Analysis and Interpretation of Protein Post-Translational Modification Site Stoichiometry

Prus, Gabriela; Hoegl, Annabelle; Weinert, Brian T; Choudhary, Chunaram

Published in:
Trends in Biochemical Sciences

DOI:
10.1016/j.tibs.2019.06.003

Publication date:
2019

Document version
Peer reviewed version

Document license:
CC BY-NC-ND

Citation for published version (APA):
Analysis and interpretation of protein posttranslational modification site stoichiometry

Gabriela Prus¹, Annabelle Hoegl¹, Brian T. Weinert¹⁺, Chunaram Choudhary¹⁺

¹Proteomics Program, the Novo Nordisk Foundation Center for Protein Research, Faculty of Health and Medical Sciences, University of Copenhagen, Blegdamsvej 3B, DK-2200 Copenhagen, Denmark

⁺Correspondence: brian.weinert@gmail.com (B. T. Weinert), chuna.choudhary@cpr.ku.dk (C. Choudhary).

Abstract:
Proteins are decorated with a diverse array of posttranslational modifications (PTMs) which regulate their spatial and temporal functions. Recent mass spectrometry (MS)-based studies have identified hundreds of thousands of PTM sites in mammalian proteomes. However, the signaling cues and enzymes regulating individual sites are often not known and their functional roles remain uncharacterized. Quantification of PTM site stoichiometry can help in prioritizing sites for functional analyses and is important for constructing mechanistic models of PTM-dependent protein regulation. Here, we review the concept of PTM site stoichiometry, critically evaluate the merits and drawbacks of different MS-based methods used for quantifying PTM site stoichiometry, and discuss the usefulness and limitations of stoichiometry in informing on the biological function of modified sites.

Key words:
posttranslational modification stoichiometry, mass spectrometry, proteomics, cell signaling
Mass spectrometry-based analysis of PTMs

Living systems are inundated with diverse intracellular and environmental signals; therefore, cells must be able to rapidly respond to programmed and unprogrammed cues. Posttranslational modifications (PTMs) provide an energy-efficient mechanism to reversibly regulate protein function on rapid timescales [1]. Indeed, PTMs are instrumental in relaying cellular signals and thereby controlling dynamic cellular processes such as cell proliferation, migration, and death [2]. The human proteome contains hundreds of different types of PTMs [3], though only few of them, such as phosphorylation, ubiquitylation, and acetylation, have been studied extensively.

Recent advances in quantitative mass spectrometry (MS)-based proteomics have revolutionized the systems-wide analysis of PTMs [4-7]. With the latest MS technologies, tens of thousands of PTM sites are routinely identified in a single study and cumulatively ~480,000 unique PTM sites have been catalogued in mammalian proteomes [8] (www.phosphosite.org). Given their widespread occurrence and rapid dynamics, PTMs have emerged as an important regulatory layer which greatly expands the functional diversity of proteomes.

Most large-scale PTM-mapping studies use a bottom-up proteomics (see Glossary) approach, in which proteins are first proteolyzed into small peptides using a protease, most commonly trypsin [7]. Modified peptides are then ‘fished-out’ with various PTM affinity-enrichment strategies and sequenced using liquid-chromatography-coupled high-resolution tandem MS (LC-MS/MS) [4]. By incorporating a variety of quantification methods, the relative and absolute abundance of PTMs can be determined in different samples or biological states. For further reading about different MS-based quantification approaches, we refer the reader to some excellent reviews [9-11].

Stoichiometry is an important property of PTMs (Box 1). In this article, we review different MS-based approaches for determining PTM site stoichiometry, discuss their strengths and limitations, and point out relevant practical considerations for selecting appropriate methods for stoichiometry quantification and for data interpretation (Box 2). Using phosphorylation and acetylation as examples,
we illustrate the usefulness of site stoichiometry in revealing global properties of PTMs. We conclude by discussing the importance of PTM site stoichiometry in developing a mechanistic understanding of PTM-dependent protein regulation, as well as its limitations in indicating the functional relevance of PTMs.

**Methods for quantifying PTM site stoichiometry**

PTMs differ in their chemical structures and target different amino acids in proteins [1]. Accordingly, tailored MS-based methods have been developed to quantify PTM site stoichiometry. These methods require different sample preparation steps, use different quantification strategies, and differ in their accuracy, throughput, and applicability to different PTMs.

**Quantification of PTM site stoichiometry using isotope-labeled standards**

The absolute amount of a peptide can be accurately measured by spiking in a known quantity of an **stable isotope-labeled peptide** standard and comparing the intensities of the native and reference peptides in MS; this method is termed **absolute quantification (AQUA)** [12] (Figure 1A). AQUA-based stoichiometry measurement requires the synthesis of an isotope-labeled peptide bearing the PTM (hereafter termed modified peptide), as well as a corresponding unmodified peptide (hereafter abbreviated as CP) for a given protein. After proteolysis of the native proteins, a known quantity of AQUA peptides is spiked into the sample of interest and the relative intensities of the native and AQUA peptides are measured. Because the native and AQUA peptides are chemically identical, their intensities are directly comparable and proportional to their abundance.

Similar to AQUA, the abundance of a protein can also be quantified by spiking in an isotope-labeled recombinant protein standard which can be generated using in-vivo metabolic labeling [13] or in-vitro protein synthesis [14]. The isotope-labeled proteins can be modified, chemically or enzymatically, to generate a PTM-containing protein reference. Alternatively, site-specific PTM-containing reference proteins can be generated by combining metabolic isotope-labeling with **genetic code expansion** technologies [15]. This technology has been successfully used to produce proteins bearing site-
specific acetylation, methylation, and phosphorylation [16]. In addition to structure-function analyses, site-specifically modified recombinant proteins can be used to quantify PTM site stoichiometry of native proteins [17, 18].

Several technical considerations are necessary when using AQUA. Before selecting AQUA peptides, the sample of interest should be analyzed to account for all observable peptide species, such as oxidized, non-oxidized, fully-cleaved, and mis-cleaved forms of the peptide. Incomplete proteolytic digestion, or the presence of modified forms of the peptide, can result in underestimation of peptide abundance. In particular, the presence of oxidizable methionine residues can result in different degrees of oxidization of native and AQUA peptides. Therefore, it is advisable to avoid selecting methionine-containing peptides, or, if that is not possible, to chemically oxidize all methionines to reduce the number of peptide species [19]. The use of two or more AQUA peptides provides greater confidence in protein quantification, although for specific PTM sites, the peptide choice may be restricted.

Standard AQUA-based PTM stoichiometry analysis requires quantification of both the modified peptide and the corresponding protein. However, if the native modified peptide is low-abundant, it may be undetectable using data-dependent acquisition (DDA) MS analyses. This problem can, to some extent, be overcome using targeted MS approaches, such as single reaction monitoring (SRM) or multiple reaction monitoring (MRM), which enable the quantification of selected ions during MS measurement and have been used for PTM site stoichiometry determination [17, 20, 21]. Also, a protein of interest can be affinity-enriched prior to AQUA analysis to reduce sample complexity and thus facilitate the detection of low-abundant PTM sites [12]. Similarly, detectability of low-abundant PTMs can be improved by affinity-enrichment of modified peptides [18, 22]. However, in this case, the abundance of the corresponding protein must be quantified separately in the sample without PTM enrichment. Furthermore, the PTM-bearing AQUA peptide must be spiked in prior to the enrichment step in order to account for incomplete enrichment efficiency.
A drawback of using isotope-labeled peptides is the high cost of synthesizing reference standards, which restricts their application on a proteome-wide scale. Moreover, the use of PTM-containing recombinant proteins is further limited because expression of full-length proteins is not always feasible, or technologies for the site-specific incorporation of certain PTMs are not available. Nonetheless, isotope-labeled reference-based methods are arguably the most accurate, and represent the method of choice for quantifying stoichiometry of a limited number of sites with very high precision. Indeed, the application of AQUA-based stoichiometry analysis has been demonstrated for a range of PTMs, including phosphorylation \cite{12}, lysine methylation \cite{23}, acetylation \cite{24}, and ubiquitylation \cite{25}.

**PTM site stoichiometry by label-free quantification**

Theoretically, the stoichiometry of a PTM site can be determined by comparing the intensity of modified and unmodified peptides in MS. However, despite a general correlation between peptide abundance and MS intensity \cite{26}, the intensity of individual peptides is variable and can be influenced by unpredictable factors, such as peptide ionization efficiency. For this reason, the absolute abundance of individual peptides cannot be directly compared based on their MS intensities.

Steen et al. devised a **label-free quantification** approach for determining differences in the ‘flyability’ of a phosphorylated peptide and its CP, and used this information to quantify phosphorylation site stoichiometry \cite{27}. The rationale behind this approach is that an increase or decrease in phosphorylation should result in a proportional and reciprocal change in CP abundance. The MS intensities of the phosphorylated peptide and CP are measured in two different samples that contain an equal amount of total protein, but differ in phosphorylation stoichiometry (**Figure 1B**). Since the protein abundance is kept equal between the two samples, the difference in phosphorylated peptide abundance should equal to the difference in CP abundance. Thus, the relative ‘flyability’ of the phosphorylated peptide and CP can be determined by measuring the difference in peptide intensity from the two samples (**Figure 1B**). Peptide intensities normalized with the ‘flyability ratio’ are directly...
proportional to their abundance; therefore, stoichiometry can be calculated using the normalized intensities of phosphorylated peptide and CP (Figure 1B).

A critical requirement of this method is that the ‘flyability’ ratio is determined accurately [27]. To reliably quantify phosphorylated and unmodified peptide intensities, a protein must be phosphorylated to a substantial degree (>10%), but should not be fully phosphorylated. Also, the amount of total protein in the two samples must be kept equal. Due to its laborious nature, this approach is not suited for systems-wide analysis of PTM stoichiometry.

**Stoichiometry determination by quantifying the relative abundance of modified peptide and CP**

To estimate PTM site stoichiometry on a global scale, Olsen et al. employed a stable isotope-labeling by amino acids in cell culture (SILAC)-based approach [28]. Similar to the logic behind the above discussed label-free method, this approach assumes that if the stoichiometry of a PTM site changes, the abundance of the modified and CP should change in a reciprocal manner (Figure 2A). Thus, PTM site stoichiometry can be determined by quantifying the relative abundance of modified peptide and CP, provided that PTM abundance differs substantially in the two samples and changes in protein abundance are also accounted for. Because the relative change in modified peptide and CP abundance can be directly quantified with isotope-labeling, this alleviates the need to determine peptide ‘flyability’.

An advantage of this approach is its easy implementation in any data processing pipeline that includes relative abundance information for the modified peptide, the CP, and the protein. Indeed, this feature is already included in the computational MS data analysis software MaxQuant [29], making it convenient for calculating the stoichiometry of PTMs that occur at sufficiently high stoichiometry and show robust changes upon perturbation, such as phosphorylation [30]. However, as discussed below, caution should be exercised in using this approach for PTMs that mostly occur at a low stoichiometry, such as acetylation and other types of lysine acylations.
A critical requirement of the above method is that the modified peptide and CP must show a measurable change upon a given perturbation; stoichiometry cannot be quantified for PTM sites that remain unchanged. For protein phosphorylation, peptides can be treated with a non-specific protein phosphatase to remove all phosphate groups and thereby induce a measurable change in the abundance of modified peptide and CP over control (Figure 2B). Control and phosphatase-treated samples are then compared using MS-based relative quantification approaches, such as dimethyl-labeling or tandem mass tags (TMT). This approach was initially applied to quantify the stoichiometry of a limited number of sites on purified proteins [31-33]. Later, Wu et al. applied this method to determine phosphorylation stoichiometry of thousands of sites in Saccharomyces cerevisiae [34]. However, due to the stochasticity in peptide sequencing in DDA-based approaches, detection of dephosphorylated peptides in complex sample mixtures can be challenging and this can limit the depth of analysis in organisms with high complexity.

To overcome this limitation, Tsai et al. devised a method to deepen stoichiometry analysis of a subset of phosphorylation sites that are targeted by specific protein kinases [35]. Similar to the approach of Wu et al., a sample of interest is split into two portions, which are either mock- or phosphatase-treated (Figure 2C). Following proteolysis, the samples are differentially labeled with stable isotopes, mixed, and phosphorylated peptides are enriched using immobilized metal ion affinity chromatography (IMAC) (hereafter 1st IMAC step). Analysis of the eluates from 1st IMAC provides a list of phosphorylated peptides that are present in the native sample. The identified phosphorylation site sequences can be analyzed to determine consensus motifs and to predict their upstream regulatory kinases. In the second part of the procedure, the flow-through from the 1st IMAC, which contains unphosphorylated CP, is treated with a selected kinase that phosphorylates a subset of CPs containing the target motif of interest. After the kinase reaction, phosphorylated peptides are again enriched using IMAC (hereafter 2nd IMAC) and quantified by MS (Figure 2C). Phosphorylation site stoichiometry is deduced from the ratio of phosphorylated peptides in the two samples [35]. An advantage of this approach is its high sensitivity for quantifying the stoichiometry of phosphorylation sites targeted by a specific kinase. A drawback is that it requires additional steps for the kinase reaction and IMAC.
enrichment, and stoichiometry can only be calculated for sites phosphorylated by a specific kinase. The methods used by Wu et al. and Tsai et al. further required the use of a highly efficient and promiscuous enzyme (phosphatase) that removes nearly 100% of the modifications on all sites. This is feasible for some PTMs, such as phosphorylation, but may not be extendible to other PTMs due to unavailability of suitable PTM-removing enzymes. In some instances, phosphorylation near Arg and Lys can reduce the kinetics of trypsin proteolysis [36, 37]; this may lead to different phosphorylated and CP sequences and could affect stoichiometry estimation.

Most critically, the above discussed approaches require that changes in CP abundance are measurable. This is a general concern and it is particularly important for sites with low stoichiometry. In shotgun proteomics, the accuracy of peptide measurement is related to its abundance: more abundant peptides have a better signal-to-noise ratio and are thus quantified with a greater accuracy [38]. Because most PTMs occur at a low stoichiometry, modified peptides often have low intensity and are therefore impacted more strongly by poor quantification. In comparison, protein abundance is mostly calculated from the median abundance of multiple independent peptides, minimizing the impact of erroneous peptide quantification on protein quantification. Because the quantification of modified peptides and CP mostly relies on a single peptide, the quantitative accuracy is often lower than for proteins.

It is therefore critically important to ensure that observed changes in peptide abundance are statistically significant when inferring PTM site stoichiometry. The changes in modified peptide and CP abundance must be greater than the inherent quantitative variability of the method used, otherwise the calculation will be based on the normal variability of the measurements. Using phosphatase treatment as an example, if phosphorylation occurs at a stoichiometry of 0.1%, 1%, or 5%, the respective CP ratio (phosphatase treated/control) will increase to 1.001, 1.01, and 1.05. However, most MS methods used for global PTM analyses are not accurate enough to confidently measure such small ratio changes, particularly if PTM site quantification is based on a single quantitative
measurement. As an example, upon genetic and metabolic perturbations, we noted vast increases (greater than an order of magnitude) in relative acetylation abundance in bacteria, yet a corresponding change in CP abundance was unmeasurable by MS [39]. Similarly, Tatham et al. showed that treatment of Hela cells with aspirin increased acetylation of many sites by several orders of magnitude; however, a change in CP abundance was unmeasurable by SILAC MS and therefore it was impossible to use the CP-based approach for calculating acetylation site stoichiometry [40]. Furthermore, if quantification error is greater than PTM site stoichiometry, it can yield negative stoichiometry values or stoichiometry greater than 100%, both of which are theoretically impossible [41]. In conclusion, while CP-based methods are useful for quantifying stoichiometry of abundant sites and for understanding global patterns of PTM stoichiometry, it is important to understand their limitations in terms of quantification accuracy and suitability for quantifying stoichiometry of low abundant modifications.

**PTM site stoichiometry quantification with isotope-labeled CP reference**

Variations of the above discussed methods have been used to quantify phosphorylation stoichiometry. In an approach termed phosphatase-based phosphopeptide quantitation (PPQ), Domanski et al. combined AQUA with phosphatase treatment [42]. Stoichiometry is determined by comparing the intensity of CP with and without phosphatase treatment, rather than comparing the intensity of phosphorylated peptide and unmodified peptide as in classical AQUA. An advantage of this method is that it requires the synthesis and quantification of only a single unmodified peptide; it nonetheless requires measurement of the sample with and without phosphatase treatment.

Instead of chemically synthesizing isotope-labeled peptides, Pratt et al. generated isotope-labeled peptide standards in *E. coli* by expressing an artificial protein containing concatenated peptide sequences, called QconCAT [43, 44]. An unmodified peptide corresponding to the phosphorylated peptide (CP) was concatenated to reference peptides [44], which are derived from the protein regions that lack phosphorylation, and hence can be used to quantify total protein amount. The QconCAT
reference is spiked in with a protein of interest, digested with a protease, and the peptides are quantified by MS. Stoichiometry is inferred by comparing the abundance of total protein to the abundance of CP, and a reduction in CP abundance was attributed to the stoichiometry of phosphorylation. This is based on the assumption that in a completely unmodified state, all peptides derived from a protein should be equally abundant. Therefore, an important consideration in using this approach is to identify reference peptides that reflect the true abundance of a protein.

Instead of producing concatenated peptides, Singh et al. developed the FLEXIQuant (full-length expressed stable isotope-labeled proteins for quantification) approach to generate isotope-labeled full-length protein references for absolute PTM quantification [14]. Using a workflow similar to the above discussed QconCAT method, PTM stoichiometry is quantified by ‘spiking-in’ a FLEXIQuant reference protein. An advantage of this approach is that by measuring the abundance of all quantifiable peptides, it is possible to quantify PTM site abundance across an entire protein. Indeed, by using this approach, stoichiometry was quantified for >20 phosphorylation sites on Tau protein expressed in cultured cells as well as in post-mortem brain samples from individuals with Alzheimer's disease [45].

**PTM site stoichiometry quantification by comprehensive chemical modification**

Other approaches have analyzed PTM site stoichiometry by comprehensive chemical modification of unmodified sites with stable isotope-labeled reagents, and comparing the abundance of native and chemically modified sites. Smith et al. initially used deuterated acetic anhydride to comprehensively modify all non-acetylated lysine residues on purified histone H4 to quantify the stoichiometry of acetyl-histone sites [46] (Figure 3A). Recently, several studies extended this approach to estimate acetylation site stoichiometry on a larger scale [47-50]. For example, Baeza et al. used deuterium-labeled acetic anhydride to chemically acetylate whole cell proteins under denaturing conditions [47]. Acetylation was quantified by comparing parent ion (MS) intensities from native heavy-isotope-labeled, chemically acetylated samples.
Meyer et al. used comprehensive chemical acetylation and SWATH-based **data-independent acquisition** (DIA) to quantify acetylation site stoichiometry [49] (**Figure 3A**). Acetylation site stoichiometry was calculated by comparing acetyl-lysine-specific fragment ion intensities from native acetylated and comprehensively acetylated samples. Quantification of multiple acetyl-lysine-derived fragment ions in MS2 was reported to provide improved quantitative accuracy as compared to MS1-based quantification [49]. While these studies validated their approach using acetylated BSA, the accuracy of their measurements of endogenous acetylation site stoichiometry was not validated using independent methods such as AQUA.

Nakayasu et al. used heavy carbon (C13)-labeled acetic anhydride to chemically acetylate peptides after endoproteinase Arg-C digestion [48] (**Figure 3B**). Acetylation site stoichiometry was calculated by comparing the intensity of the diagnostic immonium ion, which is generated upon fragmentation of acetyl-lysine [51]. To quantify the relative abundance of the diagnostic acetyl-lysine ion in MS2, the method requires a wider mass isolation window to simultaneously isolate native and chemically acetylated peptides for fragmentation. This can be problematic because it increases the chances of co-isolating unrelated neighboring acetylated peptides, which can adversely impact the quantitative accuracy, as shown for MS2-based quantification using isobaric mass tags [52-55].

Additional general restrictions limit the application of comprehensive chemical acetylation-based approaches. Firstly, comprehensive chemical acetylation creates many acetylated peptides that are not normally present in unperturbed samples, and causes a massive increase in isotope-labeled acetyl-peptide signal, which hampers detection of native acetylated peptide due to increased sample complexity and the very high dynamic range. Also, native peptides that do get quantified show a high quantitative error rate due to the high dynamic range [40, 56]. Secondly, because comprehensive acetylation prevents trypsin cleavage at lysine residues, stoichiometry analyses are limited to arginine-flanked peptides unless a different protease, such as Glu-C or Asp-N, is used. Thirdly, the use of trypsin or Arg-C proteases generates longer peptides [48]. As a result, some peptides will contain
more than one acetylation site and the stoichiometry then represents the average of these sites. Finally, stoichiometry measurement of low-abundant sites is impacted by the purity of isotope-labeled acetylating agent, as the presence of unlabeled (1-2%) forms of acetylation reagents limits quantification of sites with low (<1%) stoichiometry. This is a particular concern, since we, and others, found that most acetylation sites occurred at a stoichiometry of lower than 1% in *E. coli* [56], yeast [57], mouse liver tissue [58], and human cells [49, 59].

**PTM site stoichiometry quantification by partial chemical modification**

While studying acetylation dynamics in bacteria, we observed large relative increases in acetylation without a clear measurable decrease in CPs [47], suggesting a very low stoichiometry of acetylation. Because the above described CP-based approaches are not suitable for measuring the stoichiometry of low-abundant sites, we developed an alternative ‘AQUA-like’ method to estimate acetylation site stoichiometry (Figure 4A). SILAC-labeled protein lysates were partially chemically acetylated using acetyl-phosphate [39] and the degree of acetylation was estimated using reference AQUA peptides [57], comparison to 100% acetylated protein [56], or SILAC and TMT-based quantification of the median decrease in CP abundance upon partial chemical acetylation [58, 59]. Native acetylated SILAC lysates were mixed with the chemically acetylated SILAC proteins and digested with trypsin, generating ‘AQUA-like’ peptide standards from the partially chemically acetylated proteins (Figure 4A). The relative abundance of native acetylated and chemically acetylated peptides was quantified by MS, and stoichiometry was calculated assuming a constant degree of chemical acetylation at all sites.

Isotope dilution mass spectrometry is widely used to accurately quantify (bio)chemical analytes [60-62]. To further improve the accuracy of PTM stoichiometry estimation by chemical acetylation, we used this concept and developed the serial dilution SILAC (SD-SILAC) approach [56]. SILAC labeled proteins were partially acetylated and ‘spiked’ into lysates at different concentrations (Figure 4B). SD-SILAC improves stoichiometry measurements by ensuring that peptide quantification occurs at
multiple independent data points using different concentrations of chemically acetylated peptide. In addition, while comprehensive acetylation prevents trypsin proteolysis at lysine residues, the use of partial chemical acetylation allows trypsin cleavage at unmodified lysine sites and thereby enables direct comparison to native acetylated peptides.

One disadvantage of this method is that the low degree of partial chemical acetylation results in small differences when quantifying high stoichiometry sites (5% partial chemical acetylation will result in just a 1.05x increase for a site with 50% native stoichiometry), thus limiting the resolution of stoichiometry measurements for sites with high (>10%) native stoichiometry [59]. In addition, variation in the degree of partial chemical acetylation at individual lysine residues may result in inaccurate measurements, though we found good agreement between the stoichiometry quantified using partial chemical acetylation and AQUA peptides [58, 59]. In both the methods (comprehensive and partial chemical modification), it is important to accurately determine the degree of chemical modification. This is especially important for partial chemical modification because it can be difficult to control the exact extent of partial chemical modification and the degree of chemical modification can vary somewhat from experiment-to-experiment.

Generally, chemical acetylation-based approaches include several advantages, such as the relatively low cost, applicability on a proteome-wide scale, and the requirements to measure a single ratio (native acetylated/chemically acetylated peptide). However, the application of in-vitro chemical modification-based approaches is limited to PTMs that can be introduced chemically, which is not feasible for many biologically-relevant PTMs. For example, it is relatively easy to introduce acetylation and other forms of lysine acylation, while introducing specific forms of lysine methylation (i.e. mono, di, or tri-methylation) has remained challenging. Further, if a modification targets frequently occurring amino acids, chemical modification will generate a very large number of multiply modified peptides which limits their usefulness in estimating site-specific stoichiometry of singly modified peptides. In this respect, most of the above discussed global approaches are limited to measuring stoichiometry of peptides with a single PTM site, although this problem can be overcome by
developing computational algorithms that allow stoichiometry estimation from multiply modified peptides [50, 63].

**Relative PTM site stoichiometry by abundance-corrected modified peptide intensity (ACI)**

Because peptide intensity in MS measurements is largely proportional to peptide abundance, the intensity of the modified peptide can serve as a proxy for relative PTM site abundance. However, since the intensity of modified peptides also depends on the cellular abundance of the corresponding protein, it is necessary to correct for this bias. This can be achieved by dividing the intensity of the modified peptide by the abundance of the corresponding protein, which can be estimated using intensity-based quantification (iBAQ) [64]. This protein-normalized intensity of a modified peptide is referred to as the abundance-corrected modified peptide intensity (ACI) [57]. We found that ACI correlated with acetylation stoichiometry estimates obtained from partial chemical acetylation approach [57, 59]. However, due to the inherent variability in peptide intensity, the ACI-based approach suffers from poor accuracy and has limited predictive power. Also, this approach only provides information on the relative abundance of modified peptides in a sample, and not on the absolute PTM site stoichiometry. In spite of these limitations, the simplicity of the method means that relative PTM abundance can be estimated by measuring the proteome and PTM sites from any sample, without the need for SILAC or other relative quantification approaches.

**Biological insights from global PTM site stoichiometry measurements**

Historically, PTM site stoichiometry analysis was limited to a few sites and was used in conjunction with other functional assays to explain the mechanistic impact of a modification on protein function. Recent MS-based studies have extended this type of analysis to a proteome-wide scale and have analyzed stoichiometry of thousands of phosphorylation sites in evolutionarily diverse organisms, including budding yeast, fission yeast, and human cells [28, 30, 34, 35]. These studies revealed that although the stoichiometry of phosphorylation varies greatly among different sites, many phosphorylation sites occur at a relatively high stoichiometry (>10%). Interestingly, in comparison to
eukaryotes, the stoichiometry of phosphorylation in exponentially growing \textit{E. coli} is low (median 4%), but stoichiometry of most phosphorylation sites is increased (median 11%) during the stationary phase [65].

Analysis of >5,000 phosphorylation sites in human Hela cells during cell division showed that the stoichiometry of many phosphorylation sites greatly increases during mitosis, with more than half of the sites reaching a stoichiometry of >70% [28]. Consequently, many high-stoichiometry sites in the mitotic phosphoproteome contained a consensus motif for cell cycle-regulating cyclin-dependent kinases (CDKs) [28]. Similar to human cells, stoichiometry of phosphorylation doubled in mitosis of fission yeast, reaching median stoichiometry of 44% [30]. These studies provided detailed insights into phosphorylation site stoichiometry on a global scale and tied physiologically-relevant changes in relative phosphorylation abundance to absolute changes in phosphorylation stoichiometry.

Quantification of stoichiometry for >5,000 phosphorylation sites in asynchronously growing yeast cells revealed that ~10% of sites were nearly completely modified (>90%) [34]. In general, nuclear proteins had a higher phosphorylation stoichiometry than proteins localized to the cytoplasm, ribosome, or mitochondria. Furthermore, the stoichiometry of phosphorylation sites containing acidic motifs, which are targeted by casein kinase 2 (CK2), was substantially higher than proline-directed sites that are phosphorylated by mitogen-activated protein kinases (MAPKs) and CDKs [34]. Also, in human cells the stoichiometry of acidic motif-containing phosphorylation sites is higher than the sites that are phosphorylated by MAPK and EGFR [35]. Thus, the high stoichiometry of acidic motif-containing phosphorylation sites appears to be an evolutionarily conserved feature from yeast to human [34, 35]. Furthermore, comparative analysis of absolute phosphorylation site stoichiometry in drug-sensitive and drug-resistance cancer cells revealed quantitative changes in phosphorylation stoichiometry and identified several sites with elevated stoichiometry in drug-resistant cells [35].

Proteome-wide studies have measured lysine acetylation stoichiometry in bacteria [47, 49, 56], yeast [57], mouse [48, 58], and human cells [50, 59, 66], and we estimate the median stoichiometry of
acetylation in bacteria, yeast, mouse, and human cells to be less than one tenth of a percent (ie. < 0.1 %) [39, 57-59]. Meyer et al. concluded that stoichiometry of lysine succinylation in E. coli is similarly low as acetylation [49]; although their acetylation site stoichiometry estimates are considerably higher than observed in our work [56]. These differences likely arise from differences in methods used stoichiometry quantification. Notwithstanding the variability of stoichiometry estimates among different studies, it is clear that, in contrast to phosphorylation, most acetylation sites occur at a very low stoichiometry. In yeast and human cells, acetylation sites with high stoichiometry are significantly biased to occur on nuclear proteins and proteins associated with chromosome organization, transcription, and DNA binding (Gene Ontology) [57, 59]. More than half of the high-stoichiometry sites in yeast cells occur on histones, transcription factors, histone deacetylase (HDAC) complex members, or histone acetyltransferase (HAT) complex members [57, 59].

Notably, lysine acetylation and several other types of lysine acylation can also occur through non-enzymatic mechanisms [39, 57, 58, 67, 68]. Interestingly, James et al. showed that proximity of a cysteine residue can promote the non-enzymatic acetylation of a nearby lysine [69]. This involves first the non-enzymatic S-acetylation of cysteine followed by transfer to the ε-amino group of a nearby lysine [69]. Examination of high-stoichiometry acetylation sites showed that, indeed, acetylation sites occurring proximal to cysteine residues had a higher stoichiometry [58, 59, 70], providing important evidence for the in-vivo relevance of the cysteine proximity-based mechanism of acetylation.

Together, these examples illustrate the usefulness of proteome-wide PTM stoichiometry analyses in revealing fundamental features of PTM-based regulatory mechanisms, and provide interesting insights into specific protein categories that harbor a higher stoichiometry of modification.

**Usefulness and limitations of PTM site stoichiometry in functional interpretation**

The knowledge of PTM site stoichiometry is useful for understanding the mechanistic implications of PTMs in protein regulation. For example, if a PTM site is directly associated with inactivation of enzyme activity, a reduction in the enzymatic activity will be expected to be commensurate with PTM
stoichiometry. While stoichiometry itself is not an absolute indicator of PTM function, high stoichiometry sites are unlikely to occur spuriously and are therefore more likely to have a function. Such PTMs have the potential to simultaneously affect a large pool of the modified protein and may be involved in regulating protein activity, impacting subcellular localization, or mediating protein-protein interactions.

While it is easier to rationalize the possible functional relevance of high stoichiometry sites, it should be clarified that low-stoichiometry PTM sites can, and do, have important functional roles. This is particularly true for sites that are modified in a conditional manner and induce a gain-of-function on the modified protein. Thus, it is expected that in the absence of appropriate signals, signal-responsive sites will have a low stoichiometry in steady-state cells [22]. Low stoichiometry could also reflect occurrence of the PTM at a specific place and time, such as at sites of transcription, replication and DNA damage. Because stoichiometry measurements are performed using a population of cells, the average PTM stoichiometry will be low if a protein is modified in only a small fraction of cells, even if such cells may specifically harbor high stoichiometry modification sites. Similarly, PTM site stoichiometry will be low if a modification destabilizes a protein, such as ubiquitylation, which targets many proteins for proteasomal degradation [71]. Another possible reason for low stoichiometry may be that PTM sites function in a collective, non-site-specific manner. According to the ‘SUMO Spray’ model [72], multi-site sumoylation of spatially restricted proteins, such as at sites of DNA damage, could collectively facilitate local protein-protein interactions. In such instances, the function of the individual sites adds up and the exact site-specificity of modification may not be of critical relevance. Thus, while low PTM site stoichiometry does not indicate a lack of function per se, knowing stoichiometry can help in understanding how it may impact protein function.

**Concluding remarks**

In the past decade, technological advances, including MS-based proteomics, have ushered in a new era in quantitative biology. Development of MS-based methods has helped in obtaining the first global snapshots of site-specific stoichiometry for some of the most extensively studied PTMs. We
anticipate that future studies will apply these technologies to investigate PTM site stoichiometry in varied contexts, such as in evolutionarily diverse organisms, cell types, and cellular states (see Outstanding Questions). For this, it will be interesting to incorporate approaches that enable increased quantitative multiplexing and cost-efficient analysis of stoichiometry. We also anticipate that future innovative technologies will enable stoichiometry analysis of PTM types that are beyond reach of the available technologies.

A quantitative understanding of cellular regulatory systems requires detailed knowledge of PTM site stoichiometry. It is our hope that an increasing number of biological and computational researchers will integrate PTM site stoichiometry analyses in their research to study the functional impact of PTMs and to construct quantitative models of signaling pathways. Just as knowing the concentration of relevant enzymes is useful in constructing metabolic flux models, accurate PTM stoichiometry data could help in developing and refining mathematical models of signaling pathways.

**Box 1: What is PTM site stoichiometry and why is it important to understand it?**

PTM site stoichiometry, also referred to as PTM site occupancy or fractional PTM occupancy, indicates the fraction of a protein that is modified with a given PTM at a specific site (Figure 1). Historically, the identification of PTM sites has relied on low-sensitivity methods, and thus, the identified sites either had a high stoichiometry or occurred on highly abundant proteins. Not surprisingly, these sites were regulated by enzymes and had a biological function. Extrapolating from these early findings, many biologists started to assume that all PTMs were catalyzed by enzymes, targeted to specific residues in proteins, occurred at a substantial stoichiometry, and had a biological function. Based on these assumptions, the function of PTM sites were often investigated using proteins that are site-specifically modified or engineered to contain a PTM-mimicking residue. Indeed, these approaches have proven instrumental in uncovering the functional roles of PTMs. However, a major limitation of these technologies is that proteins are often irreversibly modified to a complete stoichiometry which could overestimate the functional impact of low-abundant PTM sites.
With the increased sensitivity of modern MS instruments and improved PTM-enrichment approaches, researchers are now picking out thousands of very low stoichiometry sites, some of which contradict the widely-held assumptions stated above. The functional interpretation of PTM sites often relies on analyzing the relative change in PTM abundance, as quantified by immunostaining with PTM-site-specific antibodies or mass spectrometry. However, if the relative abundance of a PTM site is increased by 10-fold over the control, this could either mean an increase in PTM stoichiometry from 0.1 to 1% or from 10% to 100%. It is quite obvious that such large differences in stoichiometry will have different consequence on overall protein function. Thus, understanding PTM site stoichiometry helps in interpreting the functional relevance of a modification site and its mechanism for affecting protein function.

**Box 2: Considerations for the calculation and interpretation of PTM site stoichiometry**

There are several important considerations for selecting an appropriate method for PTM site stoichiometry quantification and for interpreting the resulting data. Below, we highlight some general points to consider, in addition to the method-specific factors discussed throughout the text.

**Target PTM abundance and quantitative precision**

While it is well-understood that the stoichiometry of a given PTM varies from site-to-site, it is perhaps less appreciated that the average stoichiometry can vary greatly for the different types of PTMs. For example, N-glycosylation often occurs at a high stoichiometry [73, 74], whereas phosphorylation occurs at a medium-to-high stoichiometry [28, 30, 34, 41, 75], and acetylation occurs at a relatively low stoichiometry [47, 49, 50, 56, 58, 59, 76]. Therefore, depending on the abundance of the target PTM, attention should be paid in choosing an appropriate MS method for stoichiometry calculation – there is no ‘one size fits all’ solution.

Also, the requirement for quantitative precision is not the same for all measurements. For example, we use different weighing scales to weigh different things in our everyday lives. Similarly, not all PTMs stoichiometry quantification experiments require the highest precision. It is therefore useful to
consider the desired quantitative precision and to choose appropriate methods that match this requirement.

**Validation of quantitative accuracy**

For PTM site stoichiometry calculations, it is useful to estimate quantitative variability of measurements. Also, when developing new methods or testing existing methods on new PTMs, the stoichiometry, at least for a subset of sites, should be independently validated using alternative, well-established methods such as AQUA. Such orthogonal validation can provide added confidence in measurements.

**Impact of unknown modification and stoichiometry of multiply-modified peptides**

In most PTM site stoichiometry measurements, it is assumed that the sum of modified and unmodified peptide equals 100% of protein abundance. However, proteins containing an unknown modification, within the same peptide sequence, may cause an under- or over-estimation in PTM stoichiometry. In such instances, quantification of multiple unmodified peptides can aid in the accurate measurement of total protein amount. However, for peptides that are modified at multiple sites and with different PTMs, such as from histone tails, accurate calculation of site-specific stoichiometry can still be challenging.

**Glossary**

**Bottom-up proteomics**: a MS-based approach in which proteins are proteolyzed into small peptides prior to their sequencing in a mass spectrometer. This approach is alternative to the top-down proteomics methodology in which intact proteins are analyzed by MS.

**Data-dependent acquisition (DDA)**: a data acquisition strategy in tandem mass spectrometry in which, based on predetermined criteria, a fixed number of precursor ions are selected from the survey scan (MS1) and subjected to MS/MS analysis.
**Data-independent acquisition (DIA):** involves fragmentation of all peptide ions in a defined m/z window and recording of composite fragment ion spectra. The complex peptide spectra are identified by matching them to a peptide spectral library that contains precursor masses and fragment spectra of all peptides present in the sample.

**Genetic code expansion:** refers to a synthetic biology approach in which an artificial codon is used to introduce an unnatural amino acid at a desired position in a protein. It requires engineering of a tRNA/synthase pair that recognizes an artificial codon and installs an unnatural amino acid (i.e. an amino acid different from one of the standard 20 amino acids) at this position.

**Intensity-based absolute quantification (iBAQ):** a label-free MS quantification strategy for estimating absolute protein amounts. The sum of all observed peptide MS intensities is divided by the number of theoretically observable peptides for a given protein. The normalized protein intensity is referred to as iBAQ.

**Label-free quantification:** refers to stable isotope-free quantification by MS. Peptides from two or more samples are analyzed separately in MS, and their relative intensities are compared computationally.

**PTM affinity-enrichment:** a sample preparation step in MS-based global PTM analysis. PTM-bearing peptides are enriched using an affinity matrix that selectively binds to the modified peptides, typically independent of the peptide sequence flanking the modification site. PTM affinity-enrichment reduces sample complexity, and thus allows sequencing of low-abundant modified peptides which otherwise may not be detected.

**Shotgun proteomics:** a bottom-up-based MS strategy for sequencing a peptide mixture without any prior knowledge about the mixture content. This typically involves on-line separation of the peptides with high-performance liquid chromatography followed by their identification by tandem MS (LC-MS/MS).
**Stable isotope-labeled peptides**: synthetic peptides that are generated using one or more stable isotope-labeled ‘heavy’ amino acids, which are produced by substituting natural variants of certain atoms with their ‘heavy isotope’ variant. The most frequent substitutions are $^{12}\text{C}$ by $^{13}\text{C}$ (carbon-13), $^{14}\text{N}$ by $^{15}\text{N}$ (nitrogen-15), and $^1\text{H}$ by $^2\text{H}$ (deuterium- D).

**Stable isotope labeling by amino acids in cell culture (SILAC)**: a MS-based method for the relative quantification of proteins and PTMs. It involves metabolic labeling of proteins, in cells or small organisms, with stable isotope-labeled amino acids.

**Tandem mass tag (TMT)**: a set of isobaric chemical labels that are used for MS-based quantification of peptides from different samples. Because of their isobaric nature, peptides with different TMT labels are indistinguishable in MS1. Upon peptide fragmentation, each TMT label produces a characteristic reporter ion in MS2, which reflects the relative abundance of peptide in the respective samples.

**Figure legends**

**Figure 1. Analysis of PTM site stoichiometry by AQUA and label-free quantification.** (A) AQUA-based quantification of PTM site stoichiometry. After proteolysis of native proteins (yellow), a known concentration of synthetic, stable isotope-labeled modified and unmodified peptides (red) is ‘spiked’ into a sample containing native peptides [12]. The MS intensity of the AQUA peptides is compared with the intensity of native modified and unmodified peptides; due to isotope-labeling, the mass to charge (m/z) ratio differs between AQUA and native peptides. By knowing the absolute concentration of AQUA peptides, the abundance of native modified peptides and unmodified peptide can be quantified and PTM stoichiometry can be calculated using the indicated formula. (B) Label-free strategy for quantifying PTM site stoichiometry. The MS intensity of modified and unmodified peptides from two samples, which differ in the PTM site stoichiometry, is quantified by MS; due to PTM, the m/z ratio differs between modified and unmodified peptides. The difference in modified
and unmodified peptide intensities in the two samples is used to calculate the ‘flyability’ ratio of modified and unmodified peptides [27]. PTM stoichiometry is calculated using the intensities of modified and unmodified peptides, normalized by multiplying with the ‘flyability’ ratio.

**Figure 2. PTM site stoichiometry determination by stable-isotope-based relative quantification.** (A) Stoichiometry analysis using relative quantification of modified and the corresponding unmodified peptide (CP). Proteins from two different SILAC-labeled samples are mixed together, proteolyzed, and modified peptides are enriched using PTM-affinity enrichment. The enriched material is analyzed by MS to quantify the relative abundance of the modification from the two samples. In parallel, a fraction of the mixed SILAC peptides is analyzed without PTM affinity-enrichment to quantify the relative abundance of CP and protein. PTM site stoichiometry is calculated using SILAC ratios of modified peptide, CP, and protein, using the indicated formulas [28]. (B) Stoichiometry determination by protein dephosphorylation and MS. Following proteolysis, protein is split into two equal portions which are either mock- or phosphatase-treated. The two samples are differentially isotope-labeled, mixed, and quantified using MS. Stoichiometry of phosphorylation sites is calculated using the relative abundance change in unmodified peptide corresponding to phosphorylation site (CP), using the shown formula [34]. (C) Motif targeted proteomic approach for quantifying phosphorylation stoichiometry [35]. Following proteolysis, peptides are split into two equal portions which are either mock- or phosphatase-treated. The two samples are differentially isotope-labeled and mixed, and then PTM-containing peptides are affinity-enriched (1st IMAC). The enriched fraction is analyzed by MS to identify native phosphorylation sites in the sample and modifying kinase(s) predicted by consensus motif characterization. The flow-through fraction is subjected to in-vitro phosphorylation using a predicted kinase, the modified peptides are affinity enriched (2nd IMAC), and analyzed by MS. PTM site stoichiometry is calculated based on relative abundance of modified peptides [35].

**Figure 3. Comprehensive chemical modification-based strategies for quantifying acetylation site stoichiometry.** (A) Stoichiometry estimation using comprehensive chemical acetylation. Endogenous
proteins (yellow), containing native acetylation and unmodified forms, are comprehensively acetylated with stable isotopes (red) before proteolysis. The relative abundance of native acetylated and chemically acetylated peptide is then quantified by MS using data-dependent acquisition (DDA) or data-independent acquisition (DIA) approaches. In DDA-based methods, the relative abundance of native and chemically acetylated peptide is quantified using intensity of intact peptide ions in MS1 (the first stage of mass spectrometry). In DIA-based methods, the relative abundance of native and chemically acetylated peptides is quantified using extracted ion chromatograms (XICs) of fragment ions in MS2 (a stage of mass spectrometry) [49]. In this example, only acetyl-lysine-containing fragment ions from native (yellow) and chemical (red) are used for PTM site stoichiometry quantification, and non-acetyl-lysine containing fragment ions (grey) are not used for acetylated peptide abundance quantification. Based on the relative intensities of native and chemically acetylated peptides, stoichiometry is calculated using the shown formula. (B) Acetylation site stoichiometry determination by acetyl-lysine immonium ion. Cellular lysates, containing unmodified and native acetylated proteins (yellow), are treated with $^{13}$C-labeled acetic anhydride (red), which comprehensively acetylates all unmodified lysine residues in proteins. Following proteolysis, abundance of the native and chemically acetylated peptides is quantified by comparing intensities of the diagnostic acetyl-lysine immonium ion [48]. The diagnostic ion is generated exclusively from fragmentation of an acetylated lysine residue [51], and the relative intensity of light and heavy acetyl-lysine immonium ion in MS2 spectra reflects the relative abundance of native and chemically acetylated peptides.

Figure 4. Partial chemical acetylation-based strategies for quantifying acetylation site stoichiometry. (A) Stoichiometry estimation using partial chemical acetylation. SILAC (heavy, red) labeled protein lysate is partially (~5-10%) chemically acetylated (blue). The degree of partial chemical acetylation is determined by spiking in a known quantity of isotope-labeled acetylated peptides, or by measuring the decrease in CP abundance, after chemical acetylation, with relative quantification methods such as SILAC or TMT. Partially chemically acetylated proteins are mixed with native
acetylated proteins that are encoded with a different SILAC label (light, yellow), and proteolyzed with trypsin. The resulting acetylated peptides are affinity-enriched, quantified by MS, and stoichiometry is calculated using the given formula [76]. (B) Acetylation site stoichiometry quantification using serial dilution SILAC (SD-SILAC). SILAC labeled (heavy, red) proteins are chemically acetylated (blue) and serially diluted to prepare dilutions of chemically acetylated proteins. Chemically acetylated proteins are ‘spiked-in’ with native acetylated proteins (yellow) at different dilutions and proteolyzed with trypsin [56]. The resulting acetylated peptides are affinity-enriched with an acetyllysine antibody, quantified by MS, and stoichiometry is calculated using the indicated formula [59].

**Figure I.** Posttranslational modification site stoichiometry. The cartoon illustrates the concept of PTM site stoichiometry using three different examples with different PTM site stoichiometry. Unmodified and PTM-containing proteins are marked as indicated.

**Acknowledgements**

We thank the members of our laboratory for their helpful discussions. The Novo Nordisk Foundation Center for Protein Research is supported financially by the Novo Nordisk Foundation (grant agreement NNF14CC0001). C.C. is supported by a Hallas Møller Investigator Fellowship from the Novo Nordisk Foundation (NNF14OC0008541), the Danish Cancer Society, and the Danish Council for Independent Research (DFF), as well as from the European Research Council (ERC) under the European Union’s Horizon 2020 Research and Innovation Program (grant agreement 648039).

**References**


Spike-in AQUA peptides:

Native AQUA unmodified peptide

BA

m/z Intensity

MS analysis modified unmodified

Native AQUA m/z Intensity

modified peptide

unmodified peptide

Abbreviations:

Abs\textsubscript{native\_mod}: absolute abundance of modified native peptide

Abs\textsubscript{native\_unmod}: absolute abundance of unmodified native peptide

Int\textsubscript{native\_mod}: intensity of modified native peptide

Int\textsubscript{native\_unmod}: intensity of unmodified native peptide

Int\textsubscript{AQUA\_mod}: intensity of modified AQUA peptide

Int\textsubscript{AQUA\_unmod}: intensity of unmodified AQUA peptide

CAQUA\_mod: concentration of modified AQUA peptide

CAQUA\_unmod: concentration of unmodified AQUA peptide

Stoichiometry (\%) = \frac{Abs\textsubscript{native\_mod}}{Abs\textsubscript{native\_mod} + Abs\textsubscript{native\_unmod}} \times 100

Stoichiometry (\%) = \frac{Int\textsubscript{mod} \times f}{(Int\textsubscript{mod} \times f) + Int\textsubscript{unmod}} \times 100

Abbreviations:

Int\textsubscript{A\_mod}: intensity of modified peptide in sample A

Int\textsubscript{A\_unmod}: intensity of unmodified peptide in sample A

Int\textsubscript{B\_mod}: intensity of modified peptide in sample B

Int\textsubscript{B\_unmod}: intensity of unmodified peptide sample B

Int\textsubscript{mod}: intensity of modified peptide in a test sample

Int\textsubscript{unmod}: intensity of unmodified peptide in a test sample

Flyability ratio (f) = \frac{Int\textsubscript{A\_unmod} - Int\textsubscript{B\_unmod}}{Int\textsubscript{B\_mod} - Int\textsubscript{A\_mod}}
**A**

Condition 1: SILAC light (L)

Condition 2: SILAC heavy (H)

Mix, proteolysis → PTM affinity-enrichment → MS analysis

Stoichiometry L = \( \frac{a}{1 + a} \) x 100%

Proportion: L = \( \frac{\text{Ratio}^{\text{prot}} - \text{Ratio}^{\text{CP}}}{\text{Ratio}^{\text{mod}} - \text{Ratio}^{\text{prot}}} \)

Proportion: H = \( \frac{\text{Ratio}^{\text{mod}}(\text{Ratio}^{\text{prot}} - \text{Ratio}^{\text{CP}})}{\text{Ratio}^{\text{mod}} - \text{Ratio}^{\text{prot}}} \) = b

Stoichiometry H = \( (1 - \frac{1}{\text{Ratio}^{\text{CP}}}) \) x 100%

**B**

Proteolysis → Dimethyl-labeling → MS analysis

Intensity → m/z → Intensity

Stoichiometry = \( (1 - \frac{1}{\text{Ratio}^{\text{CP}}}) \) x 100%

**C**

Proteolysis → Phosphatase + → MS analysis

Combine samples, 1st PTM affinity-enrichment → Flow-through

Identification of phosphorylation sites in native sample

Stoichiometry = \( (1 - \text{Ratio}^{\text{CP}}) \) x 100%

Abbreviation:

\( \text{Ratio}^{\text{prot}} \): intensity ratio of heavy (H) and light (L) SILAC samples

\( \text{Ratio}^{\text{CP}} \): intensity ratio of heavy (H) and light (L) SILAC samples

\( \text{Ratio}^{\text{mod}} \): intensity ratio of heavy (H) and light (L) SILAC samples

Identification of phosphorylation sites in native sample

Abbreviation:

\( \text{Ratio}^{\text{CP}} \): intensity ratio of heavy (H)- and light (L)-labeled unmodified corresponding peptide (CP) following their in vitro phosphorylation by a kinase.
A

**Protein-level**

Comprehensive chemical acetylation (with stable isotope-labeled acetic anhydride or NHS-acetate)

Proteolysis

MS analysis, MS1-based quantification (data-dependent acquisition)

<table>
<thead>
<tr>
<th>Int native</th>
<th>Int chemical</th>
<th>Stoichiometry</th>
</tr>
</thead>
<tbody>
<tr>
<td>Int native: intensity of modified peptide in native sample</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Int chemical: intensity of modified peptide peptide in comprehensively chemically modified sample</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**Abbreviations:**

Comprehensive chemical acetylation

MS analysis, MS2-based quantification (data-independent acquisition)

MS1

MS2

**Label-free quantification**

B

**Proteolysis**

Comprehensive chemical acetylation with $^{13}$C-labeled acetic anhydride

MS analysis

<table>
<thead>
<tr>
<th>Int AcK_Im_native</th>
<th>Int AcK_Im_chemical</th>
<th>Stoichiometry</th>
</tr>
</thead>
<tbody>
<tr>
<td>Int AcK_Im_native: intensity of the diagnostic acetyl-lysine (AcK) immonium ion (Im) in native sample</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Int AcK_Im_chemical: intensity of the diagnostic acetyl-lysine (AcK) immonium (Im) in comprehensively chemically modified sample</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**Abbreviations:**

Int AcK_Im_native: intensity of the diagnostic acetyl-lysine (AcK) immonium ion (Im) in native sample

Int AcK_Im_chemical: intensity of the diagnostic acetyl-lysine (AcK) immonium (Im) in comprehensively chemically modified sample
Intensity
m/z

**Stoichiometry**

**Ratio**_{H/L_{mod}} = \frac{P}{\text{Ratio}^{\text{H},\text{mod}} - (1 - P)} \times 100\%

Quantify the degree of partial chemical acetylation (P)

Abbreviations:
- Dilution factor (D): 20, 100, 1000
- Acetylation stoichiometry: 10%, 1%, 0.1%, 0.01%

**Figure 4**

**A**
- Partial chemical acetylation
- MS analysis
- Quantify the degree of partial chemical acetylation (P)
- 'Spike-in' partial chemical acetylated proteins
- Proteolysis
- PTM affinity-enrichment
- MS analysis

**B**
- SILAC heavy (H)
- Partial chemical acetylation
- 'Spike-in' different dilutions of chemically acetylated SILAC heavy (H)
- Native acetylated SILAC light (L)
- Serial dilution
- Acetylation stoichiometry: 10%

Abbreviations:
- Ratio^{\text{H},\text{mod}}: ratio of native acetylated (SILAC light (L)) and chemically acetylated (SILAC heavy (H)) peptide intensity

Stoichiometry = \frac{P}{\text{Ratio}^{\text{H},\text{mod}} - (1 - P)} \times 100\%
Modification site stoichiometry = \frac{\text{Abundance of protein with a site-specific modification}}{\text{Total protein abundance}} \\
\quad = \frac{\text{abundance of modified protein} + \text{unmodified protein}}{	ext{abundance of modified protein} + \text{unmodified protein}}

Modification site stoichiometry:

- 5%: Protein with a site-specific posttranslational modification
- 50%: Protein with a site-specific posttranslational modification
- 100%: Protein with a site-specific posttranslational modification

Unmodified protein