13 (high in G2 and M, low in G1 and S, **Supplementary Table 9**), whose degradation at the transition from M to G1 is well known (39) **(Figure 2C)**. Strongly fluctuating was also the septation protein Uds1 (highest in S). This had been observed by MS previously (40) and we confirmed by immunoblotting and fluorescence microscopy **(Figure 2C, Supplementary Figure 4 and 5)**. On the phosphoproteome level, we recovered the well-known cell cycle-regulated phosphorylation of histone H3-Ser10 (41) **(Figure 2C)** as well as several previously known cell cycle-specific phosphorylations, including mitosis-specific phosphorylation of monopolin Mde4 (42), survivin Bir1 (43) and previously identified sites phosphorylated by Aurora kinase (17). Interestingly, we observed considerable cell cycle-specific fluctuation in the phosphorylation of nucleoporins, which mimics observation in human cells (44), despite the fact that fission yeast, unlike human cells, undergoes a closed mitosis **(Figure 4A; Supplementary Table 8)**.
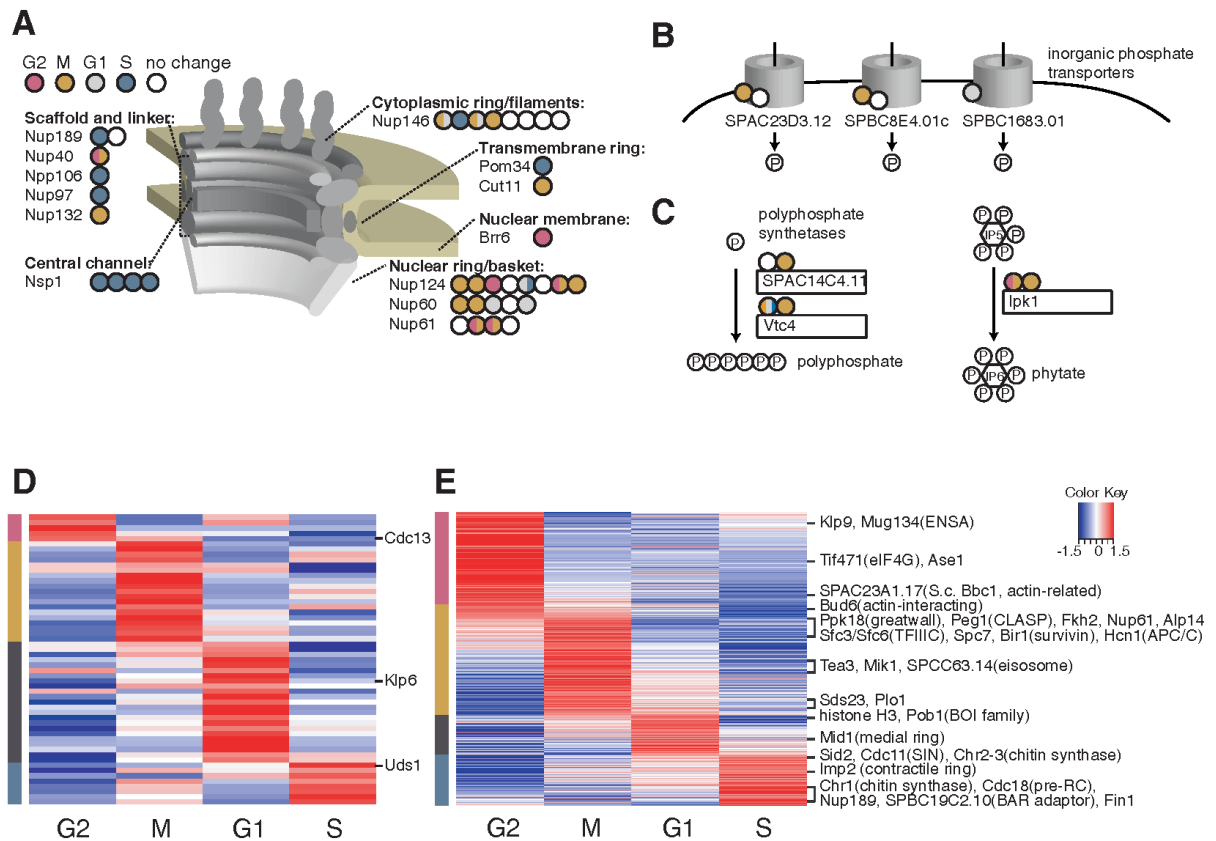
Figure 4. *Selected cell cycle dependent phosphorylation events and sorting of protein and phosphorylation variation according to peak time index.* (A) Cell cycle dependent variation in the phosphorylation of nuclear pore or nuclear membrane related proteins, (B) inorganic phosphate transporters, or (C) phosphate metabolism enzymes. Each circle represents a phosphorylation event detected on the respective protein and colored according to the cell cycle phase with the highest phosphorylation; circles with a P represent phosphate. (D) Proteins or (E) phosphorylation events within the quartile of highest fluctuation were clustered in the cell cycle phases according to their peak time index (see methods). Protein and phosphorylation events in this figure can be found and are highlighted in Supplementary Tables 8 and 9.

We performed functional enrichment analysis for the proteome and the phosphorylation data separately for the different quartiles **(Supplementary Figures 6A-B)**. As expected, cell cycle-related domains and processes such as regulation of cyclin−dependent protein kinase activity and septum were overrepresented among the fluctuating proteins. Housekeeping processes like the TCA, citric acid and urea cycles, glycolysis, ribosome and translation were overrepresented among non-fluctuating proteins. At the phosphorylation level, cell cycle cytokinesis, mitotic cell cycle spindle assembly checkpoint, spindle assembly and regulation of transcription during mitosis were overrepresented, whereas translation and mitochondrial translation were among the underrepresented functional classes. Our global analysis also indicated which kinases have a large influence on cell-cycle dependent fluctuations (CDK1, CDK2, Polo kinase), and which kinases act mostly in a cell-cycle independent fashion (CK1, CK2) **(Supplementary Figure 6C)**.
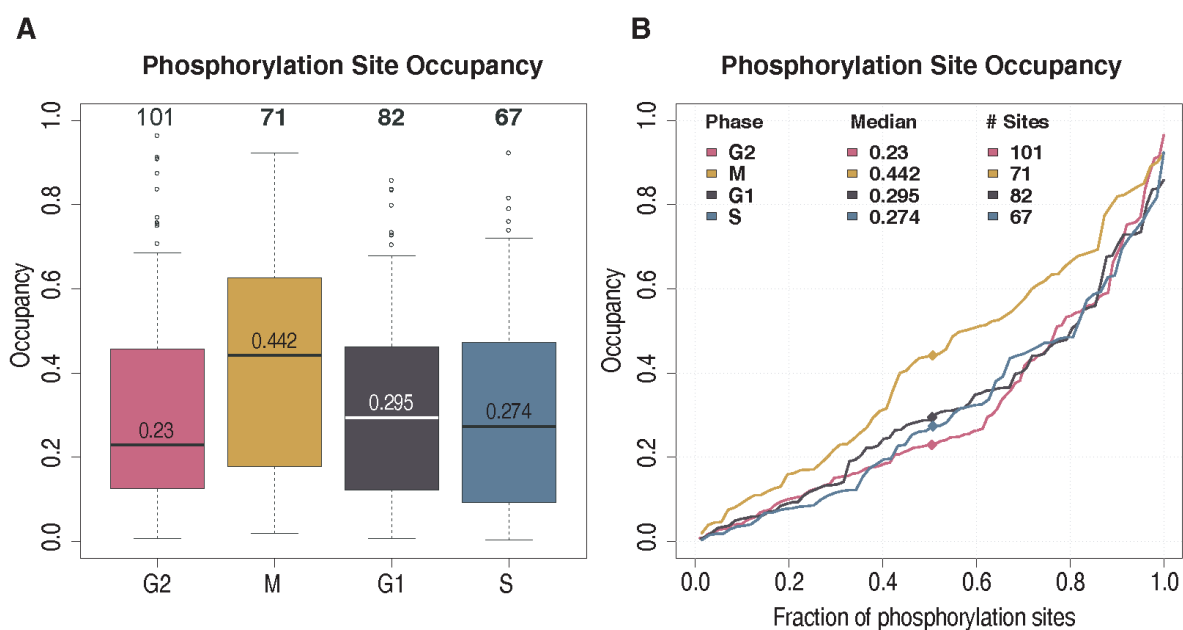
Figure 5. **_Phosphorylation site stoichiometry._** *(A) Boxplots of phosphorylation site stoichiometry calculations for each cell cycle phase (see methods). The number of phosphorylation sites is shown at the top of the figure, while the numbers in the boxes represent the median values of all calculated phosphorylation site occupancies. (B) Cumulative diagram of the phosphorylation site occupancy in the studied cell cycle phases.*

For each of the fluctuating proteins or phosphorylation events, we determined the cell cycle phase with the highest value and used the peak time indices (4) as similarity measure for hierarchical cluster analysis. The resulting four clusters contained the proteins and phosphorylation events that were most prominent in each analyzed cell cycle phase **(Figure 4D-E)**. Functional enrichment analysis of proteins and phosphorylation events in the resulting phase-specific clusters revealed that kinase-related functions increased specifically in the G2 phase. Furthermore, several mitosis-related biological processes were increased during mitosis, further validating our approach **(Supplementary Figure 6D-E).**

Using the measured SILAC ratios of phosphoproteins, phosphopeptides and their unmodified counterparts, we calculated stoichiometry (occupancy) of 101 phosphorylation sites in G2 phase, 71 sites in M phase, 82 sites in G1 phase and 67 sites in S phase. In agreement with previous studies (4), the highest median occupancy of 44.2 % was measured in M phase. Average occupancy was reduced in the subsequent phases reaching a minimum of 23 % in G2 phase **(Figure 5)**.

Absolute quantification

We estimated absolute levels of *S. pombe* proteins using the iBAQ approach (24). This method correlates the MS signal intensities and the number of observable peptides of each protein to a spiked-in protein standard. We spiked 48 proteins from 0.5 fmol to 50,000 fmol into one fraction of our measurements (Off-Gel fractions of M and S phase triple SILAC measurement) and detected 23 of the spiked-in proteins at concentrations between 50 fmol and 50,000 fmol. This pointed to an underrepresentation of very low abundant proteins. The detected standard proteins enabled the calculation of iBAQ values for 3,178 *S. pombe* proteins **(Supplementary Figure 7A)**.

The dynamic range of protein copy numbers spanned almost five orders of magnitude, which was in agreement to the dynamic range determined in asynchronously growing *S. pombe* cells (18, 22) **(Figure 6A, Supplementary Table 10, Supplementary Figure 8).** The total number of protein molecules in an average *S. pombe* cell was 176 million, with single proteins having between 92 (Snf59) and 1.4 million (enolase) copies. Eighty percent of the whole protein content was constituted by twenty percent of the most abundant proteins, consistent with the Pareto 80-20 principle, also observed by Marguerat et al. (18) **(Figure 6B)**.

We compared our absolute quantification to several studies that used different technologies to derive absolute protein amounts in *S. pombe*: MS-intensity based absolute quantification (18, 19, 22), and fluorescence microscopy-based quantitation (27, 45). It has been reported that *cdc25-22* cells under the used growth conditions are 3.51 fold larger, which was taken into account when normalizing the protein copy numbers (see methods) (27). Overall, we found a good correlation of protein copy numbers compared to previous studies (Pearson-correlation > 0.7 **Supplementary Figure 7B-F**). Our numbers were higher when compared with Gunaratne et al. (19), however they were in agreement with the other four studies used for comparison (18, 22, 27, 45). The stoichiometry of several known stable protein complexes was also compared and subunit stoichiometry was in agreement with previous studies **(Table I)**. All studies dealing with estimation of absolute protein numbers are exposed to multiple sources of error (see Discussion) and all absolute numbers (from our and others' studies alike) should therefore be treated as estimates.
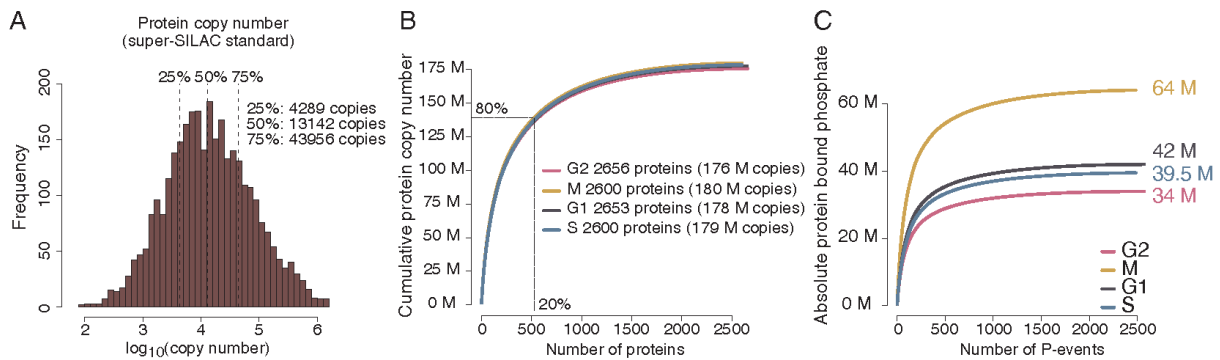
Figure 6. *Absolute quantification.* (A) Histogram of the protein copy number in log_{10} from the super-SILAC standard with dashed lines indicating the quartiles. (B) Cumulative diagram of protein copy number in each cell cycle phase. (C) Cumulative diagram of absolute protein-bound phosphate groups in each cell cycle phase (see methods).

Estimation of the absolute amount of protein bound phosphate groups

The availability of phosphorylation site occupancies and absolute protein copy numbers enabled us to address the absolute amount of Ser/Thr/Tyr phosphorylation in the different phases of the cell cycle. In order to estimate the absolute level of protein-bound phosphate we imputed missing occupancy values by bootstrapping (see methods) and calculated the number of protein-bound phosphate for all proteins that we absolutely quantified. Our results showed that the total number of protein-bound phosphate groups fluctuates between 0.056 and 0.106 fmol or between 34 million (G2 phase) and 64 million phosphate groups (M phase). The cumulative distribution of the estimated number of protein-bound phosphate groups had a very similar distribution to the cumulative distribution of the protein copy numbers and both followed the 80/20 Pareto principle. **(Figure 6C, Supplementary Figure 9)**.

In order to assess the amount of ATP required for protein phosphorylation during the G2/M transition we measured the intracellular amount of ATP in asynchronous cells. We then used this information to address the absolute consumption of ATP needed for protein phosphorylation across the cell cycle. We measured the amount of ATP to be 217 ± 14 amol per cell, which is in agreement with values reported before (46, 47). The amount of ATP required for protein STY phosphorylation during the G2/M transition is 30 million phosphate molecules (50 amol), which corresponds to 23 % of the average amount of ATP in asynchronous cells. Hence, a significant amount of ATP is needed to phosphorylate proteins required to drive the cell cycle through the G2/M transition. Consistently, we observe phosphorylation (and hence presumably regulation) of phosphate transporters and enzymes of phosphate metabolism during mitosis **(Figure 4B-C, Supplementary Table 6 and 9)**.

3.5   Discussion

The cell cycle is a complex dynamic process involving multiple layers of regulation and its global analysis is challenging. The super-SILAC approach combined with high accuracy mass spectrometry is a powerful tool that enables integrative analysis of several samples by comparing them to a common standard. Super-SILAC was used previously to analyze specific proteomes of various cell lines (48), cancer cell lines (26) and mouse tissues (49). Here we applied it to measure relative dynamics of the proteome and phosphoproteome during the four major phases of the cell cycle (G1, S, G2 and M) in fission yeast. A similar approach was used by Olsen et al. to quantify phosphoproteome/proteome dynamics of HeLa cells, albeit using unsynchronized cells instead of an equal mixture of labeled synchronized cells as internal standard (4). The 3,753 proteins identified in our study represent 74 % of the predicted fission yeast proteins, making it one of the most comprehensive published protein dataset of this organism to date. To enable triple-SILAC labeling and avoid quantification problems arising from arginine to proline conversion (40), we exclusively used SILAC labeling with lysine and digestion with endoproteinase Lys-C. Our data complements the recently published PeptideAtlas of the fission yeast proteome obtained using trypsin (19), and increases the coverage of the *S. pombe* proteome from 3,542 to a total of 4,068 proteins.

In the replicate analyzed we were able to relatively quantify 80 % of all identified proteins in at least one cell cycle phase and 65 % of the proteins in all phases. Surprisingly, despite detecting many proteins for which the mRNA was shown to fluctuate during the cell cycle (50-52), the variations in protein abundance were low. Analysis of phosphoproteome dynamics showed a different picture: of 3,682 phosphorylation events, at least 30 % were fluctuating, confirming that considerable cell cycle regulation occurs through phosphorylation. We suspect that the mRNA fluctuations are not translated into proteome changes because we use a culture that is cycling at maximum speed. Slower progression through the cell cycle would provide the time needed to reach a new steady state of protein abundance based on the modified mRNA level. Possibly, phosphorylation is used for fast changes in the cell cycle, whereas changes in the proteome only occur on a slower time scale. Our data also suggest that most cell cycle-dependent changes on the mRNA level are not essential for cell cycle progression. It should be noted that our synchronization strategy involved a temperature shift. This and sample processing may elicit stress responses, which could lead to proteome and phosphoproteome changes that are superimposed to the cell cycle changes. However, this does not affect our conclusion that there is surprisingly little protein fluctuation in these synchronized, fast cycling cells.

To complement the relative quantification we estimated copy numbers of 3,178 proteins using iBAQ (24). Proteins that fluctuate across the cell cycle have a low abundance and represent only 0.35 % of the total amount of protein in the cell (4th quartile). In contrast to the proteome, fluctuating phosphorylation events are present on proteins of the entire abundance range. We could also confirm that smaller proteins tend to be more abundant than larger proteins, as had previously been seen in budding and fission yeast (18, 19, 53) **(Supplementary Figure 10)**. The use of SILAC-based relative quantification enabled us to address the stoichiometry of detected phosphorylation events. As expected, the overall phosphorylation level was peaking in mitosis, where it was about two-fold higher than in the other cell cycle phases. Our median occupancy of single sites in M phase (44 %) was significantly lower than reported in HeLa cells (~70 %) (4). The reasons could include biological (species-related) differences as well as experimental differences, since Nocodazole was used to arrest HeLa cells, whereas our cells progressed through mitosis unperturbed.

It is important to point out that calculations of absolute levels depend on many different parameters such as cell counting, cell size measurement, iBAQ measurement, determination of UPS standard and protein concentration. As each of these parameters presents a source of error, only approximate absolute values can be determined. Nevertheless, our dataset shows good correlation with previously reported similar datasets and satisfactory internal control parameters, such as stoichiometry of known complex subunits. Hence, the data reported here will be a useful resource for classical and systems biologists alike.

3.6 Contributions

Silke Hauf and Boris Macek designed the experiments. Sabine Graf and André Koch performed cell culture, immunoblot analyses, microscopy, and phosphoproteome experiments. Karsten Krug performed bioinformatic analysis. Sasa Popic calculated phosphorylation site stoichiometry with my assistance. I performed MS measurements, proteome analysis, ATP quantification and data processing and prepared most of the figures. Silke Hauf, Boris Macek and I wrote the manuscript. My contribution to this manuscript was about 50%.

3.7 References

1. Kastan, M. B., and Bartek, J. (2004) Cell-cycle checkpoints and cancer. *Nature* 432, 316-323

2. Malumbres, M., and Barbacid, M. (2009) Cell cycle, CDKs and cancer: a changing paradigm. *Nat. Rev. Cancer* 9, 153-166

3. Lapenna, S., and Giordano, A. (2009) Cell cycle kinases as therapeutic targets for cancer. *Nat. Rev. Drug Discov.* 8, 547-566

4. Olsen, J. V., Vermeulen, M., Santamaria, A., Kumar, C., Miller, M. L., Jensen, L. J., Gnad, F., Cox, J., Jensen, T. S., Nigg, E. A., Brunak, S., and Mann, M. (2010) Quantitative Phosphoproteomics Reveals Widespread Full Phosphorylation Site Occupancy During Mitosis. *Sci. Signal.* 3, ra3-

5. Pagliuca, F. W., Collins, M. O., Lichawska, A., Zegerman, P., Choudhary, J. S., and Pines, J. (2011) Quantitative proteomics reveals the basis for the biochemical specificity of the cell-cycle machinery. *Mol. Cell* 43, 406-417

6. Forsburg, S. L., and Nurse, P. (1991) Cell Cycle Regulation in the Yeasts Saccharomyces Cerevisiae and Schizosaccharomyces Pombe. *Annu. Rev. Cell Biol.* 7, 227-256

7. Ghaemmaghami, S., Huh, W., Bower, K., Howson, R. W., Belle, A., Dephoure, N., O'Shea, E. K., and Weissman, J. S. (2003) Global analysis of protein expression in yeast. *Nature* 425, 737-741

8. Ross, P. L., Huang, Y. L. 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. *Mol. Cell. Proteomics* 3, 1154-1169

9. 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. *Nat. Biotechnol.* 17, 994-999

10. Ho, Y., et al. (2002) Systematic identification of protein complexes in Saccharomyces cerevisiae by mass spectrometry. *Nature* 415, 180-183

11. de Godoy, L. M. F., Olsen, J. V., Cox, J., Nielsen, M. L., Hubner, N. C., Froehlich, F., Walther, T. C., and Mann, M. (2008) Comprehensive mass-spectrometry-based proteome quantification of haploid versus diploid yeast. *Nature* 455, 1251-U1260

12. Flory, M. R., Lee, H., Bonneau, R., Mallick, P., Serikawa, K., Morris, D. R., and Aebersold, R. (2006) Quantitative proteomic analysis of the budding yeast cell cycle using acid-cleavable isotope-coded affinity tag reagents. *Proteomics* 6, 6146-6157

13. Breitkreutz, A., Choi, H., Sharom, J. R., Boucher, L., Neduva, V., Larsen, B., Lin, Z. Y., Breitkreutz, B. J., Stark, C., Liu, G. M., Ahn, J., Dewar-Darch, D., Reguly, T., Tang, X. J.,

Almeida, R., Qin, Z. S., Pawson, T., Gingras, A. C., Nesvizhskii, A. I., and Tyers, M. (2010) A Global Protein Kinase and Phosphatase Interaction Network in Yeast. *Science* 328, 1043-1046

14. Schmidt, M. W., Houseman, A., Ivanov, A. R., and Wolf, D. A. (2007) Comparative proteomic and transcriptomic profiling of the fission yeast Schizosaccharomyces pombe. *Mol. Syst. Biol.* 3

15. Matsuyama, A., Arai, R., Yashiroda, Y., Shirai, A., Kamata, A., Sekido, S., Kobayashi, Y., Hashimoto, A., Hamamoto, M., Hiraoka, Y., Horinouchi, S., and Yoshida, M. (2006) ORFeome cloning and global analysis of protein localization in the fission yeast Schizosaccharomyces pombe. *Nat Biotech* 24, 841-847

16. Nurse, P. (1985) Cell cycle control genes in yeast. *Trends in genetics : TIG* 1, 51-55

17. Koch, A., Krug, K., Pengelley, S., Macek, B., and Hauf, S. (2011) Mitotic Substrates of the Kinase Aurora with Roles in Chromatin Regulation Identified Through Quantitative Phosphoproteomics of Fission Yeast. *Sci. Signal.* 4, rs6

18. Marguerat, S., Schmidt, A., Codlin, S., Chen, W., Aebersold, R., and Bähler, J. (2012) Quantitative Analysis of Fission Yeast Transcriptomes and Proteomes in Proliferating and Quiescent Cells. *Cell* 151, 671-683

19. Gunaratne, J., Schmidt, A., Quandt, A., Neo, S. P., Sarac, O. S., Gracia, T., Loguercio, S., Ahrne, E., Li Hai Xia, R., Tan, K. H., Loessner, C., Bahler, J., Beyer, A., Blackstock, W., and Aebersold, R. (2013) Extensive Mass Spectrometry-Based Analysis of the Fission Yeast Proteome: The S. pombe PeptideAtlas. *Mol. Cell. Proteomics*

20. Wilson-Grady, J. T., Villen, J., and Gygi, S. P. (2008) Phosphoproteome analysis of fission yeast. *J. Proteome Res.* 7, 1088-1097

21. Beltrao, P., Trinidad, J. C., Fiedler, D., Roguev, A., Lim, W. A., Shokat, K. M., Burlingame, A. L., and Krogan, N. J. (2009) Evolution of phosphoregulation: comparison of phosphorylation patterns across yeast species. *PLoS Biol.* 7, e1000134

22. Kulak, N. A., Pichler, G., Paron, I., Nagaraj, N., and Mann, M. (2014) Minimal, encapsulated proteomic-sample processing applied to copy-number estimation in eukaryotic cells. *Nat Meth* 11, 319-324

23. Geiger, T., Cox, J., Ostasiewicz, P., Wisniewski, J. R., and Mann, M. (2010) Super-SILAC mix for quantitative proteomics of human tumor tissue. *Nat Meth* 7, 383-385

24. Schwanhausser, B., Busse, D., Li, N., Dittmar, G., Schuchhardt, J., Wolf, J., Chen, W., and Selbach, M. (2011) Global quantification of mammalian gene expression control. *Nature* 473, 337-342

25. Moreno, S., Hayles, J., and Nurse, P. (1989) Regulation of p34cdc2 protein kinase during mitosis. *Cell* 58, 361-372

26. Geiger, T., Cox, J., Ostasiewicz, P., Wisniewski, J. R., and Mann, M. (2010) Super-SILAC mix for quantitative proteomics of human tumor tissue. *Nat. Methods* 7, 383-385

27. Heinrich, S., Geissen, E. M., Kamenz, J., Trautmann, S., Widmer, C., Drewe, P., Knop, M., Radde, N., Hasenauer, J., and Hauf, S. (2013) Determinants of robustness in spindle assembly checkpoint signalling. *Nat. Cell Biol.* 15, 1328-1339

28. Franz-Wachtel, M., Eisler, S. A., Krug, K., Wahl, S., Carpy, A., Nordheim, A., Pfizenmaier, K., Hausser, A., and Macek, B. (2012) Global Detection of Protein Kinase D-dependent Phosphorylation Events in Nocodazole-treated Human Cells. *Mol. Cell. Proteomics* 11, 160-170

29. Choudhary, C., Kumar, C., Gnad, F., Nielsen, M. L., Rehman, M., Walther, T. C., Olsen, J. V., and Mann, M. (2009) Lysine acetylation targets protein complexes and co-regulates major cellular functions. *Science* 325, 834-840

30. Rappsilber, J., Mann, M., and Ishihama, Y. (2007) Protocol for micro-purification, enrichment, pre-fractionation and storage of peptides for proteomics using StageTips. *Nature protocols* 2, 1896-1906

31. Olsen, J., and Macek, B. (2009) High Accuracy Mass Spectrometry in Large-Scale Analysis of Protein Phosphorylation. In: Lipton, M., and Paša-Tolic, L., eds. *Mass Spectrometry of Proteins and Peptides*, pp. 131-142, Humana Press

32. Hubner, N. C., Ren, S., and Mann, M. (2008) Peptide separation with immobilized pI strips is an attractive alternative to in-gel protein digestion for proteome analysis. *Proteomics* 8, 4862-4872

33. Rappsilber, J., Mann, M., and Ishihama, Y. (2007) Protocol for micro-purification, enrichment, pre-fractionation and storage of peptides for proteomics using StageTips. *Nat. Protocols* 2, 1896-1906

34. Wiśniewski, J. R., Zougman, A., and Mann, M. (2009) Combination of FASP and StageTip-Based Fractionation Allows In-Depth Analysis of the Hippocampal Membrane Proteome. *Journal of Proteome Research* 8, 5674-5678

35. Schroeder, M. J., Shabanowitz, J., Schwartz, J. C., Hunt, D. F., and Coon, J. J. (2004) A Neutral Loss Activation Method for Improved Phosphopeptide Sequence Analysis by Quadrupole Ion Trap Mass Spectrometry. *Anal. Chem.* 76, 3590-3598

36. 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 Biotech* 26, 1367-1372

37. Thakur, S. S., Geiger, T., Chatterjee, B., Bandilla, P., Frohlich, F., Cox, J., and Mann, M. (2011) Deep and highly sensitive proteome coverage by LC-MS/MS without prefractionation. *Molecular & cellular proteomics : MCP* 10, M110.003699

38. Benjamini, Y., and Hochberg, Y. (1995) Controlling the False Discovery Rate: A Practical and Powerful Approach to Multiple Testing. *Journal of the Royal Statistical Society. Series B (Methodological)* 57, 289-300

39. Yamano, H., Gannon, J., and Hunt, T. (1996) The role of proteolysis in cell cycle progression in Schizosaccharomyces pombe. *EMBO J.* 15, 5268-5279

40. Bicho, C. C., de Lima Alves, F., Chen, Z. A., Rappsilber, J., and Sawin, K. E. (2010) A Genetic Engineering Solution to the "Arginine Conversion Problem" in Stable Isotope Labeling by Amino Acids in Cell Culture (SILAC). *Mol. Cell. Proteomics* 9, 1567-1577

41. Hsu, J. Y., Sun, Z. W., Li, X., Reuben, M., Tatchell, K., Bishop, D. K., Grushcow, J. M., Brame, C. J., Caldwell, J. A., Hunt, D. F., Lin, R., Smith, M. M., and Allis, C. D. (2000) Mitotic phosphorylation of histone H3 is governed by Ipl1/aurora kinase and Glc7/PP1 phosphatase in budding yeast and nematodes. *Cell* 102, 279-291

42. Choi, S. H., Peli-Gulli, M. P., McLeod, I., Sarkeshik, A., Yates, J. R., 3rd, Simanis, V., and McCollum, D. (2009) Phosphorylation state defines discrete roles for monopolin in chromosome attachment and spindle elongation. *Curr. Biol.* 19, 985-995

43. Tsukahara, T., Tanno, Y., and Watanabe, Y. (2010) Phosphorylation of the CPC by Cdk1 promotes chromosome bi-orientation. *Nature* 467, 719-723

44. Laurell, E., and Kutay, U. (2011) Dismantling the NPC permeability barrier at the onset of mitosis. *Cell Cycle* 10, 2243-2245

45. Wu, J.-Q., and Pollard, T. D. (2005) Counting Cytokinesis Proteins Globally and Locally in Fission Yeast. *Science* 310, 310-314

46. Pluskal, T., Nakamura, T., Villar-Briones, A., and Yanagida, M. (2010) Metabolic profiling of the fission yeast S. pombe: quantification of compounds under different temperatures and genetic perturbation. *Mol. Biosyst.* 6, 182-198

47. Ingram, S. W., and Barnes, L. D. (2000) Disruption and overexpression of the Schizosaccharomyces pombe aph1 gene and the effects on intracellular diadenosine 5',5'''-P1, P4-tetraphosphate (Ap4A), ATP and ADP concentrations. *Biochem. J.* 350, 663-669

48. Geiger, T., Wehner, A., Schaab, C., Cox, J., and Mann, M. (2012) Comparative Proteomic Analysis of Eleven Common Cell Lines Reveals Ubiquitous but Varying Expression of Most Proteins. *Mol. Cell. Proteomics* 11

49. Geiger, T., Velic, A., Macek, B., Lundberg, E., Kampf, C., Nagaraj, N., Uhlen, M., Cox, J., and Mann, M. (2013) Initial Quantitative Proteomic Map of 28 Mouse Tissues Using the SILAC Mouse. *Mol. Cell. Proteomics* 12, 1709-1722

50. Oliva, A., Rosebrock, A., Ferrezuelo, F., Pyne, S., Chen, H., Skiena, S., Futcher, B., and Leatherwood, J. (2005) The cell cycle-regulated genes of Schizosaccharomyces pombe. *PLoS Biol.* 3, e225

51. Peng, X., Karuturi, R. K., Miller, L. D., Lin, K., Jia, Y., Kondu, P., Wang, L., Wong, L. S., Liu, E. T., Balasubramanian, M. K., and Liu, J. (2005) Identification of cell cycle-regulated genes in fission yeast. *Mol. Biol. Cell* 16, 1026-1042

52. Rustici, G., Mata, J., Kivinen, K., Lio, P., Penkett, C. J., Burns, G., Hayles, J., Brazma, A., Nurse, P., and Bahler, J. (2004) Periodic gene expression program of the fission yeast cell cycle. *Nat. Genet.* 36, 809-817

53. Warringer, J., and Blomberg, A. (2006) Evolutionary constraints on yeast protein size. *BMC Evol. Biol.* 6, 61

3.8 Supporting information

Supplementary tables S1 – S10 and Supplementary Data Files I and II are contained in the CD attached.
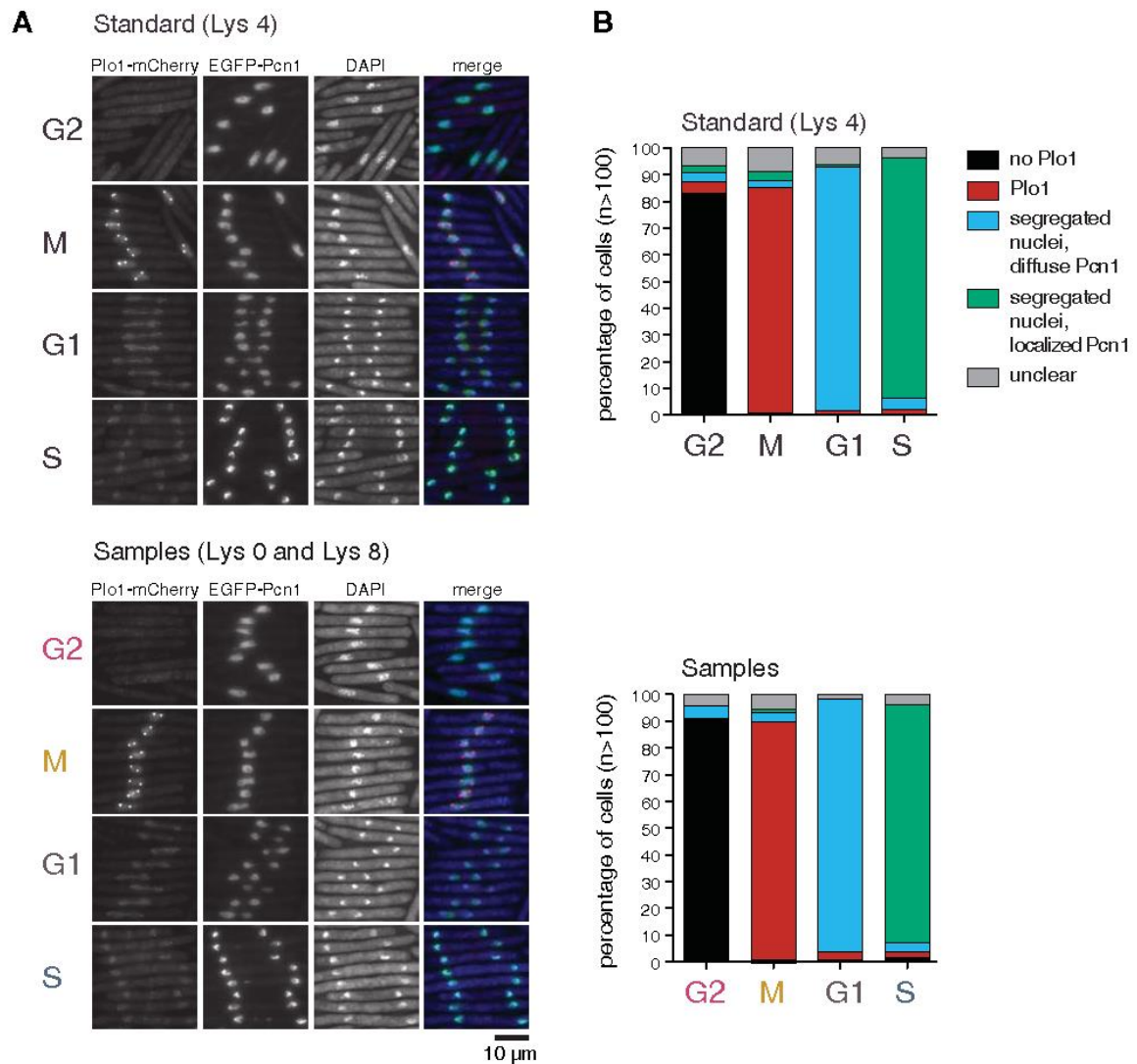
Supplementary Figure 1: Cell synchronization. *(A) Fluorescence microscopy for the super-SILAC standard and the experimental samples using Plo1 kinase tagged with mCherry, proliferating cell nuclear antigen Pcn1 tagged with EGFP and DAPI staining to label DNA. (B) Synchronicity for all analyzed cell cycle phases based on manual counting.*
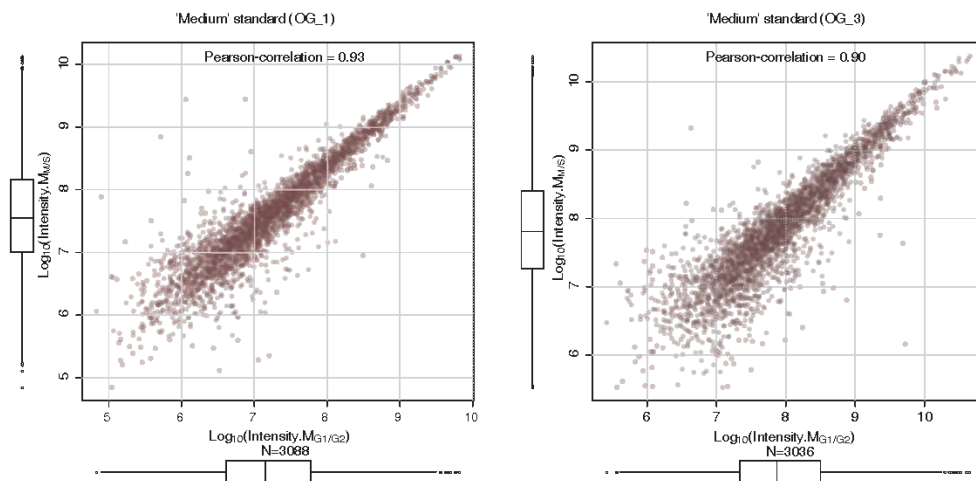
Supplementary Figure 2: Comparison of super-SILAC standards. Plot showing the correlation of the intensity of proteins identified in the super-SILAC standard from the G1G2 experiment against the intensity of proteins from the MS experiment in the log10 scale for both replicates.
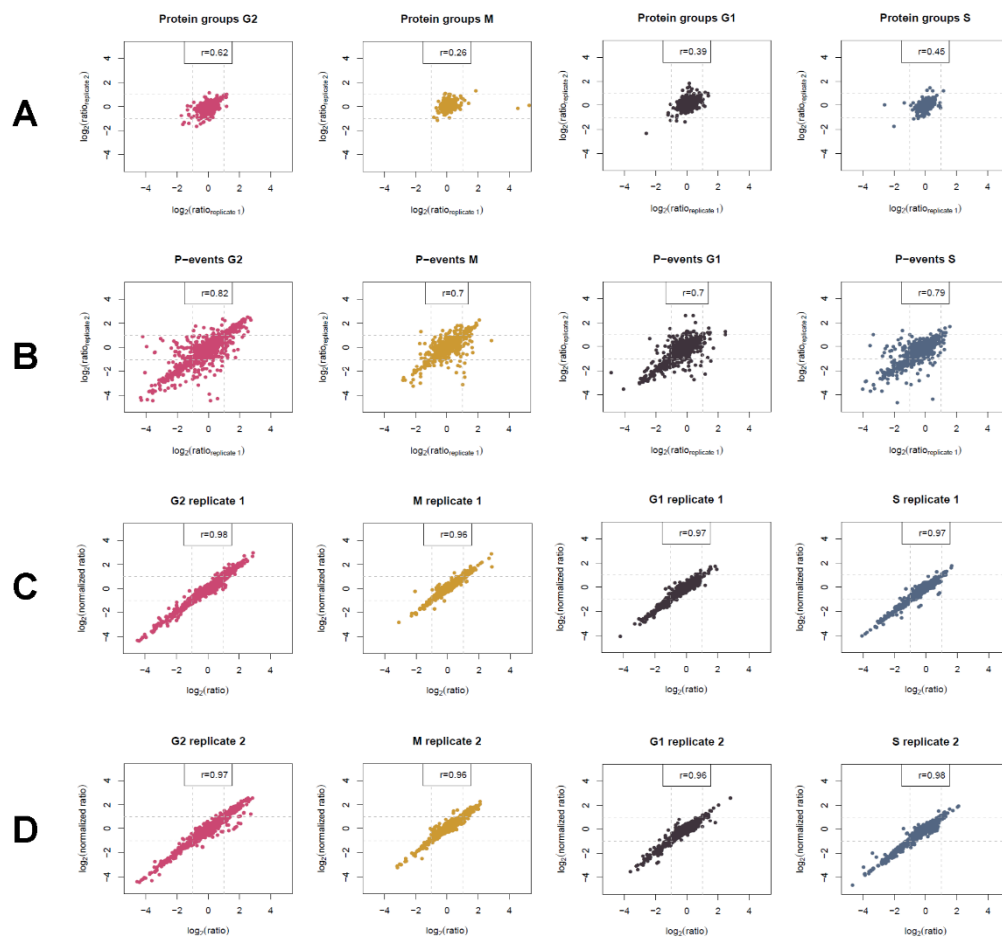
Supplementary Figure 3: Comparison of the SILAC ratios and of phosphopeptide ratios before and after normalization for both replicates. The ratios of both replicates for the studied cell cycle phases were compared for (A) protein groups and (B) phosphorylation events. The Pearson's correlation coefficient for proteome measurements is low because most protein groups have low variance; nevertheless high reproducibility was observed. The effect of protein abundance on phosphorylation ratios was compared by plotting the phosphorylation events ratios before (x-axis) and after (y-axis) normalization by protein abundance for both (C) replicate 1 and (D) replicate 2.
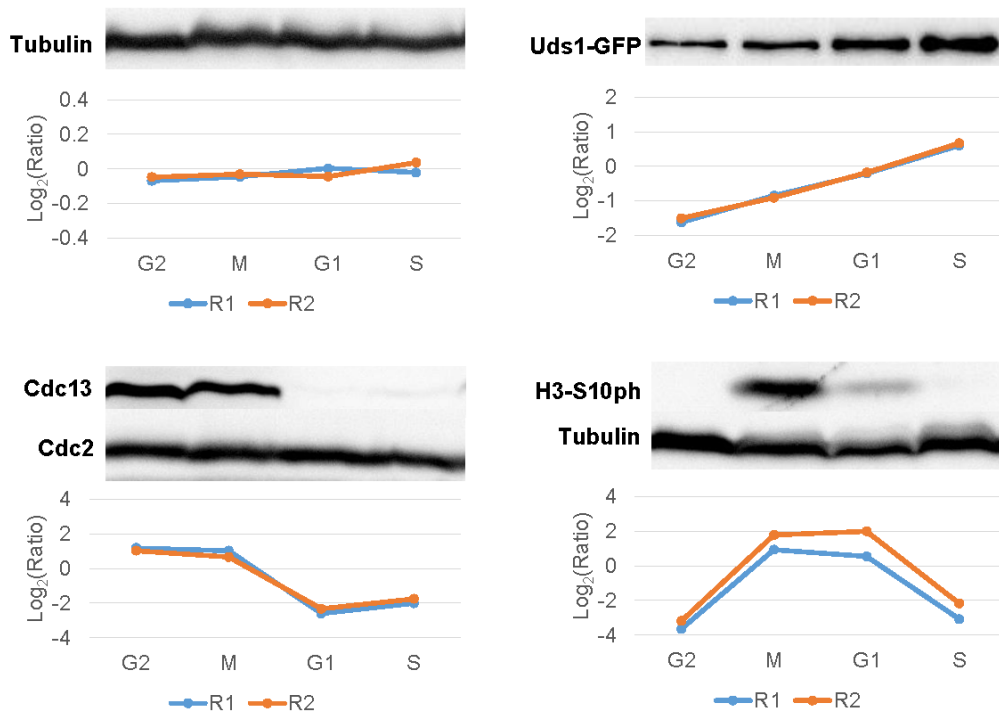
Supplementary Figure 4: Correlation of immunoblots with MS data. *Immunoblots of tubulin, Uds1-GFP, Cdc13 (cyclin B), S10-phosphorylation on histone H3, and their correlation with MS data in both replicates. Cdc2 servers as loading control. The tested proteins showed good agreement to the MS data, supporting the high quality of quantification across cell cycle phases.*
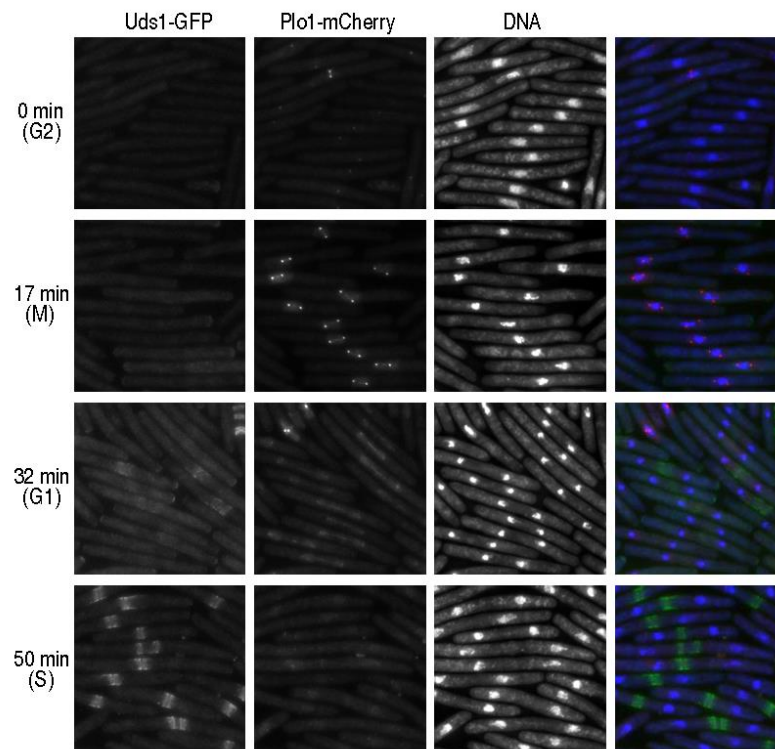
Supplementary Figure 5: Correlation of fluorescence microscopy with MS data. *Fluorescence microscopy of Uds1-GFP, Plo1-mCherry (which marks mitotic cells) and the DNA-stain DAPI across the studied cell cycle phases. In accordance with immunoblot and MS data (Supplementary Fig. 4), Uds1 (up-regulated during septation) increased in abundance from G2 to S, where it marks the region of septation. These data are in agreement with Bicho et al.*

Supplementary Figure 6: Functional enrichment analysis. *Functional enrichment analysis of the (A) proteome data, (B) phosphorylation data and (C) canonical kinase motifs performed separately for the different classes of fluctuating proteins. The canonical kinase motifs were extracted from the Human Protein Reference Database. Functional enrichment analysis of (D) proteins and (E) phosphorylation sites in the cell cycle phase-specific clusters.*

Supplementary Figure 7: Absolute quantification validation. *(A) The linear regression used for absolute quantification was obtained by plotting the MS signals against known concentrations of the UPS2 protein standard. (B) Protein abundance in this study against abundance of selected proteins measured by fluorescence correlation spectroscopy in a study by Heinrich et al. (C) Comparison of protein abundance obtained in this study to values reported in Marguerat et al. The MA-plot in the left panel depicts the average (x-axis) plotted against the difference (y-axis) in log₂-scale between measured protein abundance. The scatterplot in the right panel shows the correlation of protein abundance (log₁₀-scale) together with a trendline obtained by linear regression. Small blue lines at the axes represent the abundance of proteins exclusively detected in one of the studies. (D) Comparison to Gunaratne et al. (E) Comparison to Kulak et al. (F) Comparison to Wu et al.*
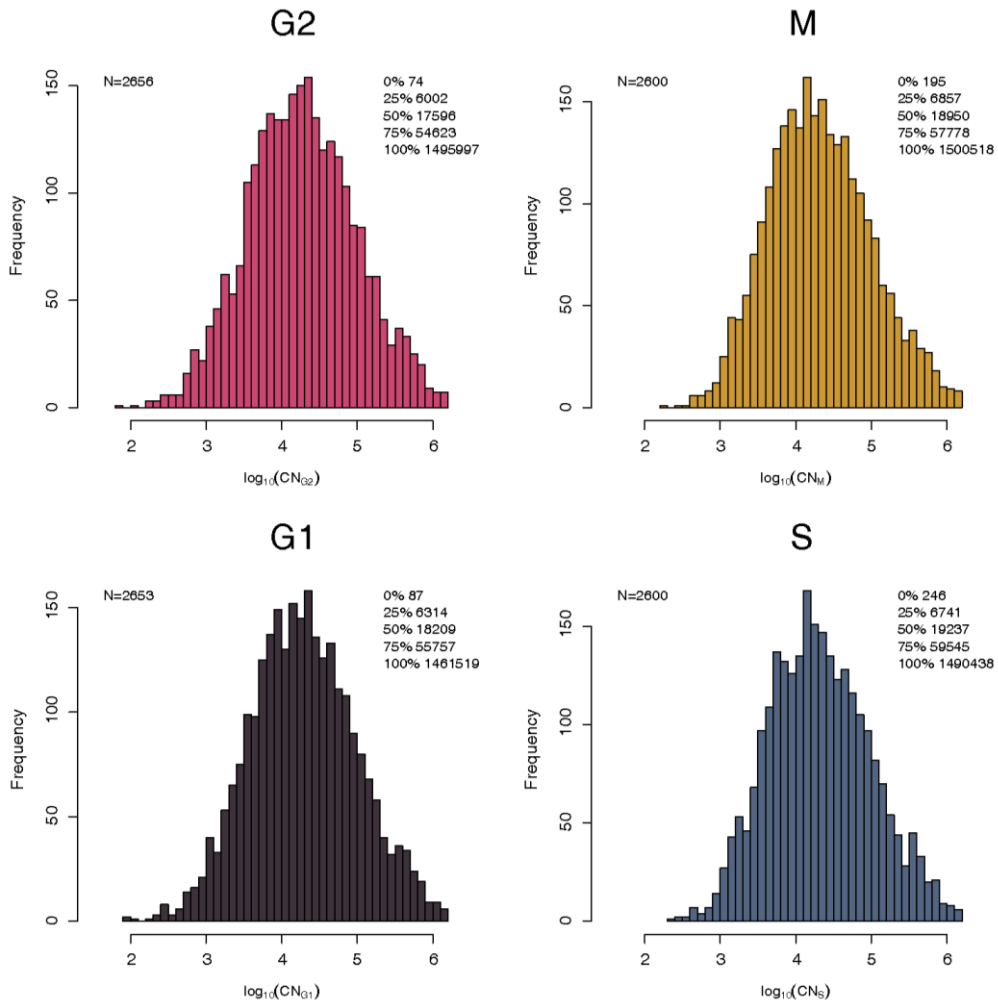
Supplementary Figure 8: Protein copy number distribution. *Histograms of protein abundance in the studied cell cycle phases.*
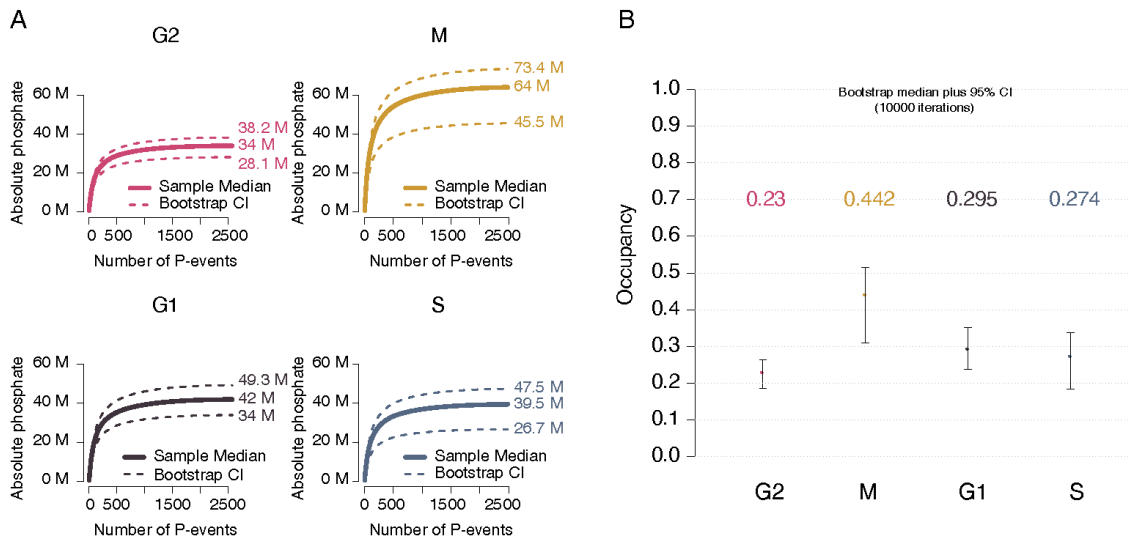
Supplementary Figure 9: Protein bound phosphate. *(A) Cumulative diagram of absolute protein bound phosphate groups for the separate cell cycle phases with bootstrap confidence intervals. (B) Median occupancy values resulting from 10,000 bootstrap iterations. The error bars correspond to a 95% bootstrap confidence interval used to calculate the lower and upper limits as shown in (A).*
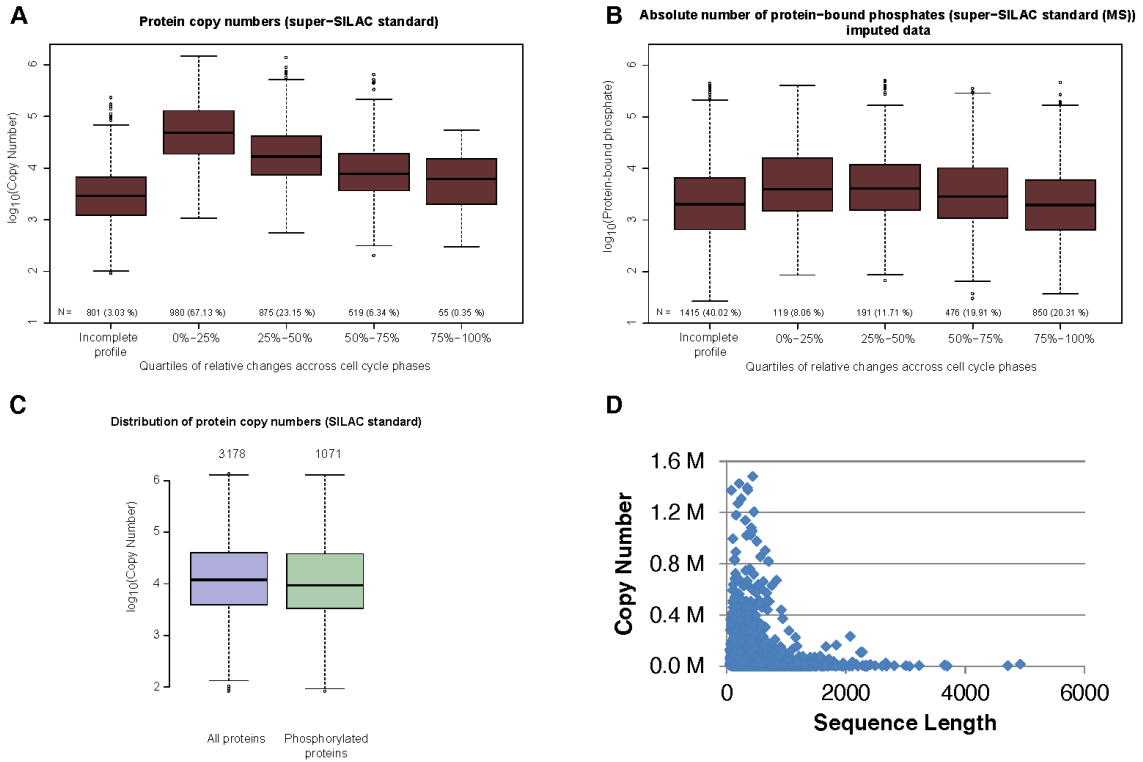
Supplementary Figure 10: Absolute quantification. *(A) Box plots of the protein copy number distribution and (B) absolute protein bound phosphate groups in the quartiles (bins) of standard deviations, the number in brackets is the contribution of each quartile in percent to the absolute amount of protein in the cell. (C) Box plot diagram of the distribution of protein abundance for all protein and phosphorylated proteins. (D) Dependence of sequence length and protein abundance.*

4. *Nic1* inactivation enables SILAC labeling with $^{13}C_6$$^{15}N_4$-Arginine in *S. pombe* (Manuscript III)

Alejandro Carpy, Avinash Patel, Tay Ye Dee, Iain Hagan, Boris Macek

4.1 Abstract

The inability to use $^{13}C_6$$^{15}N_4$-arginine in SILAC-based proteomics limits the exploitation of SILAC technology in fission yeast. The guanidinium group of $^{13}C_6$$^{15}N_4$-arginine is catabolized to $^{15}N_1$-ammonia that is used as a precursor for general amino acid biosynthesis. We show that disruption of Ni^{2+}-dependent urease activity, through deletion of the sole Ni^{2+} transporter Nic1 in ammonium supplemented medium blocks this re-cycling to enable $^{13}C_6$$^{15}N_4$-arginine labeling for SILAC strategies in *S. pombe*.

4.2 Introduction

Stable Isotope Labeling by Amino acids in Cell culture (SILAC) is one of the most widely used methods in quantitative proteomics (1). It involves *in vivo* metabolic labeling of cell cultures (or small organisms) with different versions of stable isotope-labeled essential amino acids (2). To maximize the number of peptides that can be quantified after proteome digestion with trypsin, proteins are usually differentially labeled with different forms of lysine and arginine (3): L- lysine (K0) and L-arginine (R0), $^2H_4$-lysine (K4) and $^{13}C_6$-arginine (R6), or $^{13}C_6$-$^{15}N_2$–lysine (K8) and $^{13}C_6$-$^{15}N_4$-arginine (R10). In yeast cells, where lysine and arginine are not essential, the use of auxotrophic mutants allows complete labeling of all tryptic peptides (4, 5). SILAC has been used in quantitative proteomics in several yeast species, but most widely in *Saccharomyces cerevisiae* (6) (budding yeast) and *Schizosaccharomyces pombe* (fission yeast)*. S. pombe* is extensively exploited to study cell cycle control (7), heterochromatin (8) and differentiation (9) and is increasingly the subject of large-scale quantitative proteomics studies (10, 11). One of the major challenges faced when applying SILAC to fission yeast is metabolic conversion of arginine to other amino acids such as proline, glutamine and lysine (5). This results in the partial labeling of the respective amino acids and produces a mass spectrum with a complex isotope cluster that makes the downstream analysis challenging and error-prone. However, inactivation of the "arginine conversion pathway" by removal of the orthinine transferase, Car2, overcomes this problem to enable the use of arginine labeling in SILAC-based experiments

(5). Although the use of *car2.Δ* has facilitated the application of SILAC technology in fission yeast, the choice of amino acids that can be exploited remains limited. Only one form of heavy arginine (R6) can be used (12) alongside three forms of heavy lysine (K4, K6 and K8) (5, 13). Surprisingly, we could not find any instance of reported use of arginine (R10) in fission yeast in the literature although this is a common reagent for labeling other cell types (1).

4.3   Materials and methods

Yeast strains

Strain No.	Genotype	Source
IH 5224	h⁺ pku70::kan ura4.d18 leu1.32 ade6.704	Lab Stock
IH 6113	pku80::ura⁺ ura4.d18 leu1.32	Lab Stock
IH 8988	h⁺ car2::NatMX6	This study
IH 8987	h⁻ car2::NatMX6	This study
IH 6297	h⁻ arg3.D4	Lab Stock
IH 1540	h⁺ lys1.131	Lab Stock
IH 9091	h⁻ car2::NatMX6 lys1.131 arg3.D4 ura4.d18 leu1.32	This study
IH 9092	h⁺ car2::NatMX6 lys1.131 arg3.D4 ura4.d18 leu1.32	This study
IH 9281	plo1.ts41:ura4+ cut7.24 lys1.131 arg3.D4 car2::NatMX6 ura4.d18 leu1.32	This study
IH 11849	h⁻ nic1::KanMX6 car2::NatMX6 lys1.131 arg3.D4	This study

Standard yeast culture

Strains used in this used study are listed above. Cell culture and maintenance were according to Moreno et al. (14). The amino acid supplements were uracil, adenine, leucine and histidine at a concentration of 200μg/mL. Yeast strains were streaked onto YES plates (Solid Yeast Extract agar with supplements) from the frozen glycerol stocks at -80 °C, and grown at the permissive temperature (25 °C) until single colonies formed. One single colony was used to start liquid cultures. For all the physiological experiments, asynchronous cultures of cells were grown in variants of Edinburgh minimal media 2 (EMM2), with appropriate supplements from log phase starter cultures ($1 \times 10^6 - 5 \times 10^6$ cells/mL). In EMMG, the 5g/l ammonium chloride of EMM2 is replaced with 5g/l monosodium glutamate (15). In EMMGn, EMMG is supplemented with 12 mM NH₄Cl.

S. pombe transformation using lithium acetate

Cultures that had been grown in YES (nutrient rich) to mid-log phase ($4 \times 10^6 - 8 \times 10^6$ cells/mL) were pelleted by centrifugation at 3400 g for 2 min, washed once in water followed by re-suspension in appropriate volume of lithium acetate (0.1 M, pH 4.9) to achieve 2×10^8 cells/ml.

After 1 hour at 25 °C 10-15 µg of DNA was added to 100 µl of competent cells alongside 2.9 volumes of PEG 4,000 (50% (w/v) in 0.1 M, pH 4.9 lithium acetate). After 1 hour at 25 °C and 15 minutes at 42 °C cells were washed once in water and plated onto selective minimal media.

SILAC-Labeling

L-lysine:2HCl (4,4,5,5-D4, 96-98%, Cambridge Isotopes, cat no. DLM-2640-0.5) and L-arginine:HCl (U-13C6, 99%, Cambridge Isotopes, cat no. CLM-2265-H-0.5) were used as medium isotopic labels; L-lysine:2HCl (U-13C6, 99%; U-15N2, 99%, Cambridge Isotopes, cat no. CNLM-291-H-0.25) and L-arginine:HCl (U-13C6, 99%; U-15N4, 99%, Cambridge Isotopes, cat no. CNLM-539-H-0.5) were used as heavy isotopic labels.

Cell Culture

EMMG media is equivalent to EMM2(16) except that 5 g/l ammonium chloride is replaced with 5 g/l mono-sodium glutamate (15). EMM2 media was also prepared with a range of ammonium chloride concentrations up to 96 mM. EMMGn is EMMG containing 12 mM NH_4Cl. To label the cells with heavy or medium isotopic labels for Arginine (R) and Lysine (K), starter cultures were prepared by inoculation into EMMG/EMMGn supplemented with 37.5 µg/mL each of light and medium/heavy isotopic versions of arginine and lysine. After inoculation from log phase starter cultures, main cultures were grown in EMMG/EMMGn supplemented with 75 µg/mL of heavy isotopic labeled R and K for at least 10 generations before sampling to ensure the maximum labeling of the proteome. Care was taken that the culture never exceeded the mid-log phase $(1 - 5 \times 10^6)$ at which they were ultimately harvested. For direct comparisons with the EMM2 experiments of Bicho et al. (5) it was necessary to make starter cultures in YES media before inoculation into EMM2 as we were repeatedly unable to get growth in main cultures of either the *SILAC* strain when both the starter and main culture were made in minimal selective medium.

SILAC strain

The 100 bp primer pair (CAR2_FORW, AAGTCAGATTATCAAAGGAGTTAAGAGATTATTGTGTAATCTTGAATCAGTTTTCTTTAGGTCTTAGTAAATTCTTCATTCGGATCCCCGGGTTAATTAA; CAR2_REV, TGTAGTCATCGAACCCTTAACATTAGAAGGGAGGTAAGTAAGAATGAATCGTCTAGTTTTTTTTGTTTAAAGCATTGTTTTGAATTCGAGCTCGTTTAAAC) were designed such that the first 80 nucleotides were homologous to the sequences flanking the *car2+* gene and the last 20 nucleotides were homologous to the pYM14 template plasmid with the *natMX6* marker (*pFA6natMX6*)(17). The PCR product was transformed into the strain 5224 (*pku70::kan ura4.d18 leu1.32 ade6.704*) and colonies that were resistant to Nourseothricin/ClonNat (Werner BioAgents, 96736-11-7) were selected. After PCR analysis

confirmed the desired integration event two rounds of back- crossing to wt removed the *pku70::kan* before arginine and lysine auxotrophic markers *arg3.d4* and *lys1.131 were* introduced to generate the *SILAC* strain (*car2::natMX6 arg3.d4 lys1-131*; IH 9091 and 9092).

SILACn strain

The 100 bp primer pair (NIC1_FORW, AAAATTGTTATTATTATTGGAATCCTAAGTGGATACAGGTTTATGTGAACGCAATCTATTCAAATTACGCGTTTATTTGTTTAATTAAGGCGCGCCAGAT; NIC1_REV, ACGACTGGGAATTTCTGTTTCCTTTTTTTTTTCTTCAGTATTCTCGATTTGTTCTCTTAACCTGTTGCTTGGGTTAAAAGTGTTTAAACTGGATGGCGGCG) were designed such that the first 80 nucleotides were homologous to the sequences flanking the *nic1+* gene and the last 20 nucleotides were homologous to the pYM14 template plasmid with the *natMX6* marker (*pFA6natMX6*) (5, 17, 18). The PCR product was transformed into the strain IH6113 (*pku80::ura+ ura4.d18 leu1.32*) and colonies that were resistant to Geneticin (MP Biomedicals, 158782) were selected. These were further confirmed for the integration at the desired locus by two rounds of back- crossing to wt before crossing into the *SILAC* base strain generated the *SILACn* base strain (*h- nic1::KanMX6 car2::NatMX6 lys1.131 arg3.D4*; IH 11849).

Cell extracts for proteomics experiments

5×10^7 cells in log phase of growth were harvested at 4000 rpm for 2 minutes in screw cap 1.5 mL microfuge tubes. Cell pellets were washed once in ice-cold STOP buffer (10 mM EDTA, 50 mM NaF, 150 mM NaCl, 1 mM NaN$_3$) and then flash frozen in liquid nitrogen and kept in -80°C. Proteins were extracted by breaking the cells with 200 µL acid washed glass beads in the presence of 200 µL of denaturation buffer (6M Urea, 2M Thio-urea in 10mM Tris buffer pH 8.0) in a ribolyser (Fast prep FP120-BIO101, ThermoSavant) used for 30 sec at 6.5 rpm. Then 200 µL of denaturation buffer was added to the glass beads and the tubes were incubated on a rotating rocker at RT for 30 mins. In the bottom of each tube a hole was made and each tube was placed on the top of the empty 1.5 mL microfuge tube and spun at 4800 rpm for 2 min at RT. The insoluble debris was separated by centrifugation at 13,000 rpm for 10 min at RT. The supernatant was collected in a fresh 1.5 mL microfuge tube and comprised the sample that would be processed for LC-MS/MS. Protein concentration was measured using Bradford.

In-solution protein alkylation and digestion

Reduction buffer [1 M dithiothreitol (DTT) in 50 mM ammonium bicarbonate (ABC)] was added to the sample to a final concentration of 1mM DTT; and incubated at room temperature (RT) for 1 hour. After reduction, alkylation buffer [550mM iodoacetamide (IAA) in 50 mM ABC] was added to the sample to a final concentration of 5.5 mM IAA and incubated at RT for 1 hour in

the dark. Sample was then digested with endopeptidase LysC (Waco chemicals, 1:100 enzyme:protein) at RT for 3 hours. After digestion with LysC at high urea concentration, the sample was diluted to a final concentration of 1.2 M urea / 0.4 M thiourea with MiliQ water and digested with trypsin (Promega, 1:100 enzyme:protein) at RT overnight. The samples were acidified to inactivate trypsin by adding 1% TFA and desalted using C18 StageTips (19).

MS analysis

Desalted peptide fractions were separated using a nano-LC (Proxeon Biosystems) coupled to either an Orbitrap Elite or a LTQ Orbitrap XL (Thermo Fisher Scientific) mass spectrometer (MS). Peptides were separated on a 15 cm PicoTip fused silica emitter with a 75 µm ID (New Objective) packed with ReproSil-Pur C18-AQ 3 µm resin (Dr. Maisch GmbH) in-house. Peptides were loaded in the column with Buffer A (0.5% acetic acid) at 700 nL/min using a maximum pressure of 280 Bar. Peptides were then eluted using a 57 or 87 min segmented gradient of 5-50% Buffer B (80% ACN in 0.5% acetic acid) at a flow rate of 200 nL/min. The mass spectrometers were operated in the positive ion mode. Spectra were recorded with data-dependent mode using a dynamic exclusion of 90 sec (LTQ Orbitrap XL) or 60 sec (Orbitrap Elite). Resolution was set at 60,000 (LTQ Orbitrap XL) or 120,000 (Orbitrap Elite) with an accumulation target value of 1E6 charges. For peptide analysis the 10 (LTQ Orbitrap XL) or 20 (Orbitrap Elite) most intense peaks were selected from the survey scan (300-2,000 Thompson) and fragmented with collision induced dissociation at a target value of 5,000 charges and recorded in the linear ion trap.

MS data processing

The MS data were analyzed with the MaxQuant Software Package version 1.3.0.5(20) and searched using the Andromeda peptide search engine (21). The MS/MS spectra were searched against a database containing the proteome sequence of the *S. pombe* which was downloaded from http://www.pombase.org (03.01.2013). The database contained 5,145 entries from the S. pombe proteome, 248 frequently observed contaminants as well as the reverse version of all sequences. Mass tolerance was set to 6 ppm at the MS and 0.5 Da at the MS/MS level. A minimal length of seven amino acids and full tryptic enzyme specificity were required. Two enzyme missed cleavages were allowed. Oxidation of methionine and protein N-terminal acetylation were defined as variable modifications and carbamidomethylation of cysteins was used as fixed modification. False discovery rate was set to 1% on the peptide and protein level.

Immunofluoresence for microtubules and spindle pole body

Immunofluorescence was performed as described by Hagan and Yanagida (22). The Sad1 (spindle pole body component) and 1:80 TAT1 (Trypanosome Alfa Tubulin) (23) antibodies were diluted 1:25 and 1:80 respectively and used as primary antibodies. FITC anti-mouse antibody (Sigma Aldrich) for TAT1 and Cy3 anti-rabbit antibody (Sigma Aldrich) for Sad1 in PEMBAL were used as secondary antibodies at a dilution of 1:100 and 1:2500 respectively.

4.4   Results and discussion

To assess the feasibility of using arginine (R10) for quantitative proteomics of fission yeast, we performed SILAC labeling of *S. pombe* cells as described by Bicho et al (5). We used *arg3.d4 lys3.131 car2::natMX6* background and refer to this genotype as *"SILAC"* in the following text. Consistent with previous observations (5), the *SILAC* strain grows poorly in the EMM2 medium that uses ammonium as a nitrogen source (**Figure 1a**). Ammonium ions inhibit one of two arginine uptake pathways, to severely compromise the growth of arginine auxotrophs such as
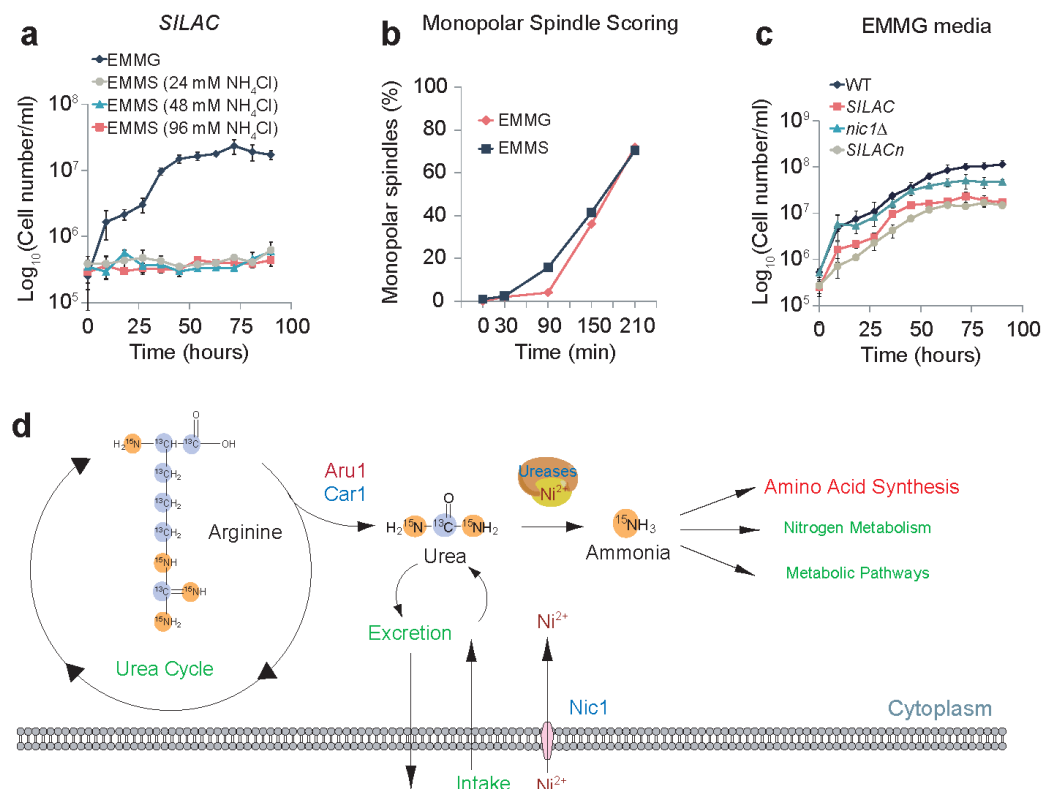
Figure 1.   **Metabolic characterization of S. pombe.** (a) Growth curve of the SILAC strain in EMMG and EMMS media supplemented with different concentrations of ammonium chloride. (b) Monopolar spindle scoring, plo1.ts41 cut7.24 SILAC cells were grown in the EMMG and EMMS media at 25oC to a cell density of 1 x 106. Cells were shifted to 37oC and immunostained to score for monopolar spindles, a phenotype associated with plo1.ts41. (c) Growth curve of WT, SILAC, nic1☐ and SILACn strains in EMMG media. (d) Metabolism of arginine in S. pombe. For full description of genotype of mutant strains see methods section.

the *SILAC* strain (24). Substitution of ammonium for monosodium glutamate in the medium EMMG enables both arginine import pathways to be used. Our experimental approach therefore deviated from that of Bicho by growth in EMMG rather than the reduced ammonium EMM2 used in their study. Monitoring the formation of monopolar spindles in a *plo1.ts41 cut7.24 double* mutant (25, 26) established that this switch of nitrogen source did not appear to alter phenotypic expression in the SILAC mutant background (**Figure 1b**). Furthermore, the *SILAC* strain growth in EMMG media is comparable to that of the WT (**Figure 1c**).

While performing a standard triple SILAC experiment using "light" (K0R0), "medium" (K4R6) and "heavy" (K8R10) amino acids in EMMG we noticed that almost all heavy (K8R10)-labeled peptides had complex isotope clusters that impede the identification of heavy peptides and proteins. Importantly, this complex isotope clusters were present in both arginine- and lysine-
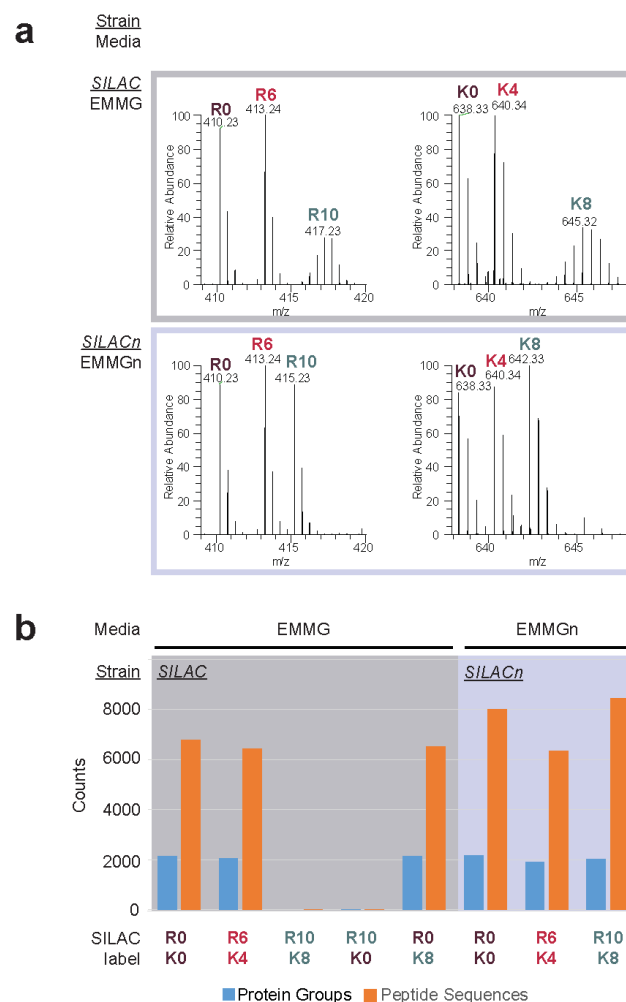
Figure 2. *Assessment of arginine conversion using MS. Isotopic distribution of peptides from a triple-SILAC experiment for an arginine and a lysine containing peptide using the SILAC strain (top) and an arginine and a lysine containing peptide using the SILACn strain (bottom). (b) Peptide sequences identified in individual experiments using different isotopic versions of lysine and arginine. The lines on the top indicate the media used and the colored background indicates the strain used.*

containing peptides (**Figure 2a**), pointing to a global conversion of the stable isotope label introduced by heavy arginine and/or lysine. To pinpoint the heavy amino acid responsible for the conversion, we measured samples labeled with a single heavy label (R0K8 or R10K0). Only the sample containing heavy arginine (R10K0) presented the complex isotope cluster observed before and the identification rate was reduced (**Figure 2b;** Summary of the proteins identified in each experiment is presented in the **Supplementary Tables 1 - 12**).

We reasoned that this extensive mislabeling of multiple amino acids probably arose from catabolism of heavy arginine (R10) to heavy urea. Catalysis by Car1 and Aru1 would convert the guanidine group from arginine into urea which could be further catabolized by ureases into ammonia (**Figure 1d**) **(5)**. The $^{15}N_1$-ammonia derived from $^{15}N_2$-urea would then be incorporated into multiple amino acids, that each contains different numbers of ^{15}N atoms to alter the isotope distributions of all peptides. Unlike the medium-heavy $^{13}C_6$-arginine (R6), only heavy $^{15}N_4^{13}C_6$-arginine (R10) has isotopically labeled nitrogen and so only samples labeled with arginine (R10) are vulnerable to this urease catabolism problem. Consistently, in measurements of samples containing only medium-heavy amino acids arginine (R6) and lysine (K4), no such conversion was observed (**Figure 2b**).

Ureases require Ni^{2+} as the major cofactor for their catalytic activity. *S. pombe* encodes one characterized urease (Ure1*)* (27) alongside several other proteins that are predicted to have a urease activity. Nic1 is the sole plasma membrane Ni^{2+} transporter and supports urease activity of fission yeast (28). We therefore reasoned based on observations from Eitinger et al. (28), that global reduction of the urease activity by abolition of Nic1 activity would reduce amino acid mislabeling when the *SILAC* strain was grown on heavy arginine (R10). We therefore deleted *nic1⁺* in the *SILAC* strain to generate "*SILACn*" before repeating the SILAC labeling as described above. *nic1.Δ* introduction resulted in an increase in the identification of arginine (R10)-labeled peptides, suggesting that urease-dependent arginine catabolism was indeed a key factor in arginine (R10) mislabeling (**Supplementary Figure 1**). However, the quantity of peptides identified in samples labelled with heavy amino acids was still lower than in samples containing light amino acids and the isotopic distribution of peptides was unusual, suggesting a partial incorporation of ^{15}N atoms. The incomplete suppression suggested that residual traces of Ni^{2+} persist in *nic1.Δ* cells or that the urea is catabolized to ammonium ions by Ni^{2+} independent urease/pathways.

The increased identification rate of (R10) labeled peptides in the *SILACn* background indicated that, although ^{15}N-ammonium has not been completely eliminated by removal of Nic1, its levels were severely reduced. We therefore asked whether we could outcompete the exploitation of

these low levels of ^{15}N-ammonium ions in amino acid anabolic pathways by supplementing the EMMG growth medium with unlabeled ^{14}N-ammonium ions that are tolerated by SILAC strains(5). The *SILACn* strain grew in EMMG media that contained 12mM NH$_4$Cl (EMMGn) alongside stable isotope-labelled amino acids with a doubling rate of 7 hours. For reasons that are not clear, the original *SILAC* strain (*nic1$^+$*) grew very poorly in EMMGn (data not shown). Importantly, however, the regular isotope pattern of the heavy-labeled peaks established that growth of *SILACn* cells in EMMGn successfully resolved the mislabeling problem (**Compare figure 2a top spectra with bottom spectra**). This improved detection rate enabled us to identify comparable numbers of heavy labeled peptides as detected in cells labeled with light or medium amino acids (**Figure 2b**).

We conclude that removal of Nic1 from cells grown in EMMGn finally renders fission yeast fully compatible with SILAC labeling to greatly expand the options for interrogation of the *S. pombe* proteome*.* The ability to routinely conduct triple SILAC analysis in this genetically malleable model organism will rapidly advance our quantitative understanding of many core principles of molecular cell biology.

4.5   Contributions

Avinash Patel and I exposed the mislabeling problem, while realizing label incorporation checks. With the help of Boris Macek and Iain Hagan, we reasoned the pathway responsible for the mislabeling and designed experiments. Avinash Patel and Tay Ye Dee transformed and grew the cells. I realized the protein extraction, LC-MS/MS measurements and analyzed the data. Figures were also prepared by myself. I wrote the manuscript with Boris Macek and Iain Hagan with support from all the authors. My contribution to this manuscript was about 50%.

4.6 References

1. Ong, S. E., and Mann, M. (2005) Mass spectrometry-based proteomics turns quantitative. *Nature Chemical Biology* 1, 252-262

2. 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. *Mol Cell Proteomics* 1, 376-386

3. Olsen, J. V., Blagoev, B., Gnad, F., Macek, B., Kumar, C., Mortensen, P., and Mann, M. (2006) Global, In Vivo, and Site-Specific Phosphorylation Dynamics in Signaling Networks. *Cell* 127, 635-648

4. de Godoy, L. M. F., Olsen, J. V., de Souza, G. A., Li, G. Q., Mortensen, P., and Mann, M. (2006) Status of complete proteome analysis by mass spectrometry: SILAC labeled yeast as a model system. *Genome Biology* 7

5. Bicho, C. C., Alves, F. D., Chen, Z. A., Rappsilber, J., and Sawin, K. E. (2010) A Genetic Engineering Solution to the "Arginine Conversion Problem" in Stable Isotope Labeling by Amino Acids in Cell Culture (SILAC). *Molecular & Cellular Proteomics* 9, 1567-1577

6. de Godoy, L. M., Olsen, J. V., Cox, J., Nielsen, M. L., Hubner, N. C., Frohlich, F., Walther, T. C., and Mann, M. (2008) Comprehensive mass-spectrometry-based proteome quantification of haploid versus diploid yeast. *Nature* 455, 1251-1254

7. Elledge, S. J. (1996) Cell cycle checkpoints: Preventing an identity crisis. *Science* 274, 1664-1672

8. White, S. A., and Allshire, R. C. (2008) RNAi-mediated chromatin silencing in fission yeast. *Current topics in microbiology and immunology* 320, 157-183

9. Harigaya, Y., and Yamamoto, M. (2007) Molecular mechanisms underlying the mitosis-meiosis decision. Chromosome research : an international journal on the molecular, supramolecular and evolutionary aspects of chromosome biology 15, 523-537

10. Marguerat, S., Schmidt, A., Codlin, S., Chen, W., Aebersold, R., and Bahler, J. (2012) Quantitative analysis of fission yeast transcriptomes and proteomes in proliferating and quiescent cells. *Cell* 151, 671-683

11. Gunaratne, J., Schmidt, A., Quandt, A., Neo, S. P., Sarac, O. S., Gracia, T., Loguercio, S., Ahrne, E., Xia, R. L., Tan, K. H., Lossner, C., Bahler, J., Beyer, A., Blackstock, W., and Aebersold, R. (2013) Extensive mass spectrometry-based analysis of the fission yeast proteome: the Schizosaccharomyces pombe PeptideAtlas. *Molecular & cellular proteomics : MCP* 12, 1741-1751

12. Koch, A., Krug, K., Pengelley, S., Macek, B., and Hauf, S. (2011) Mitotic substrates of the kinase aurora with roles in chromatin regulation identified through quantitative phosphoproteomics of fission yeast. *Science signaling* 4, rs6

13. Koch, A., Krug, K., Pengelley, S., Macek, B., and Hauf, S. (2011) Mitotic Substrates of the Kinase Aurora with Roles in Chromatin Regulation Identified Through Quantitative Phosphoproteomics of Fission Yeast. *Sci. Signal.* 4

14. Moreno, S., Klar, A., and Nurse, P. (1991) Molecular genetic analysis of fission yeast Schizosaccharomyces pombe. *Methods Enzymol* 194, 795-823

15. Fantes, P., and Nurse, P. (1977) Control of cell size at division in fission yeast by a growth-modulated size control over nuclear division. *Experimental cell research* 107, 377-386

16. Nurse, P. (1975) Genetic control of cell size at cell division in yeast. *Nature* 256, 547-551

17. Hentges, P., Van Driessche, B., Tafforeau, L., Vandenhaute, J., and Carr, A. M. (2005) Three novel antibiotic marker cassettes for gene disruption and marker switching in Schizosaccharomyces pombe. *Yeast* 22, 1013-1019

18. Bahler, J., Wu, J. Q., Longtine, M. S., Shah, N. G., McKenzie, A., 3rd, Steever, A. B., Wach, A., Philippsen, P., and Pringle, J. R. (1998) Heterologous modules for efficient and versatile PCR-based gene targeting in Schizosaccharomyces pombe. *Yeast (Chichester, England)* 14, 943-951

19. Rappsilber, J., Mann, M., and Ishihama, Y. (2007) Protocol for micro-purification, enrichment, pre-fractionation and storage of peptides for proteomics using StageTips. *Nat. Protoc.* 2, 1896-1906

20. 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

21. Cox, J., Neuhauser, N., Michalski, A., Scheltema, R. A., Olsen, J. V., and Mann, M. (2011) Andromeda: A Peptide Search Engine Integrated into the MaxQuant Environment. *J. Proteome Res.* 10, 1794-1805

22. Hagan, I., and Yanagida, M. (1995) The product of the spindle formation gene sad1+ associates with the fission yeast spindle pole body and is essential for viability. *The Journal of cell biology* 129, 1033-1047

23. Woods, A., Sherwin, T., Sasse, R., MacRae, T. H., Baines, A. J., and Gull, K. (1989) Definition of individual components within the cytoskeleton of Trypanosoma brucei by a library of monoclonal antibodies. *Journal of cell science* 93 (Pt 3), 491-500

24. Fantes, P. A., and Creanor, J. (1984) Canavanine resistance and the mechanism of arginine uptake in the fission yeast Schizosaccharomyces pombe. *J Gen Microbiol* 130, 3265-3273

25. Grallert, A., Patel, A., Tallada, V. A., Chan, K. Y., Bagley, S., Krapp, A., Simanis, V., and Hagan, I. M. (2013) Centrosomal MPF triggers the mitotic and morphogenetic switches of fission yeast. *Nature cell biology* 15, 88-95

26. Hagan, I., and Yanagida, M. (1990) Novel potential mitotic motor protein encoded by the fission yeast cut7+ gene. *Nature* 347, 563-566

27. Lubbers, M. W., Rodriguez, S. B., Honey, N. K., and Thornton, R. J. (1996) Purification and characterization of urease from schizosaccharomyces pombe. *Canadian journal of microbiology* 42, 132-140

28. Eitinger, T., Degen, O., Bohnke, U., and Muller, M. (2000) Nic1p, a relative of bacterial transition metal permeases in Schizosaccharomyces pombe, provides nickel ion for urease biosynthesis. *The Journal of biological chemistry* 275, 18029-18033

4.7   Supporting information

Supplementary tables are contained in the CD attached.
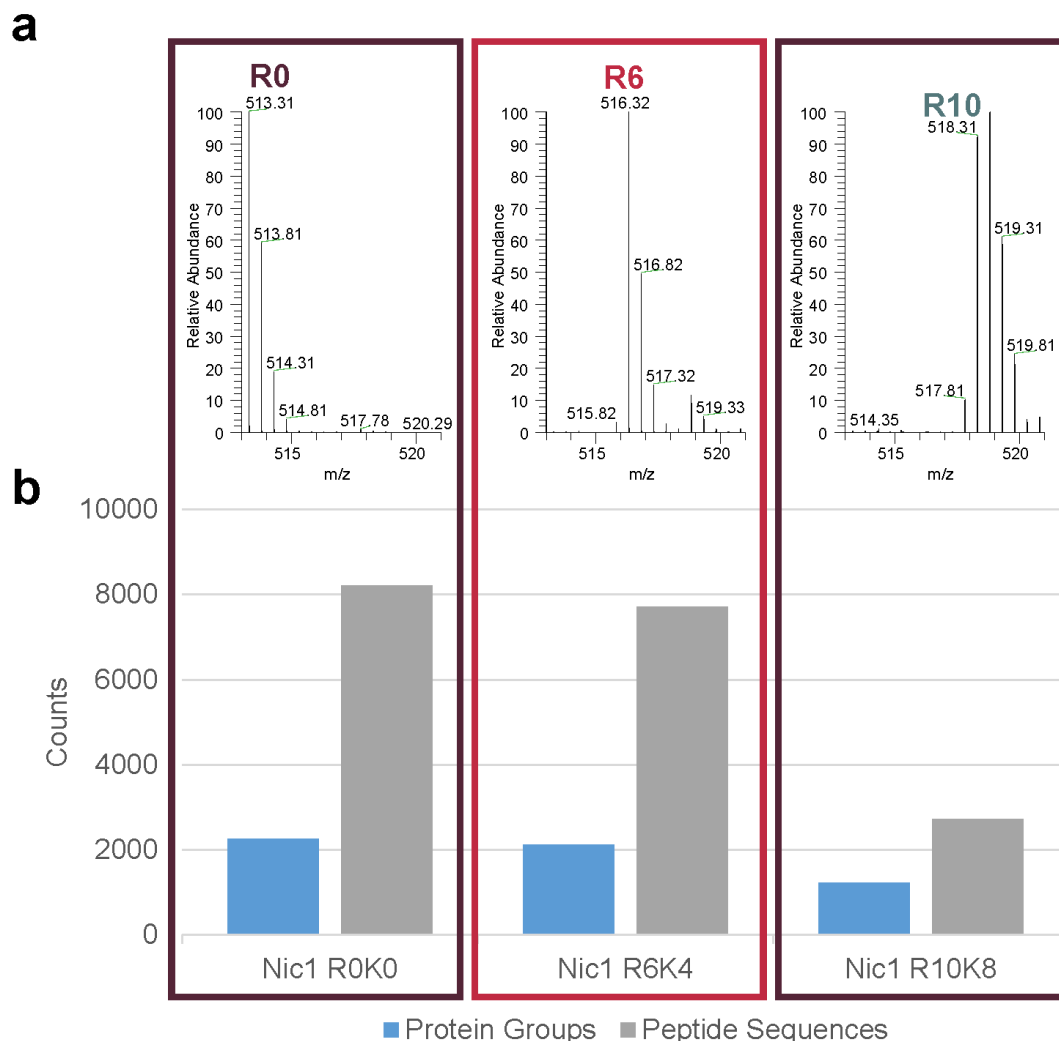
Supplementary figure 1. SILACn strain grown in EMMG media. *(a) Spectra from the isotopic distributions for an arginine containing peptide labelled with light, medium or heavy amino acids. (b) Peptide sequences and protein groups identified.*

5. Conclusions

In this thesis I developed, optimized and applied several methods in order to characterize fission yeast using MS-based quantitative proteomics. From this studies I can conclude the following:

1. Optimization of quantitative proteomics and phosphoproteomics methods
    a. The use of DHB can be efficiently avoided by using high concentrations of TFA and acetonitrile (ACN).
    b. The sequential enrichment of fractions containing high amount of phosphopeptides leads to higher number of phosphosite identifications.
    c. This approach is universal and can be applied to several organisms.

2. Study of the absolute proteome and phosphoproteome dynamics during the cell cycle of the fission yeast
    a. The approach used is adequate to characterize the fission yeast proteome with a high coverage. The 3,753 proteins identified in our study represent 74% of the predicted fission yeast proteins, making it one of the most comprehensive published protein dataset of this organism to date. Furthermore, this approach is adequate to characterize fission yeast since we were able to relatively quantify 80% of all identified proteins in at least one cell cycle phase and 65% of the proteins in all phases.
    b. Cell cycle regulation occur mainly via phosphorylation. Despite detecting many proteins for which the mRNA was shown to fluctuate during the cell cycle, the variations in protein abundance were low. At the phosphorylation level from 3,682 phosphorylation events, at least 30% were fluctuating. Our data also suggest that most cell cycle-dependent changes on the mRNA level are not essential for cell cycle progression.
    c. iBAQ can be used to absolute quantify fission yeast proteome. We estimated copy numbers of 3,178 proteins. Proteins that fluctuate across the cell cycle have a low abundance and represent only 0.35% of the total amount of protein in the cell (4th quartile) while fluctuating phosphorylation events are present on proteins of the entire abundance range.
    d. The use of SILAC-based relative quantification enabled us to address the stoichiometry of detected phosphorylation sites.
    e. In combination with absolute quantification it was possible to calculate the ATP dynamics during the cell cycle. 50 amol of ATP is required for protein STY phosphorylation during the G2/M transition. This corresponds to 23% of the

average amount of ATP in the cell. This approach can also be applied to other PTMs.

3. Solving the "heavy labelling" problem in fission yeast
    a. The removal of Nic1 from cells grown in EMMGn renders fission yeast fully compatible with SILAC labeling and greatly expands the options for interrogation of the *S. pombe* proteome. The ability to conduct triple SILAC analysis will allow understanding of many core principles of molecular cell biology.
    b. This study may point the way towards solving of arginine-related labeling problems in SILAC experiments in other species.

6. Appendix

Curriculum Vitae

Personal Information

Name	Alejandro Carpy Gutierrez Cirlos
Date of Birth	February 20th, 1985
Nationality	Mexican / German

Education

10/2010 – 07/2010	**PhD Student in Biochemistry**
	Proteome Center Tübingen, University of Tübingen, Germany
10/2008 – 09/2010	**MSc in Applied Chemistry**, Reutlingen University, Germany
	Thesis: «Development and Optimization of a Capillary Zone Electrophoresis Method for the Routine Analysis of Monoclonal Antibodies in Quality Control»
10/2005 – 07/2008	**BSc in Chemistry,** Reutlingen University, Germany
	Thesis: «Purification and Crystallization of *Schistosoma japonicum* Aspartic Protease Cathepsin D for Structural Study»
07/2004 – 06/2005	Lüneburg University, Germany
	German as a foreign language
07/2003 – 06/2004	Universidad Iberoamericana, Mexico City
	2 Semesters Chemical Engineering
07/2000 – 06/2003	High School, ITESM-CEM, Mexico City

List of publications

1. Oberoi, T. K., Dogan, T., Hocking, J. C., Scholz, R. P., Mooz, J., Anderson, C. L., Karreman, C., Meyer zu Heringdorf, D., Schmidt, G., Ruonala, M., Namikawa, K., Harms, G. S., **Carpy, A.,** Macek, B., Koster, R. W., and Rajalingam, K. (2012) IAPs regulate the plasticity of cell migration by directly targeting Rac1 for degradation. *The EMBO journal* 31, 14-28

2. Franz-Wachtel, M., Eisler, S. A., Krug, K., Wahl, S., **Carpy, A.,** Nordheim, A., Pfizenmaier, K., Hausser, A., and Macek, B. (2012) Global detection of protein kinase D-dependent phosphorylation events in nocodazole-treated human cells. *Molecular & cellular proteomics : MCP* 11, 160-170

3. Krug, K., **Carpy, A.**, Behrends, G., Matic, K., Soares, N. C., and Macek, B. (2013) Deep coverage of the Escherichia coli proteome enables the assessment of false discovery rates in simple proteogenomic experiments. *Molecular & cellular proteomics : MCP* 12, 3420-3430

4. Takeda, A.N., Oberoi, T.K., Glatz, G., Peter-Scholz, R., **Carpy, A.**, Macek, B., Remenyi, A., Rajalingam, K. (2014) Ubiquitin dependent regulation of MEKK2/3-MEK5-ERK5 signaling module by IAPs. *The EMBO journal*

5. **Carpy, A.**, Krug, K., Graf, S., Koch, A., Popic, S., Hauf, S., and Macek, B. (2014) Absolute proteome and phosphoproteome dynamics during the cell cycle of fission yeast. *Molecular & cellular proteomics : MCP* M113. 035824

Accepted Manuscripts

1. Macek, B.**, Carpy, A.**, Koch, A., Bicho, C.C., Borek, W.E., Hauf, S., Sawin, K.E. (2014) SILAC-based Quantitative Proteomics and Phosphoproteomics in Fission Yeast. *Cold Spring Harbor Protocols*, accepted

2. Koch, A., Bicho, C.C., Borek, W.E., **Carpy, A.**, Macek, B.**,** Hauf, S., Sawin, K.E. (2014) Construction, Growth and Harvesting of Fission Yeast SILAC strains. *Cold Spring Harbor Protocols*, accepted

3. **Carpy, A.**, Koch, A., Bicho, C.C., Borek, W.E., Hauf, S., Sawin, K.E., Macek, B. (2014) SILAC-based Quantitative Proteomics and Phosphoproteomics in Fission Yeast. *Cold Spring Harbor Protocols*, accepted

Submitted Manuscripts

1. **Carpy, A.**, Patel, A., Dee T.Y., Hagan, I., Macek, B. (2014) Nic1 inactivation enables SILAC labeling with $^{13}C_6^{15}N_4$-Arginine in S. pombe. Under review in *Molecular & cellular proteomics : MCP*

2. Krug, K., Popic, S**., Carpy, A.**, Taumer, C., Macek, B., (2014) Construction and assessment of individualized proteogenomic databases for large-scale analysis of non-synonymous single nucleotide variants. Under review in *Proteomics*

Oral Presentations

1. **Carpy, A.,** Krug, K., Graf, S., Koch, A., Hauf, S., Macek, B. (2012) Cell cycle: global quantitative acetylome, phosphoproteome and absolute proteome analysis in fission yeast. Talk: *PTMs in Cell Signaling.* Copenhagen, Denmark.

Poster Presentations

1. Franz-Wachtel, M., Eisler, S., Krug, K., **Carpy, A.**, Hausser, A., Macek, B. (2011) Global analysis of Protein Kinase D targets in human cells. Poster presented at: *10th Human Proteome Organization Annual World Congress.* Geneva, Switzerland.

2. **Carpy, A.**, Krug, K., Graf, S., Koch, A., Hauf, S., Macek, B. (2012) Global proteome, phosphoproteome and acetylome dynamics in the cell cycle of fission yeast. *Proceedings of the 60th ASMS Conference on Mass Spectrometry and Allied Topics*, Vancouver, Canada, May 20 - 24, WP03-074.

3. **Carpy, A.**, Patel, A., Dee, T., Hagan, I., Macek, B. (2013) Genetic modification of the urease activity enables full compatibility of fission yeast with SILAC labeling. *Proceedings of the 61th ASMS Conference on Mass Spectrometry and Allied Topics.* Minneapolis, MN, June 9 -13, ThP09.