Author Manuscript
Author Manuscript
Author Manuscript
Author Manuscript

HHS Public Access
Author manuscript
Chem Rev. Author manuscript; available in PMC 2015 July 15.

Published in final edited form as:
Chem Rev. 2015 March 25; 115(6): 2376–2418. doi:10.1021/cr500491u.

Quantitative Proteomic Analysis of Histone Modifications

He Huang[†], **Shu Lin**[‡], **Benjamin A. Garcia**[‡], and **Yingming Zhao**[†,*]

[†]Ben May Department of Cancer Research, The University of Chicago, Chicago, Illinois 60637, United States

[‡]Department of Biochemistry and Biophysics, University of Pennsylvania, Philadelphia, Pennsylvania 19104, United States

1. INTRODUCTION TO THE BIOLOGY OF HISTONE POST-TRANSLATIONAL MODIFICATIONS

1.1. Nucleosome and Chromatin

In eukaryotic cells, chromosomal DNA is packaged into a compact structure, chromatin, with the use of four core histones (H2A, H2B, H3, and H4). The fundamental repeating unit of chromatin is the nucleosome, which is composed of an octamer of the core histones, around which ~147 base pairs of DNA are wrapped. Nucleosomes are in turn folded into progressively higher-order structures. Dynamic chromatin remodeling plays a critical role in regulating diverse DNA-based biological processes, such as transcription of RNA, DNA replication, and DNA repair, as well as chromosome condensation and segregation.[1]

The core histone proteins (not histone octamer) are small (10–20 kDa) and highly basic. They are predominantly globular except for their N-terminal “tails”, which are unstructured and protrude from the surface of the chromatin polymer. Amino acid sequence analysis shows that histone proteins are highly conserved in eukaryotic cells from yeast to human, implying that most amino acid residues, if not all, are likely to be important for structure or function. Indeed, studies among histone variants as well as mutational evidence in cancers suggest that a change of a single amino acid residue can lead to very different biological output and even disease, such as cancer.[2]

Histone post-translational modification (PTM), or histone mark, in combination with DNA modifications, histone variants, and ATP-dependent protein complex formation, is used by cells to dynamically modulate chromatin structure and function. Because PTMs alter the properties of the substrate amino acid residue, typically more significant than a mutation, they are likely to affect histone structure and therefore function.[3] Indeed, PTMs are abundant in histones, especially at their N-terminal tails, and have roles in modulating chromatin dynamics and diverse DNA-templated biological processes (Figure 1).[1] Dysregulation of these processes has been intimately associated with the development of diseases such as cancer.[4]

[*]Corresponding Author Tel.: (773) 834-1561. Fax: (773) 702-3701. yingming.zhao@uchicago.edu.
The authors declare no competing financial interest.

Author Manuscript

1.2. Biological Mechanism of Histone PTMs

As of this writing, 20 types of histone PTMs had been reported: phosphorylation, acetylation, monomethylation, dimethylation, trimethylation, propionylation, butyrylation, crotonylation, 2-hydroxylisobutyrylation, malonylation, succinylation, glutarylation, formylation, hydroxylation, ubiquitination, SUMOylation, O-GlcNAcylation, ADP-ribosylation, proline isomerization, and citrullination (Figure 1).[5] In more recent times, known PTM sites on histones have been identified either by sequence-specific antibodies or by mass spectrometry (MS) methods in an unbiased manner.[6] The function and dynamic regulation of these PTMs have been the subject of extensive investigations over the past decade.

Histone PTMs are thought to regulate chromatin structure and function by two mechanisms.[1a,b] First, histone PTMs can directly modulate the packaging of chromatin by either altering the charge state of histones or through inter nucleosomal interactions, thereby regulating chromatin higher-order structure and the access of DNA-binding proteins, such as transcription factors. Additionally, histone PTMs can modify chromatin structure and function either by recruiting PTM-specific binding proteins (also called "readers") and their associated binding partners ("effector proteins") or by inhibiting the binding of a protein to the chromatin. PTM-induced changes in protein interactions between chromatin and its binding proteins are in turn translated into biological outcomes.[7] Proteins are recruited to histone PTMs through direct binding to specific domains. For example, chromo, Tudor, PHD, MBT, PWWP, WD, ADD, zf-CW, BAH, and CHD domains are all known to bind methyllysine,[8] while the bromodomain binds acetyllysine.[9] Proteins containing these PTM-specific binding domains may recruit additional protein factors to execute their functions. Alternatively, they may carry enzymatic activities that can further modify chromatin structure and function.

Histone marks are known to be critical in regulation of diverse DNA-templated biological processes.[1] Interestingly, some of these histone PTMs correlate with transcriptional activation or repression, depending on the types and the locations of the PTMs.[1b,10] To execute DNA-templated processes, histone PTMs coordinate the unraveling of chromatin to carry out specific functions. For example, histone lysine acetylation (K_{ac}) typically correlates with transcriptional activation, while lysine deacetylation correlates with transcriptional repression.[1b,11] Lysine methylation (K_{me}) is implicated in both gene activation (H3K4, H3K36, and H3K79) and transcriptional repression (H3K9, H3K27, and H4K20).[12] As examples, some monomethylation (e.g., H3K9me1 and H3K27me1) is involved in transcriptional activation, while trimethylation at the same sites (H3K9me3 and H3K27me3) is linked to repression.[13] Likewise, some other histone PTMs also correlate with DNA repair (e.g., H2AS129 phosphorylation and H4S1 phosphorylation)[14] and replication (e.g., acetylation).[15] Dysregulation of each step of histone PTMs, including adding the histone marks by a "writer", removing the histone mark by an "eraser", and misinterpretation by a "reader" protein, has shown to be associated with disease, such as cancer.[4a,e]

These histone PTMs are proposed to contribute a "histone code" or "histone language" that dictates the functions of the proteins in gene expression and chromatin dynamics.[1a,c,d]

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

Addition and removal of histone PTMs are regulated by diverse groups of enzymes that were initially identified in the past decade, but still are being discovered in recent times. These enzymes are responsible for adding ("writing") or removing ("erasing") the histone PTM "code". The resulting histone marks are in turn translated into biological outputs by different mechanisms.

Chromatin dynamics are mainly controlled by ATP-dependent chromatin remodeling enzymes/complexes and histone PTMs.[16] The "histone code" can facilitate the recruitment of diverse chromatin remodeling enzymes to regulate chromatin dynamics. Conversely, chromatin remodeling enzymes can also influence the histone PTMs.[17] For example, an ATP-dependent nucleosome remodeling complex, nucleosome remodelling and deacetylation complex (NuRD), can facilitate the deacetylation of the target histones.[18]

Some histone PTMs, if not all, are inheritable during cell division and correlate with gene expression. Therefore, histone PTMs are linked with epigenetic phenomena and are generally considered to be a major type of epigenetic marks.[19]

1.3. Histone PTMs in the Context of Epigenetics and Diseases

Epigenetics has been generally defined as the study of inheritable phenotype changes that do not involve changes in DNA sequence.[20] Dynamic changes in the epigenome are crucial for novel development and differentiation to the various cell types in an organism, as well as for diverse cellular phenomena such as genomic imprinting, paramutation, polycomb silencing, and position effect variegation. Dysregulation of epigenetics programs is associated with diverse physiological conditions and diseases, including but not limited to aging, neurodegenerative disorders, autoimmune diseases, and cancers.[4a,b,d,e]

It is believed that three major types of changes underlie dynamic epigenetic phenomena: histone PTMs, DNA methylation, and noncoding RNAs.[21] The three major epigenetic pathways have independent functions, but also influence each other. For example, MeCP2, a methyl-CpG binding protein, recruits the H3K9 methyltransferase SUV39H1 to target genes, leading to methylation of H3K9 and gene silencing.[22] Likewise, it has been reported that under certain circumstances H3K9 methylation is a prerequisite for DNA methylation.[23]

Histone PTMs have been generally viewed as an important group of epigenetic marks, for several reasons.[1c,19,21,24] First, histone PTMs correlate with global gene expression and are known to be associated with expression of cell-type specific proteins, such as embryonic stem (ES) master regulators (e.g., Oct4) during ES cell differentiation.[19b,21,24a] Second, some histone PTMs are highly stable and inheritable during cell division. For example, arginine methylation and lysine methylation, while reversible, have low turnover rates (at least in certain restricted regions of chromatin and in some cellular states). H3K27me and H4K20me1 are good examples of such epigenetic marks that, at least in some specific chromatin regions, are stably methylated over several generations of cell division.[24a] Third, epigenetic marks do not have to be completely stable. Diverse lines of evidence indicate that DNA methylation is reversible, while its epigenetic nature is unquestioned.[25] Thus, not only histone methylation but also other less stable histone PTMs, such as lysine acetylation, may also contribute to epigenetic phenomena. For this reason, histone deacetylase inhibitors,

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

such as suberoylanilide hydroxamic acid (SAHA) that has been approved as a clinical antitumor drug, are considered as epigenome-modifying agents. Finally, it has been reported that DNA methylation is sometimes dependent on histone PTMs, including lysine methylation.[23] Accordingly, histone PTMs may modulate the epigenetic program indirectly through regulation of DNA methylation.

1.4. The Focus of This Review: Proteomic Analysis of Epigenetic Marks and PTMs

Given the critical roles of histone marks in chromatin structure and functions, it is essential to precisely characterize this type of chemical events, and correlate the modifications with biological outcomes. To this end, we believe that proteomics can contribute to the studies of histone biology at least in five ways: (1) identification and validation of all of the known histone marks; (2) determination of quantitative changes of histone marks under diverse cellular physiology, and disease settings; (3) characterization of crosstalk among histone marks; (4) pinpointing histone marks that can be regulated by a specific “eraser” or “writer”; and (5) identifying binding proteins of histone marks.

In this Review, we try to offer an updated summary of MS-based proteomic approaches, including strategies, techniques, and applications, to address these issues. In addition, we will also review current literature, which uses MS-based proteomic methods for global analysis of nonhistone proteins and identifying substrates and PTM sites whose PTM statuses are controlled by a PTM-regulatory enzyme.

2. MASS SPECTROMETRY AND ITS APPLICATIONS TO IDENTIFY HISTONE PTMs

2.1. Introduction of Mass Spectrometry

Over the past two decades, significant increases in the mass accuracy and resolution of mass spectrometers have brought these instruments to the forefront of protein identification as means to analyze complex protein samples. MS is an analytical technique that measures the mass-to-charge (*m*/*z*) ratio of ions in the gas phase for molecules constituting a sample. A mass spectrometer is an instrument for MS analysis, usually composed of three parts: ion source, mass analyzer, and detector.[26] Typically a sample is ionized into the gas phase in the ion source. The widely used ionization methods for proteins and peptides are electrospray ionization (ESI)[27] and matrix-assisted laser desorption/ionization (MALDI).[28] The sample molecules lose electrons, and gain or lose protons or adduct cations during the electrospray ionization. They are then introduced to a vacuum environment and guided through the mass analyzer, which are a series of electric or magnetic fields. During this process, ions with different *m*/*z* ratios can be separated or filtered as the beam of ions by the electric/magnetic fields. Five major types of mass analyzers are commonly used: time-of-flight (TOF),[29] quadrupole,[30] ion traps,[31] orbitraps,[32] and Fourier transform (FT) ion cyclotron resonance.[33] Finally, the ions reach the detector (usually composed of electron multipliers), and then the detected signals are converted to digital output as a mass spectrum, a plot of different *m*/*z* ratios, and their intensities for sample ions. The generated *m*/*z* ratios can be used to calculate the molecular weights by using the determined number of charges and correcting for the number of attached protons (which is equivalent to *z*) for each *m*/*z* ratio.

Author Manuscript

For complex biological studies, MS is very often coupled with high performance liquid chromatography (HPLC). Typically, the proteins in a complex sample are prepared. This can be a protein lysate form whole cells, or an organelle or a protein complex. The proteins are digested with a proteolytic enzyme, for example, trypsin. HPLC then separates proteolytic peptides in a sample before they are introduced to the ion source in a mass spectrometer. HPLC have two major functions, enhancing concentration of a peptide and reducing sample complexity, therefore offering advantages of higher sensitivity and selectivity. For example, in reverse-phase HPLC, peptide samples bound to C18 resin can be gradually released from the resin, by increased percentage of organic solvent in aqueous solvent. The more hydrophilic a peptide is, the earlier it is released. The eluted peptides are then ionized and introduced into a mass analyzer for MS analysis.

Tandem mass spectrometry (also called MS/MS (or MS^2) or tandem MS) has been applied to protein identification and peptide sequencing for over decades.[34] In tandem MS, the peptide ions are introduced into a mass analyzer, and the *m*/*z* ratio is first determined in MS. In its coupled second stage of MS, the ion specie of interest, called “precursor ion”, is isolated and fragmented by collision with an inert gas (such as nitrogen, argon, or helium atoms) to generate peptide fragment ions (also called “product ions” or “daughter ions”). The *m*/*z* ratios are subsequently determined. Thus, the mass spectrometer alternates between determining the masses of peptides and generating their fragmentation patterns by MS/MS analysis. The most frequently used fragmenting techniques in proteomics studies include collision-induced dissociation (CID),[35] higher-energy collision dissociation (HCD),[36] electron capture dissociation (ECD),[37] and electron transfer dissociation (ETD).[38] Under CID condition, collision with a neutral gas typically leads to amide bond fragmentation along the peptide backbone, generating b- and y-type fragment ions or leading to neutral losses of water and/or ammonia. CID is more effective for small, low-charged peptides. HCD also generates b- and y-type fragment ions, and the b-type ions can be further fragmented to a-type ions or smaller species.[39] As compared to the traditional ion trap CID, HCD does not suffer from low mass cutoff restriction and therefore is useful for observing reporter ions in isobaric tag-based quantification.[40] While CID and HCD fragmentation produces predominantly b- and y-type ions (Figure 2B, C),[41] ECD generates radical cations for a multiply protonated protein/peptide, and ETD transfers electron to a multiply protonated peptide/protein, leading to the cleavage of the N–Cα backbone bonds and to generation of c- and z-type fragment ions.[38] Some PTMs, such as phosphorylation at serine and threonine residues, are labile under CID conditions because these PTM groups compete with the peptide backbone as the preferred site of cleavage. However, ECD/ETD can leave these labile PTMs intact, serving as useful methods to mapping these PTMs in large scale.[42] Additionally, they are effective for the fragmentation of longer peptides or even intact proteins, and thus can be used for top-down proteomics studies.[38,42a,43] Depending on the fragmentation methods, covalent bonds in a peptide are broken, therefore generating a-x, b-y, or c-z ions of the parent peptides (Figure 2).[44] By comparing the series of daughter ion masses, we can deduce, in good cases, the primary sequence as well as any PTMs of the peptide.

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

In most cases, however, it is difficult to directly read out a peptide sequence from the MS/MS spectrum as described above. Nevertheless, the molecular weights of the parent tryptic peptide and its fragment ions are specified by the peptide sequence. Each peptide generates a unique set of fragment ions, thus serving as a mass fingerprint for the peptide. Accordingly, experimental MS/MS data can be used to identify peptide sequence. The resulting spectra are processed by a sequence alignment algorithm for identifying peptides. The technology enables researchers to identify thousands of proteins in a single analysis.

During a sequence alignment, the mass fingerprints (MS/MS data) generated in the HPLC/MS/MS analysis are compared to the theoretical mass fingerprints database of all possible tryptic peptides derived from the protein sequence database, to find the best match (Figure 3). In the past two decades, MS has become the method of choice for identifying proteins and mapping their modification sites.[44b,46] The principles and procedures of such analyses have already been carefully reviewed.[44b,47]

To quickly and accurately analyze large volumes of MS data, many database search engines have been developed. The most popular ones are based on protein sequence alignment of MS data against the known protein sequence databases.[48] Some other software tools use de novo sequencing approaches,[49] which are well-suited for proteomic data mining from organisms without a known genome. Some tools take advantages of known spectra and use spectral library databases.[50] Many software packages are well-established for these purposes: examples include Sequest,[51] Mascot,[52] X!Tandem,[53] pfind,[54] Skyline,[55] Sonar,[56] ProbID,[57] Popitam,[58] and Andromeda.[59] For complete lists of software tools, please refer to the following recent reviews.[48–50,60] Mascot and Sequest are the most widely used commercial search engines. Popular protein sequences database, for example, Uniprot, NCBInr, and International Protein Index (IPI), can be used to construct mass fingerprints database. Both Mascot and Sequest embedded in Proteome Discoverer (Thermofisher Scientific Inc.) have intuitive interfaces, and their latest versions (Mascot v2.2 or higher) now support the peptides identification at fixed false discovery rate (FDR) via decoy database search. As compared to Sequest, Mascot searches the database a little faster, but its HTML report regeneration is time-consuming. The two algorithms usually yield comparable results at the peptide/protein level, although the results from different mass spectrometers may be complementary.[61] A comprehensively comparison between Sequest and Mascot by Kapp et al. demonstrated that Sequest is more sensitive, while Mascot is more specific in peptide identification.[62] Because each individual search engine provides some unique correct matches, combining the results from multiple search engines can serve to improve peptides or proteins identification.[60a]

2.2. Mass Spectrometry for Identifying PTMs

The same procedure described above can be used for identifying PTM peptides and mapping PTM sites. Any time an amino acid residue is modified by a PTM, its chemical structure (or element composition) is changed, which in turn leads to an alteration of its molecular weight. We can consider a modified amino acid reside as a new amino acid residue. Accordingly, the original amino acid residue will have two possibilities, unmodified versus

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

modified. For example, K_{ac} leads to a 42.0106-Da increase in the mass of the modified residue (K_{ac}) relative to the unmodified residue (K).

The same procedure for identifying proteins using MS/MS data can be modified to map PTMs, each of which is associated with a specific mass shift (Δmass) (Figure 4). When there is a PTM, the structure of the substrate residue is changed that induces a mass shift, which is different from its unmodified residue. If it is suspected that a lysine residue has been acetylated, the software can be directed to consider the possibility of a 42.0106-Da accurate mass shift on lysine residues when generating theoretical mass fingerprints during the protein database search. The same set of algorithms previously mentioned for protein identification also has the capability for mapping PTM sites during protein sequence database searching. These algorithms play a key role in using MS data for protein identification and mapping PTM sites. A major weakness of the algorithms lies in the difficulty of searching simultaneously for multiple PTMs (e.g., >10). In addition, a database search with many PTMs will result in exponentially increased search time and a number of false positives.[63]

The most commonly detected PTMs include mono-, di-, and trimethylation on lysine residues (K_{me1}, K_{me2}, and K_{me3}), me1 and me2 on arginine residues (R_{me1} and R_{me2}), acetylation on lysine residues (K_{ac}), phosphorylation on serine, tyrosine, or threonine residues (S_{ph}, Y_{ph}, and T_{ph}), and protein ubiquitination on lysine (K_{ub}). The Unimod database lists almost 1000 different protein modifications that have been detected in MS analyses.[64] Examples of identifying protein PTM by MS are given in Figures 2 and 4. Two spectra from the peptide TKQTAR are shown in Figure 2B and C. One form of the peptide has an additional monomethylation on the lysine residue. The modification adds 14.016 Da of mass, which is detected both at the MS level (the charge +2 precursor ion gains 14.016/2 = 7.008 *m*/*z*) and at the MS/MS level (y5 ion) (Figure 2B, C).

The efficient detection of PTMs by MS highly depends on two major factors: first, the chemical stability/property of the PTM; and second, the abundance of the PTM in the sample. Modifications such as lysine acetylation and methylation are quite stable and usually remain intact during sample preparation and MS experiments. As shown in Figure 2C, the examples of K_{me1} and lysine propionylation (K_{pr}) both stayed with the amino acid residue even after MS/MS fragmentation by HCD. In contrast, volatile modifications like phosphorylation are much less stable. In most CID and HCD fragmentations, the most abundant fragment ions are usually the precursor ion eliminating a phosphate group and a water molecule (97.9769 Da).[65] Although this phenomenon impairs sequence information in MS/MS spectra, it helps for confirming phosphorylation on the precursor ions. Using certain phosphorylation enrichment methods (see below) and optimization of MS/MS instrumentation, many MS-based studies have successfully identified thousands of phosphorylated proteins.[66] For very dynamic modifications such as phosphorylation, SUMOylation, and acetylation, inhibitors targeting eraser enzymes are routinely added into sample collection protocols to reduce PTM loss during sample preparation.[67]

Mass spectrometers can be operated in either positive ion mode or negative ion mode. Negative ion mode has not yet widely been used in peptide detection because it has lower

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

sensitivity and often promotes side-chain cleavages in addition to backbone dissociation.[68] However, this feature can be used to identify some specific type of peptides, such as the peptides with disulfide bonds.[69] Cleavages of cysteine side chain in peptide anions lead to efficient disulfide bond dissociation, and thus provide characteristic signatures in the fragment ion mass spectra.

To accurately map PTMs' localization, for example, determining phosphorylation sites in a peptide with multiple serine/threonine/tyrosine residues, different tools have been developed.[70] For example, Ascore measures site localization probability on the basis of the presence and intensity of site-determining ions in MS/MS spectra.[71] Ascore higher than 20 indicates a site being localized with higher than 99% certainty. Using this method, high-throughput protein phosphorylation localization was analyzed in a fully automated fashion. PTM Score embedded in Maxquant/Andromeda adopts a strategy similar to Ascore.[59] Localization score and probability are calculated after peptide identification. Usually, applying a localization probability cutoff of 0.75 leads to confident localization. Mascot Delta Score is another method for calculating PTM-localization probability by comparing the Mascot ion-score difference between peptide identifications with different site localizations.[72] A Mascot Delta Score of 10 means probabilities of 91% and 9% for two PTM localization (same PTM but localized at different residues, derived from the same peptide), respectively. In addition to these tools, many other scoring modules or softwares are developed for this purpose in recent years, including PhosphoScore,[73] Phosphorylation Localization Score (PLS) in Inspect,[74] SLoMo,[75] PhosphoRS,[76] Phosphinator,[77] SLIP score in Protein Prospector,[78] and D-score.[79]

Currently, there are two complementary methods for MS analysis of proteins, bottom-up and top-down.[80] In the bottom-up method,[47,81] proteins of interest are first digested with a proteolytic enzyme, such as trypsin.[37,57] The resulting proteolytic peptides are analyzed by HPLC/MS/MS in a data-dependent acquisition mode. An alternative strategy for protein characterization is the top-down approach, in which either an entire protein molecule or a large fragment of a protein is analyzed by MS (Figure 7).[80b,82] When the analysis is carried out for a medium-size protein fragment that is larger than a typical proteolytic peptide (e.g., 5–20 amino acid residues), it is called middle-down MS.[83] Because top-down and middle-down use similar mass spectrometers and MS fragmentation methods, we refer to both as top-down in the subsequent sections. The ability to efficiently fragment large protein ions is critical for top-down MS. Both ECD and ETD fragmentation methods have been widely used in top-down experiments.[37,38,84]

As compared to bottom-up proteomics, top-down proteomics is a relatively new method. While potentially powerful, this method currently suffers from several limitations. First, the very complex spectra generated by multiply charged proteins limit the approach to isolated proteins. Peptides are much smaller than proteins. Hence, peptides can be chromatographically separated with much higher resolution. Peptides can be enriched in HPLC for 100-fold or higher in terms of their concentration and have higher efficiency of fragmentation in MS/MS, both of which are critical to higher sensitivity. Accordingly, the bottom-up approach can be used to detect thousands of proteins in a single nano-HPLC/MS/MS analysis in a few hours. Also, it has much higher sensitivity than the top-

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

down approach. In contrast, top-down experiments are more challenging for analysis of a complex protein mixture. Second, the large size of a protein leads to dramatically increased fragmentation possibilities among the peptide bonds, reducing MS/MS sensitivity. Thus, the top-down approach has lower detection sensitivity than the bottom-up approach and typically requires a lengthy data acquisition time to generate a good MS/MS spectrum. This approach has traditionally been used to analyze a single protein or a moderately complex mixture containing a few proteins. Third, the fragmentation efficiency of favored dissociation techniques (ECT, ETD) is low, which limits the ability to couple top-down MS techniques with online separations.[85] Finally, the exact mechanisms of protein dissociation behavior are not fully characterized as those for peptide fragmentation.[38]

Nevertheless, top-down approaches do offer some advantages over bottom-up MS. First, the top-down approach fragments the whole protein, therefore generating information, either PTMs or sequence, for the whole protein structure (e.g., arguably 100% sequence coverage). In contrast, in bottom-up MS, some peptides get lost during digestion and nano-HPLC/MS/MS analysis, leading to lower protein sequence coverage (missing some peptides and their associated PTM sites). Second, the top-down approach can generate useful information about protein variants (or protein isoforms), which are highly similar to each other in the primary sequence. Third, for proteins bearing multiple PTMs on the same molecules, top-down can reveal potential combinatorial PTM crosstalk. For instance, trimethylation on histone K4 and K27 is known to be the bivalent domain in pluripotent chromatin.[86] Genomic studies have revealed that these histone PTMs are both enriched on the developmental genes. Top-down MS was used to investigate if any histone PTMs occur on the same histone H3 tail.[6c]

2.3. Mass Spectrometry for Detecting Novel PTMs

Identifying proteins and mapping PTM sites using MS/MS data relies on powerful sequence alignment algorithms.[87] Popular software like SEQUEST, MASCOT, and Andromeda can efficiently align experimental MS/MS spectra with theoretical peptide fragmentation patterns to find the best match.[44b,87] By specifying a limited number of PTMs (typically less than 10) on certain amino acids, these programs can identify modified peptides and assign the PTM sites.

Nevertheless, identifying novel PTMs poses a challenge. When the software must take into account the possibility of a PTM at an unknown site, the number of possible tryptic peptides is dramatically increased, which in turn increases the database search time and the number of search errors.[63] In addition, such algorithms cannot be used to carry out unrestricted sequence alignment to identify new PTMs that cause a novel Δmass. A novel PTM is most likely to induce an undescribed mass shift in an amino acid residue of interest. Thus, the key to discovering a new PTM is to detect a mass shift that is different from those of the existing PTMs.

Several algorithms have been described to identify peptides with unspecified mass shifts from MS/MS spectra for the discovery of protein modifications and polymorphisms.[63,88] Because the PTMs are not prespecified during the protein sequence database searching, this procedure is called “nonrestricted sequence alignment”.[63,88a–c,89] These algorithms enable

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

searches for all of the mass shifts that are caused by either PTMs or mutations, mainly based on the assumption that the modification or mutation induced mass shifts can be identified from the changes in the precursor ion masses of the peptides and these mass differences can also be localized to certain residues through comprehensive MS/MS spectra alignment. In a typical workflow, the software first generates a list of candidate peptides based on the precursor mass of a target MS/MS spectrum through direct database searching or sequence tag analysis. Next, the software calculates the mass difference between peptide and the precursor ion, and performs complex sequence alignment to determine the matching quality between the target MS/MS spectrum and the peptide carrying the mass shift as modification. As such analysis exponentially increases the search space, they often suffer from ambiguous alignments and noisy background. Careful manual examination is required to eliminate the false positives.

2.4. Mass Spectrometry for Protein Quantification

Apart from accurate identifications of proteins and PTMs, MS can also yield highly precise quantification data. Below we will summarize a number of strategies commonly used in quantitative proteomics.

2.4.1. In Vivo Labeling with SILAC in Cultured Cells and Animals—A few methods have been described to isotopically label proteins in cultured cells and living animals. Isotopic labeling of cells can be achieved using ^{14}N and ^{15}N cell culture media (for yeast)[90] or food (^{15}N-labeled algal cells for rodents).[91] Currently, stable isotope labeling by amino acids in cell culture (SILAC) is more widely used, which is a metabolic labeling approach to isotopically label proteins.[92]

In a typical SILAC experiment, cells of interest are grown under two different conditions, one supplemented with normal amino acids (“light”, $^{12}C_6$,$^{14}N_2$-lysine and $^{12}C_6$,$^{14}N_4$-arginine), whereas the other with stable isotope-labeled amino acids (“heavy”, $^{13}C_6$,$^{15}N_2$-labeled lysine and $^{13}C_6$,$^{15}N_4$-labeled arginine) (Figure 5). The two populations of cells are harvested under the same conditions. Usually equal numbers of cells or equal amounts of proteins are mixed, and the resulting proteins are digested and analyzed in an HPLC/MS system. The usage of stable-isotope labeled essential amino acids guarantees the nearly complete labeling (typically >97%) after at least six cell divisions in cell culture.[92a] If the same peptide is present in both samples, it can be detected as a pair of twin peaks with the same mass difference as that between “light” and “heavy” amino acid residues, if only one arginine or lysine residue is present in the peptide. The peptide levels in two pools of the samples can be quantified from the peak heights (or peak area under curve (AUC) of the twin peaks) from MS analysis, while the peptide identity can be determined from MS/MS analysis of the twin parent peaks. To analyze samples in more than two conditions, different combinations of stable isotopes in an amino acid such as lysine and arginine can be used (e.g., ($^{12}C_6$,$^{14}N_2$-lysine and $^{12}C_6$,$^{14}N_4$-arginine) vs ($^{2}H_4$-labeled lysine and $^{13}C_6$-labeled arginine) vs ($^{13}C_6$,$^{15}N_2$-labeled lysine and $^{13}C_6$,$^{15}N_4$-labeled arginine)). Deuterium is not as commonly used as other isotopes in SILAC, because deuterium atoms can interact with C18 stationary phase and therefore impact its retention during reversed-phase chromatography.[93] This effect can adversely influence quantitation accuracy when performing LC/MS

experiments. The SILAC technique minimizes any technical variations during sample processing and mass spectrometry. Similar to DNA microarray analysis, proteins or peptides that do not change between the conditions serve as internal controls. Since its invention, SILAC has been widely used by the proteomics community.

SILAC has also been expanded from cell culture into intact animals. Doherty et al. described nonsaturated labeling of intact chicken using an $^{2}H_{8}$-valine-containing diet, measuring protein turnover rates in vivo.[94] Yates and his colleagues developed a method to metabolically introduce ^{15}N stable isotope into the proteins of animals, by feeding the animal either ^{15}N enriched or unlabeled algal cells, to quantify proteins in animal tissues.[91,95] Mann and colleagues have successfully developed a mouse SILAC diet with $^{13}C_{6}$-lysine, which has been used to generate SILAC mouse with complete labeling in the F2 generation.[96] By using a SILAC mouse model, in vivo analysis of the entire proteome and PTMs from different tissues became possible.

Similar to SILAC, an emerging method, called "NeuCode (neutron encoding) SILAC", is able to metabolically label and quantify proteins via the subtle mass differences between different amino acid isotopologues.[97] Although this method highly relies on spectrometers with high resolution and accuracy, it greatly expands the throughput of SILAC. Theoretically, up to 39 isotopologues with different combinations of ^{13}C, ^{2}H, ^{15}N, and ^{18}O can be embedded in one +8-Da lysine, which is suitable for highly multiplexed proteome analysis. Recently, this method, in combination with dimethyl labeling, was successfully used to relatively quantify yeast proteome over 18 different conditions.[98]

2.4.2. In Vitro Labeling Using TMT, ITRAQ, and DiLeu—Very often biological samples cannot easily be labeled using in vivo labeling techniques like SILAC (e.g., quiescent or senescent cells that do not divide). An alternative approach is to use chemical labeling reactions in vitro to add an isotopic mass tag, after protein samples are digested by a protease. Typically the labeling reagents consist of three regions: (1) amine reactive group that can react with amine groups on peptides' N-terminal amine and unmodified lysine's ε-amine group; (2) a mass reporter region that releases characteristic reporter ions in MS/MS spectra; and (3) a mass normalization region (Figure 6). Two widely used in vitro labeling reagents for protein quantification are tandem mass tags (TMTs)[99] and isobaric tags for relative and absolute quantification (ITRAQs)[100] (Figure 6). They are now commercially available from ThermoFisher Scientific Pierce Inc. and AB SCIEX, respectively. These isotopic tags are valuable to those applications where in vivo metabolic labeling is not convenient. In addition, *N*,*N*-dimethyl leucine (DiLeu) 4-plex isobaric tandem mass tagging reagents were also developed and used for in vitro labeling and protein quantification.[101] The chemical structures of all of the tags, which are used in a quantification experiment, are identical, except that each contains isotopes substituted at various regions. The difference in the mass of the reporter ion is balanced with a mass normalization group. Each tag is balanced in such a way that the total molecular weights are identical and thus indistinguishable in HPLC separation and during ionization. Practically, samples are first labeled with ITRAQ, TMT, or DiLeu reagents, respectively, and then combined. The resulting samples are analyzed by HPLC/MS/MS. Relative quantification is achieved by measuring intensities of different reporter ions generated during MS/MS.

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

Both in vivo labeling methods (e.g., SILAC) and the in vitro labeling methods have the multiplex capabilities, thus minimizing run-to-run instrumentation variation. The protein quantification from the multiplexing data is typically performed on the basis of the combination of data obtained from multiple peptides of the same protein.[102]

2.4.3. Label-Free Quantification—While the labeling techniques described above are valuable, they are not easily applicable to those studies with more than 10 samples.[103] In addition, when very complex samples are analyzed in a multiplexing experiment, efficiency of identification, accuracy, and precision of quantification suffer from coeluted contaminating ions in the *m*/*z* range of isobaric ions in MS/MS spectra,[104] and additional efforts are necessary to correct for this interference.[105]

In contrast, the label-free approach does not have these shortcomings. This approach enables determination of dynamic changes among dozens or even hundreds of samples. Such analysis is likely valuable to those that involve biomarker discovery based on the analysis of hundreds of clinical samples. Major label-free methods include multiple/selective reaction monitoring (MRM or SRM), data-dependent acquisition (DDA), and data-independent acquisition (DIA).

MRM assay is a highly selective and quantitative method.[106] In this experiment, an MRM assay for a protein of interest is first established. To this end, a protein of interest is subjected to digestion using a proteolytic enzyme, such as trypsin, followed by HPLC/MS/MS analysis. Next, MS/MS spectra of the tryptic digest are screened to identify those peptides with strong signals for both parent peptide and at least three of its fragment ions. Besides, ideal quantitative peptide/transition pairs should have a linear response curve. These ion pairs are selected for MRM assay.

The MRM experiments are typically performed in a triple quadrupole mass spectrometer. In this instrument, the selected precursor ions are isolated in the first quadrupole (Q1) and fragmented in Q2. The resulting fragment ions from the isolated precursor are then selected and quantified in Q3. Monitoring more fragments in Q3 consumes more time. Thus, 3–5 fragment ions are selected for each target peptide to reach a balance between the sensitivity and signal specificity.[106b] Using this approach, dozens or even hundreds of proteins can be quantified in a MRM experiment. The MRM assay is also used to carry out absolute quantification to determine the concentration of a given protein in a sample.[106a,107] In this analysis, isotopically labeled, cognate synthetic peptides with known concentration are spiked into the sample as internal standards to determine absolute concentration for a sample of interest. A response curve is plotted on the basis of the signal of internal standards, and abundances of endogenous proteins are determined accordingly.

Another label-free quantification approach is based on the MS information from DDA, which has two methods that are based on either spectral count or peak AUC of precursor ions. Spectral counting was one of the most widely used label-free quantification methods especially when high-resolution instruments were not applied, which is based on the total number of spectra detected for a protein.[108] In this method, the tryptic digest of a protein mixture is analyzed by HPLC/MS/MS in a DDA mode. Thus, there is no special

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

requirement of the mass spectrometer. In a typical duty cycle of DDA experiment, which can be completed in one or a few seconds depending on the experimental design, the most abundant precursor ions (e.g., top 10–20) are selected for MS/MS. Thus, a short 1 h experiment can generate tens of thousands MS/MS spectra. The resulting MS/MS data are first used for identifying proteins. Subsequent quantification can be achieved by measuring spectral counts for proteins. Another popular quantification method through DDA is measurement of peak area of a precursor ion's MS intensity, in which target precursor ion chromatograms are extracted to provide quantitative information on the peptide of interest.[109] To obtain quantitative result with high reliability, DDA-based label-free quantification typically requires some repetitive MS runs (e.g., 2–3 cycles) of the same sample.

Recently, label-free quantification methods based on DIA, for example, sequential window acquisition of all theoretical spectra (SWATH) MS, have been developed.[110] In a SWATH MS experiment, across the *m/z* range of interest, all precursors are divided into multiple groups by a fixed isolation window (e.g., 25 *m/z* ratios). All precursor ions in each sequential isolation window then were fragmented together, followed by measurement of the resulting daughter ions. This method generates time-resolved MS/MS spectra for essentially all of the detectable analytes within the predefined mass range for precursor ions. A sophisticated method has been developed by the Aebersold group to identify and quantify proteins in complex SWATH-MS data sets, which enables people to explore a nearly complete picture of a sample.[111]

Like any other label-free quantification methods, this method is also associated with both advantages and disadvantages. As compared to SRM/MRM and DDA-based label-free quantification methods, SWATH-MS relies on more powerful mass spectrometers with high resolution, high mass accuracy, and high speed (e.g., AB SCIEX TripleTOF 6600 system).[112] The current data processing tool is computation consuming and less accurate in both identification and quantification. In addition, SWATH-MS is less sensitive than the SRM/MRM-based quantification method.[110e] Despite those limitations, SWATH-MS has many advantages.[110a] For example, (1) SWATH-MS can record nearly complete peptide fragmention signals for precursors generated from a sample, which allows revisiting the data to check any new targets without running new samples. (2) As compared to SRM/MRM methods, SWATH-MS can quantify a significantly larger number of peptides and allows measuring relative and absolute protein levels for hundreds of proteins in a single HPLC/MS/MS run.[110e,112] Besides, it has a much higher proteome pool coverage. (3) As compared to DDA methods, SWATH-MS has higher sensitivity and much better reproducibility. In conclusion, these advantages make SWATH-MS a powerful tool for label-free quantification in the future, when its data processing pipeline is fully optimized.

2.4.4. Top-Down Quantification—The protein quantification can also be performed at intact protein level via the top-down approach. Typically, top-down quantification can be achieved by three methods, which are similar to bottom-up quantification in principle: (1) in vivo labeling with SILAC;[113] (2) in vitro labeling using isobaric mass tags, isotopic formaldehyde, or metal-coded affinity tags;[114] and (3) label-free quantification by calculating the statistically significant differences between two or more experimental

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

conditions.[115] Nevertheless, due to the limited sensitivity of the top-down approach and possible low stoichiometry of PTMs, currently the top-down method has not yet been widely used for quantifying PTMs.

3. IDENTIFICATION OF NOVEL TYPES OF HISTONE PTMs

3.1. General Principle

Historically, histone modifications were detected by experimental accidents. A candidate PTM was afterward anticipated; the experimental strategy was designed to hunt for the histone marks. In early days, modern mass spectrometry technology was not available for analysis of peptides and proteins. Consequently, isotopic labeling and chemical methods were used to generate evidence for a protein modification. As an example, sodium acetate-$^{14}C_2$ was incubated with cells or isolated nuclei. Rapid incorporation of the ^{14}C isotope into histones suggested that the ^{14}C isotope was caused by acetylation modification at the lysine's side chain or N-terminal acetylation (as N-terminal acetylation was already known). The failure of puromycin to inhibit the isotopic incorporation indicates that the process happened after the synthesis of protein.[116] The Pronase digestion of the isotopic histones into amino acids followed by amino acid analysis indicated that the isotopic acetate-$^{14}C_2$ could be incorporated into isotopic acetyllysine.[117]

The complexity of this type of discovery has reduced dramatically with the advances in reagents and tools today (e.g., both chemical and biochemical approaches). MS is a powerful tool that can be used for detecting the modification based on the preassumed mass shift caused by the candidate modification. Alternatively, sequence-independent (or pan) and sequence-specific antibodies against histone PTM of interest can be developed and used to probe the histone marks by Western blotting analysis and immunostaining. As an example, O-GlcNAc modification was discovered decades ago.[118] Recent demonstration of its high abundance in cytosolic and nuclear proteins suggests the existence of the modification in core histones, which was eventually established recently.[119]

An alternative approach to find novel histone PTMs is unbiased screening, without the preassumed mass shifts, using MS. Nevertheless, this is not an easy task, as by mass alone, we do not know the chemical structures of the modified residues bearing the unknown PTMs. Therefore, mass information from the unknown PTMs cannot be incorporated into automated data mining of MS/MS spectra. Additionally, novel PTMs may be present at very low stoichiometry (e.g., less than 0.01%) that would prevent their detection by routine HPLC/MS/MS analysis.

To solve this problem, an integrated stepwise strategy for identifying and verifying novel PTMs in histones has been described, involving four steps (Figure 8):[6a] (1) identifying a possible novel PTM candidate by noting a novel mass shift Δmass caused by a modification by HPLC/MS/MS and nonrestrictive sequence alignment algorithm; the novel Δmass is different from that of any known PTM; (2) calculation of the PTM's element composition based on the accurately determined novel Δmass from high-resolution MS; (3) determining the possible structure isomers that fit the novel Δmass; and (4) pinpointing and verifying the novel PTM by chemical and biochemical methods.

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

More attention should be paid to methyllysine-containing peptides if singly charged ions were selected in MS/MS, especially for the identification of novel types of histone PTMs. The methyl group can migrate from the side chain of methyllysine to the C-terminal arginine residue or adjacent serine residue.[120]

3.2. Core Histones Are an Excellent Source for Studying PTMs

Histones have been used as model proteins for studying chemistry and biochemistry of PTM pathways. They are abundant in eukaryotic cells and can be easily isolated from cells or tissues.[121] Additionally, they are decorated with multiple types of PTMs, more than any other cellular proteins. Despite extensive studies and significant progress in our understanding of histone marks, it is unknown whether other types of histone marks remain to be discovered, and, if so, how many undiscovered types exist. These novel PTMs might contribute to epigenetic regulation.

Emerging evidence supports the hypothesis for undescribed histone PTMs. Novel PTMs in histones have been identified very recently. For example, the Zhao group recently described histone lysine propionylation,[122] butyrylation,[122] crotonylation,[6b] 2-hydroxyisobutyrylation,[6a] succinylation,[123] and malonylation.[123b] Additionally, histone O-GlcNAc modification was recently identified and confirmed by the Hart group.[119]

3.3. Detection of a Novel Mass Shift That Is Caused by a New PTM

To detect novel mass shifts, the protein lysate of interest is first digested by a proteolytic enzyme and then analyzed by HPLC/MS/MS. Depending on the experimental goal, proteolytic digest, typically using trypsin, of either a protein cell lysate or a single protein can be used for MS analysis. Ideally, a high-resolution mass spectrometer is used so that the parent peptide's mass and the mass shift caused by a modification can be accurately determined, with accuracy less than a few ppm. A major advantage of high-resolution mass spectrometric analysis is its high resolution and high mass accuracy.[124] For example, the Orbitrap Fusion has a resolving power of up to 450 000 (fwhm) at *m*/*z* 200, and the accuracy is less than 1 ppm using internal calibration (http://planetorbitrap.com/orbitrapfusion).[125]

The acquired MS/MS spectra will be analyzed by a nonrestrictive protein sequence alignment algorithm to hunt for novel mass shifts. Several algorithms, including InsPecT,[88a] PTMap,[63] and PILOT_PTM,[89a] are available to carry out nonrestricted sequence alignment for spotting new modifications. The Δmass of the modified residue will be obtained from the protein sequence alignment. With the high-resolution mass spectrometers, the detected mass shift will be high enough to calculate element composition responsible for the detected mass shift.

3.4. Deduction of Chemical Structure for a Novel PTM

3.4.1. Elucidation of the Molecular Formula for the PTM Chemical Moiety—

Detection of an unknown mass shift is the first step in the long journey to identify a novel PTM. To deduce the chemical structure responsible for a detected mass shift, the next step is to elucidate the molecular formula of the PTM chemical moiety. The accurate mass of the modified residue containing the novel PTM moiety will be used to deduce the molecular

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

formula (i.e., the elemental composition) of the modified residue and PTM moiety. The mass accuracy is dependent on the mass spectrometer and MS method used for MS analysis. Typically, the higher is the mass accuracy of the detected mass shift, the fewer possible elemental formula exist for the PTM candidate. More than one candidate formula usually fit each detected mass shift. General organic chemistry is then used to deduce which element composition is reasonable that follows the existing rules of organic chemistry.

Software to deduce the element composition from the mass is publically available (http://www.chemspider.com/PropertiesSearch.aspx). In this method, the detected mass shift and mass accuracy are used as inputs. This method for defining elemental composition has become routine and has been widely used in the fields of natural product chemistry, organic chemistry, and metabolomics.[126]

3.4.2. Structure Candidates for the Deduced Molecular Formula—The elemental composition derived from high-resolution MS should conclusively determine whether a previously undescribed PTM has in fact been discovered. Nevertheless, knowledge of the elemental composition does not provide information about the arrangement of the atoms within the moiety. To address this issue, ideally, NMR spectroscopy can be used to determine the chemical structure of the novel PTM. Nevertheless, the standard two-dimensional NMR experiments typically require at least 10 μg of a tryptic peptides (e.g., 5–20 residues), even with a sensitive capillary microcoil NMR (CapNMR) probe.[127] Unfortunately, it is unlikely to obtain this amount of the modified peptide from an endogenous protein.

An alternative solution is to deduce all of the reasonable structure isomers based on the molecular formula. We can then synthesize each of structure isomers and use them to carry out pairwise comparison between the in vivo peptide and each of the synthetic isomers by the chemical and biochemical methods to pinpoint the chemical structure for the in vivo PTM.

3.5. Validation of Novel Types of Histone PTMs

3.5.1. Chemical Validation of a Novel PTM—In principle, all of the possible structure isomers should be deduced on the basis of the elemental composition, and each one of them should be synthesized prior to have pairwise comparison between a synthetic one and an in vivo one by MS and HPLC. However, it is generally accepted that, if two peptides have the same MS/MS fragmentation patterns and are indistinguishable in HPLC chromatographic profiles, then they have identical structures.

For example, we identified a modified H4 peptide, DAVTYTEHAKR, containing a mass shift of +86.0354 Da at its lysine residue (H4K77). The only reasonable elemental composition responsible for this mass shift was $C_4H_7O_2$ (mass shift plus one proton) using ±0.02 Da mass tolerance and a maximum of 2 nitrogen atoms. According to the formula, there are five possible structures for the lysine modification that induces a mass shift of 86.0368 Da (theoretical mass shift): 2-hydroxyisobutyryl (K_{hib}), 2-hydroxybutyryl (K_{2ohbu}), 3-hydroxybutyryl (K_{3ohbu}), 3-hydroxyisobutyryl (K_{3ohibu}), and 4-hydroxybutyryl (K_{4ohbu}) groups. Using HPLC coelution experiment, we showed that only K_{hib} would coelute with

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

the endogenous peptide (Figure 9A).[6a] Furthermore, K_{hib} was confirmed by MS/MS experiments between the in vivo peptide and its corresponding synthetic peptide (Figure 9B).

3.5.2. Immunochemical Validation of a Novel PTM—While chemical experiments using MS/MS and HPLC are critical to pinpoint a new PTM among multiple structure isomers, it is not sufficient to fully establish a new structure. Structural isomers with a small chemical difference could lead to almost identical MS/MS spectra and the same retention time in HPLC analysis. Accordingly, it is necessary to further verify a new PTM using an orthogonal method, such as immunochemistry. To address this issue, it is important to develop pan (or sequence-independent) antibodies against a PTM and sequence-specific anti-PTM antibodies to confirm the in vivo modification using Western blotting analysis and/or immunostaining. Because immunochemical signals can be caused by nonspecific binding to other proteins or other modifications, caution should be taken to ensure that the signals are from the PTM of interest instead of artifacts. To this end, it is important to include adequate controls, for example, competition experiment using a randomized peptide library containing a fixed PTM of interest or an antigen peptide bearing the PTM of interest. In addition, the quality of antibody reagents should be fully characterized using competition experiments and binding assays.

Three independent experiments using HPLC coelution, MS/MS experiments, and Western blotting analysis should firmly establish a novel PTM. The chemical and immunochemical methods are complementary and should be able to address the weakness of each method. We would suggest carrying out all three types of experiments as long as the synthesis of the modified peptides is feasible. Indeed, the three methods have been recently used for detection and validation of several histone PTMs.[6a,b,123a,128]

3.5.3. Additional Biochemical Validation for a Novel PTM—Besides the three major approaches discussed above, some other experiments can provide additional evidence for a new modification and are critical to the biology and biochemistry of a new type of histone marks, including the following: (1) The first is identification of additional peptides and proteins bearing a new PTM. This can be typically carried out by immunoprecipitation of peptides bearing a PTM of interest with an anti-PTM antibody, followed by HPLC/MS/MS analysis of the enriched peptides. Using such a method, thousands of acetylation and succinylation peptides were identified.[129] (2) Next is identification of a possible precursor for the novel PTM. One example is the identification of the precursor for crotonylation. Two lines of evidence suggested that crotonyl-CoA is a precursor for lysine crotonylation. The first, lysine crotonylation, is enhanced when cells are treated by crotonylate. Second, crotonyl-CoA exists in cells that can be stimulated by crotonate.[6b] A candidate enzyme, AceCS1, has been suggested to catalyze the synthesis of short-chain fatty acid CoAs (such as crotonyl-CoA).[130] However, it may take years to identify a nonconventional precursor for a PTM of interest. For example, acetyl phosphate was identified as a precursor for lysine acetylation in bacteria until recently,[131] while the modification was discovered decades ago.[116] (3) Third is mapping chromatin localization and the possible contributions of the novel PTMs to epigenetics regulations. This can be carried out by chromatin

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

immunoprecipitation (ChIP)-DNA sequencing (seq) and ChIP-qPCR experiments. Specific localizations of the new histone marks would suggest an intrinsic regulatory mechanism instead of random chemical reactions between the candidate CoA and the lysine's ε-amine group. (4) Next is identification of regulatory enzymes for the PTM of interest. The well-known examples are the enzymes for lysine acetylation and lysine methylation that add and remove these marks. They play critical roles in epigenetic regulation and in functional studies of two groups of histone marks. The two modifications were discovered in 1960s. Their biology started to take off in the 1990s, when their first regulatory enzymes were discovered.[132] (5) The final is functional studies of PTM's nonhistone substrates. As soon as regulatory enzymes are known, functional consequences of a PTM can be studied in a nonhistone substrate of interest using conventional methods.

4. DETECTION AND QUANTIFICATION OF HISTONE PTMs

As introduced above, MS is an unbiased approach of analyzing histone PTMs. Unlike antibody-based techniques, MS methods enable simultaneous detection and quantification of PTMs.

4.1. Detection of Histone PTMs with High Sequence Coverage and Sensitivity

Analysis of histone PTMs is typically carried out by a bottom-up approach. Two issues are critical for detecting as many as possible histone marks in a sample of interest: high sequence coverage and high detection sensitivity. In a typical experiment, the histone proteins are first digested with a proteolytic enzyme, followed by HPLC/MS/MS analysis. During this analysis, proteolytic peptides can be lost due to three reasons: (1) They are too sticky in an eppendorf tube or HPLC column. (2) They bind too strong (which is usually caused by long peptide length, e.g., longer than 25 amino acid resides) to a C18 HPLC column so that it is difficult to be eluted from the column. In bottom-up MS, it is difficult to detect those peptides longer than 25 amino acid residues when a C18 HPLC column is used. (3) Some peptides are too hydrophilic (typically for short peptides) to be retained in a C18 HPLC column so that they are eluted in solvent front with no enrichment and with low detection sensitivity. Therefore, it is almost impossible to have 100% sequence coverage when a protein's digest is analyzed by HPLC/MS/MS. The PTM information in those missed peptides cannot be recovered.

Histones are a special class of proteins as they are highly rich in lysine and arginine residues. Thus, tryptic digestion can lead to many short, hydrophilic peptides that cannot be retained in reverse-phase HPLC and thus detected. In addition, histones carry remarkable numbers of PTMs, and many of these PTMs are on lysine residues. These features make histones intolerant to the most efficient and widely used protease, trypsin, which cuts at the carboxyl side of lysine and arginine residues unless they are followed by a proline. However, two factors make the tryptic digested histone products more complicated. (1) Trypsin would digest histones into very small peptides with poor chromatographic retention because of the common adjacency of lysine and arginine residues in histones. (2) The frequent modifications on lysine residues reduce digestion efficiency or completely block trypsin digestion.

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

Fortunately, two approaches can be used to overcome these two issues. First, multiple proteases can be used to digest histones. For instance, Arg-C protease, which digests at the carboxyl side of arginine residues, has been used for MS analyses of histones.[133] Glu-C protease was used for digesting macroH2A,[134] H2A and H2B variants,[133a] and H3.[135] Analysis of histone peptides derived from multiple proteolytic digestions enables improved sequence coverage.

The other major approach to improve sequence coverage is chemical derivatization by acylation reaction at amine groups. The derivatization usually reacts with the ε-amine group of lysine residues, blocking trypsin digestion at the lysines. The trypic products from derivatized histones are more hydrophobic and behave better in reverse-phase HPLC, enabling detection of the corresponding peptides. Deuterated acetic acid was one of the earliest reagents used to derivatize histones.[136] The derivatization adds a deuterated acetyl group to originally unmodified lysine residues. Because the acetyl groups contain deuteron instead of hydrogen, which causes a 3-Da mass shift, they therefore can be distinguished from endogenous acetylated lysine residue. Another chemical derivatization reagent, propionic anhydride, is much more widely used. Developed by Garcia and Hunt et al.,[137] derivatization of tryptic peptides with propionic anhydride not only occurs at unmodified lysine residues, but also at monomethylated lysines and N-termini of tryptic peptides (Figure 10).[67a] Propionyl groups impart additional hydrophobicity than acetyl groups. For histone peptides without lysine residues, the N-terminal propionylation can improve HPLC retention for short peptides. Propionylation on histones has been widely adapted and used by many different groups.[123b,138] With propionylation, more than 80% of histones sequence coverage can be achieved.[6b] However, a limitation of this method is endogenous propionylation sites are masked. To resolve this problem, isotopic propionic anhydride can be used to distinguish endogenous propionylated lysine residues. It causes 3-Da ($^{13}C_3$-propionylation) or 5-Da (D_5-propionylation) mass shifts in a fashion similar to the deuterated acetylation method mentioned above.[139] Recently, modified protocols using *N*-hydroxysuccinimide ester propionate instead of propionic anhydride have also been developed.[140] By analysis of histone peptides generated by multiple proteolytic enzymes and by chemical derivatization, more peptides can be detected, thus improving sequence coverage.

In addition to sequence coverage, the other key parameter is the sensitivity of the HPLC/MS/MS system. Many histone marks are present in low stoichiometry, for example, 1% or less.[141] As an example, H3K4me3 is a histone mark associated with active gene expression. While widely studied, this histone mark is not easily detected in many cell lines due to its low stoichiometry.[142] Accordingly, the HPLC/MS/MS system should be optimized with a set of standard peptides so that the maximum detection sensitivity is achieved (please see section 7.4 on the optimization of the HPLC/MS/MS system). In an ideal situation, any histone mark with a stoichiometry of 0.01% or higher should be detected.

4.2. Detection of Multiple Histone PTMs

In HPLC/MS/MS analysis, the MS instrument acquires mass spectrometric data regardless of peptide sequence and PTMs, as long as they are present in the sample. Accordingly,

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

subsequent sequence alignment of MS/MS data is critical to extract PTM information. Popular software, such as Sequest,[51,143] Mascot,[52] and Andromeda,[59] can specify a limited number of PTMs (typically less than 10) on certain amino acids. Because histones have 20 known types of PTMs, it is difficult to use the algorithms to analyze MS/MS data in one experiment to scan all of the histone marks. Two approaches can be used to address this issue. First, we can divide the 20 types of histone marks into several sets, leading to several sequence alignment experiments. However, this approach may miss those peptides containing multiple PTMs, which are present in both sets of PTMs. An alternative solution is to use unrestricted sequence alignment such as InsPecT,[88a] PTMap,[63] PILOT_PTM,[89a] MODa,[89b] and some specialized algorithms for top-down or middle-down tandem mass spectra.[144] It is not necessary to prespecify PTM during protein sequence alignment. However, this approach can have more false positives than restricted sequence alignment. Accordingly, careful manual examination of the identified peptides is necessary to ensure high accuracy.

As an example, Tan and his colleagues identified 130 histone marks in HeLa cells using PTMap and Mascot, when tryptic digests were analyzed in HPLC/MS/MS, with or without lysine propionylation, including 28 lysine crotonylation marks.[6b] Surprisingly, the study also identified new histone sites of known types of histone marks, including 18 new lysine monomethylation sites and 8 new arginine monomethylation sites.

Multiple histone PTMs can be detected in one single nano-HPLC/MS/MS analysis. Taking histone H3 and H4 as examples, the N-terminal tails of H3 and H4 are very basic and can be highly modified. Some tryptic peptides from histone H3 and H4 proteins have less than 6 residues, which are highly hydrophilic and relatively difficult to be detected in HPLC/MS/MS analysis. The bottom-up approach can typically give in-depth analyses of the peptides and PTMs (Table 1). For instance, for the histone H3 9–17 peptide, K9, S10, and K14 residues are frequently modified, providing a total of 20 different forms of this peptide from a sample. Eight of the most abundant forms of this peptide were extracted from a histone run (Figure 11). Reverse-phase HPLC efficiently separated these modified peptides including two isobaric pairs of PTMs that have the identical *m*/*z* after propionylation: unmodified and K9me1K14ac, and K9me3K14ac and K9me2.[67a] Therefore, the use of reverse-phase HPLC is very helpful to resolve peptide isoforms.

The aforementioned experimental procedure can efficiently detect multiple PTMs in histones, including detection of multiple PTMs in a peptide (Table 1). However, when two or more histone marks are present in different proteolytic peptides, this approach cannot tell if the marks are present in the same molecule. In this case, the top-down or middle-down strategy can be used to examine the coexistence of multiple PTMs in a protein of interest.[82a,83,145] For example, middle-down MS/MS analysis of 1–50 N-terminal tail peptide from histone H3.1 identified H3K23ac and H3K27me2 (Figure 12).

The bottom-up and top-down approaches are complementary. The bottom-up approach requires less sophisticated analysis of MS/MS data and offers advantages of high sensitivity. On the other hand, the top-down approach has low sensitivity. Yet it has potential to generate all of the modifications in a protein of interest.

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

Multiple PTMs can be present in a single amino acid residue, for example, methylation, acetylation, crotonylation, propionylation, and butyrylation at lysine residues. They are mutually exclusive in a single protein molecule. However, they can be present in the same protein with differential stoichiometry. In addition, different PTMs can recruit different binding proteins, and then exert diverse biological functions.[147] The protein of interest can be analyzed by either bottom-up or top-down methods to identify multiple PTMs in a residue of interest.[6b,148] Alternatively, for large-scale analysis, serial enrichment method followed by bottom-up approach allows the detection of multiple PTMs on a single localization. Using this method, Mertins et al. sequentially enriched phosphorylated, ubiquitylated, and acetylated peptides with metal-affinity or immunoaffinity techniques from tryptic proteins in bortezomib-treated human leukemia cells.[149]

4.3. Quantitative Proteomics for Determining Changes of Histone PTMs

As mentioned earlier, dynamic changes of histone marks are associated with many cellular processes such as transcription and DNA damage. Accordingly, it is becoming more frequently necessary to quantify changes of histone marks between two or more samples. Examples often include a comparison between a wild-type animal versus a mutant animal, undifferentiated stem cells versus differentiated stem cells, and among samples from different patients and cellular perturbations.

The methods for protein quantification (as described in section 2.4) can also be used for determining dynamic changes of histone marks. In these methods, a suitable labeling technique, either in vivo labeling with SILAC or in vitro labeling with TMT or ITRAQ, is needed for accurate quantification of histone marks. These methods have been used to quantify changes of histone marks in a number of studies.[6a,123c,150]

In addition to in vitro labeling with TMT or ITRAQ reagents, lysine acylation reaction with isotopic compounds has also been used for relative quantification of histone PTMs. For example, isotopic propionic anhydride (e.g., H_{10} vs D_{10} or $^{12}C_6$ vs $^{13}C_6$) can be used to label two pools of proteolytic peptides from core histones, respectively, by acylating amine groups, either at the N-termini or unmodified lysine side chain. For the H_{10}/D_{10} pair, in each round of propionylation reaction, 5 hydrogen/deuteron atoms from propionic anhydride are added onto the peptides (Figure 10). Therefore, the resulting histone peptides are labeled by a hydrogen- or deuteron-propionyl group, respectively, having a mass difference of 5 Da. The relative quantifications of the peptide twin peaks, “light” and “heavy”, are typically achieved by measuring the signal areas of precursor ions from extracted ion chromatogram. Using this method, Sridharan et al. determined dynamic changes of several histone marks on H3 and H4 during reprogramming of somatic cells into induced pluripotent stem cells (iPSCs).[151] Dai et al. quantified changes of histone K_{hib} during spermatogenesis.[6a]

Label-free quantification is another powerful method to relatively quantify histone PTMs. As compared to the stable isotopic labeling methods, this method is cost-effective. Currently, most label-free quantification experiments are based on calculation of the peak areas of the modified-peptide parent ions.[67a] In this type of analysis, all different modified peptides and their corresponding unmodified counterparts are considered. Specifically, the peak areas of the unmodified and all differentially modified forms of a histone residue of

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

interest (e.g., H3K9 residue) are summed and designated as 100%, and an individual form of a histone peptide is calculated as a fraction of the summed peak areas (e.g., dividing the peak area of the individual form by those from all of the histone peptides bearing the amino acid residue of interest).[67a] This method allows comparison across multiple experimental conditions. Also, some unmodified histone peptides (e.g., HLQLAIR from H2A, YRPGTVALR from H3, VFLENVIR and ISGLIYEETR from H4) can be used as internal standards.[138i,152] Using this method, Peters et al. relatively quantified all possible methylation states for H3K9 and H3K27, and demonstrated that H3K27me and H3K9me3 are selectively enriched in pericentric heterochromatin.[137a]

For label-free quantification, an important issue is the varied MS detection efficiency caused by PTMs, for example, variant ionization efficiency of peptides with the same sequence but different PTMs. This difference can be normalized using synthetic peptides that carry PTMs.[153] For example, Lin et al. systematically analyzed 93 histone peptides and found that the detection efficiency variation is a widespread problem. The authors then showed that the biases can be effectively corrected by applying correction factors generated by spiking in synthetic peptide standards (internal correction), or by using information derived by independent MS analyses of the standard synthetic peptide mixtures (external correction).[153] Because diverse PTMs combinations generate a huge number of multiply modified peptides, synthesis of all of these peptides is not feasible and not cost efficient. However, this method is still effective on a small scale.

4.4. Determining the Stoichiometry of Histone PTMs

Since the early 2000s, several MS-based strategies were developed to determine the stoichiometry of PTMs, and these strategies can be applied to studying both histone and nonhistone PTMs.

1. First is absolute quantification (AQUA) of modification states. The relative quantification determines dynamic changes of a protein or PTM by comparing peak areas of the targeted molecules in two or more conditions. However, this type of experiment cannot tell the absolute concentration of an analyte or the stoichiometry of a PTM event. In 2003, the Gygi group reported a method for absolute quantification of proteins by using an isotopic synthetic peptide as an internal standard.[154] With a synthetic isotopic modified peptide, individual stoichiometry of a PTM site can also be calculated. This strategy was further refined by Steen and his colleagues. They reported an isotope-free quantification strategy to determine the protein phosphorylation stoichiometry, either in a single protein or in a protein mixture.[155] This method is based on comparing the ion currents of both phosphopeptides and their corresponding unmodified counterpart. Recently, the AQUA method was also applied for determining absolute stoichiometry of lysine acetylation.[156]

2. Next is SILAC-based peptide/protein quantification. A SILAC-based method was developed for determining the stoichiometry of global PTMs.[157] In this method, “heavy” and “light” ratios of the modified peptide, its unmodified counterpart, and modified protein are generated in HPLC/MS/MS analysis. The three parameters are

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

used to calculate absolute stoichiometry for a modification of interest. This method was used to quantitative analysis of phosphoproteomes, determining the stoichiometry of more than 5000 phosphorylation sites that include a bunch of histone phosphorylation sites.[157]

A recent study using a SILAC-based method determined the stoichiometry of lysine succinylation (K_{succ}), histone K_{hib}, and histone lysine crotonylation (K_{cr}).[6a,123c] From these studies, 32% and 56% of the identified K_{succ} sites were found having stoichiometry higher than 10% in SIRT5 wild-type (WT) and knockout (KO) MEF cells, respectively. The stoichiometries of four Khib sites, H3K79, H2BK108, H4K91, and H1K62, in synchronized G2/M HeLa cells are 1.45%, 1.54%, 5.33%, and 7.79%, respectively, which are comparable to or even higher than that of many histone K_{ac} marks with known biological functions.

3. Third is isotope labeling in vitro. The PTM stoichiometry can also be determined by introducing an isotope into the modified residues. To determine the phosphorylation stoichiometry of protein phosphorylation on a large scale, the proteolytic protein lysate was divided into two identical aliquots. One aliquot was treated with a phosphatase to remove the phosphate group. Treated and untreated samples then were chemically labeled by reductive dimethylation with isotopic "heavy" and "light" formaldehyde, respectively, and then mixed. The resulting mixture was analyzed by HPLC/MS/MS analysis. The absolute stoichiometry of each phosphorylated site is encoded in the ratio of "heavy"/" light" species ((1–1/ratio) × 100%).[158] The in vitro deuteroacetylation can label unmodified lysine in samples and is distinguished from endogenous acetylation, which has been used to quantify the stoichiometry of acetylation at proximal lysine residues or specific histone PTMs.[159] Recently, this method was further developed by the Denu group to quantify the stoichiometry of site-specific acetylation in the entire proteome of *E. coli*.[160] After tryptic digestion, "heavy" (in vitro isotopically labeled) and "light" (originally acetylated) acetyl-lysine pairs across the entire proteome were determined. The ratios of the isotopic pairs were used for calculating stoichiometry. To increase the sensitivity of the quantification experiment on lysine acetylation on histone H3, this method was combined with acid-urea gel separation by Cieniewicz et al.[161] In this study, the acid-urea gels separated distinctly acetylated isoforms of histone H3 as a "ladder", then each band was excised and all unmodified lysine residues were deuteroacetylated. The acetylation levels were calculated using the ratio of "light" (endogenous) and "heavy" (chemical) acetylation signals detected by MS.[161] In addition, the Hsieh-Wilson group developed a method for quantification of individual O-GlcNAc modification stoichiometry.[162] In this method, O-GlcNAc modified proteins are chemoenzymatically labeled, attached with a PEG tag, and visualized by SDS-PAGE and immunoblotting. In this assay, the proteins with different O-GlcNAc site occupancy have different molecular weights and appear in different bands, which can be used to determine the stoichiometry.

4. The final is label-free quantification. As mentioned in section 4.3, label-free method can be used for relative quantification of histone PTMs. In this method, all

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

forms of a peptide need to be identified, and the ionization efficiency needs to be normalized because it varies among differently modified peptides. This method can be used to produce a rough estimation of histone PTM stoichiometry. For example, LeRoy et al. used this approach to determine the stoichiometry of a variety of histone PTMs from a large panel of cancer cell lines.[142] The stoichiometries of PTMs on four histone lysine residues are picked up and averaged (Figure 13). Among the four methylated or acetylated residues examined, the PTM stoichiometries varied within a wide range (e.g., H3K9me2 > 30% and H3K4me2 ≈ 1%).

In addition to the DDA-based methods, SRM/MRM-based label-free method was also used to determine stoichiometries of PTMs.[163] As compared to the former, the latter is more sensitive, which enables the detection of histone PTMs with very low stoichiometries. For example, Drogaris et al. built a MRM-based label-free method to determine the stoichiometry of H3K56ac.[164] In this study, a very low stoichiometry of H3K56ac (roughly 0.04%) was determined by using the peak areas of both the modified and the unmodified peptides in a MRM analysis.

4.5. Dynamic Flux Analysis of Histone Proteins and Modifications

To study the dynamics of histone proteins or PTMs, a simple approach is to analyze samples at multiple time-points during a biological process, and to quantify histone proteins or PTMs in steady-state experiments. For such experiments, the methods mentioned in sections 2.4 and 4.3, for example, in vivo labeling with SILAC, in vitro labeling with TMTs/ITRAQs/DiLeu, or label-free quantification, can be used and then provide some temporal information.[90a,165]

Although these methods are powerful to quantify histone proteins or PTMs in steady-state experiments, these techniques cannot measure the cellular kinetics, or dynamics/flux, of histone proteins and PTMs, because the MS cannot distinguish histone PTM and protein turnover. For example, if H3K9me2 increased from a few cellular conditions, questions would be whether the increase of H3K9me2 was caused by changes of other histone marks. Obviously, the methods described above are not possible to determine the kinetic flux of histone proteins and PTMs to answer these questions.

Fortunately, some methods are able to resolve this issue. First, to measure the kinetic flux of histone proteins (e.g., protein turnover experiments), one strategy is incorporation of metabolic analogues, for example, the methionine surrogate azidohomoalanine (Aha), in nucleosome turnover experiments.[166] Aha can be conjugated to alkynyl group containing reagents via click chemistry, which allows facile isolation of new synthesized histones.[166,167] For example, N-terminally Flag-tagged histones H3.1 and H3.3 were expressed for a period of time under the control of a tetracycline-inducible promoter, which was regarded as "old" histones. To investigate the splitting of histone H3–H4 tetramers during genome duplication (e.g., epigenetic inheritance), $^{13}C_6$,$^{15}N_2$-labeled lysine was used to metabolically label the newly synthesized proteins, including nontagged histones.[168] After affinity purification, the results indicated that significant H3.3–H4 tetramers could split, but this phenomenon was not observed in H3.1–H4 containing tetramers.

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

Second, to measure the cellular kinetic flux of histone PTMs, as we discussed above, the steady-state analyses cannot distinguish histone PTM turnover at both protein and residue levels, for example, at the protein level when the histone proteins are synthesized and degraded, and at the residue level when the PTMs are added or removed. To overcome this limitation, metabolic labeling of histone PTMs with small molecule metabolite precursors was used. Historically, this approach has been heavily used as stated above, using radio isotope-labeled metabolites to study protein and PTM turnover. Currently, various stable-isotope labeled metabolites can be used for this type of pulse-chase experiments to specifically label histone PTMs. For example, a stable isotope of methionine, $^{13}CD_3$-methionine, can be used to label histone methylation sites in vivo ("heavy" methyl SILAC, hmSI-LAC),[169] as it can be converted to the sole biological methyl donor, $^{13}CD_3$-*S*-adenosyl-methionine, and then used by histone methyltransferases. Because there are only a few methionine residues on histones,[170] the majority of histone peptides are not labeled by the "heavy" methionine except those peptides with methylated residues. This approach was used to determine the dynamics of several histone methylation sites.[171] It was found that histone methylation turnover was generally slow (e.g., hours), but heavily depended on the modified residue. In addition, methyl marks associated with active gene such as H3K36 turned over much faster than those associated with silenced genes (e.g., H3K9). Similarly, comprehensive steady-state methylation and demethylation rate constants were derived using hmSILAC and a targeting MRM approach providing the information for an integrated kinetic model that was shown to correctly predict observed abundances.[138k]

As canonical histones are primarily generated and incorporated into chromatin during the S phase of cell cycle,[172] strategies combining metabolic labeling and synchronized cell populations can be used to pulse-chase label and therefore distinguish the newly synthesized histones. This type of experiment enables one to distinguish modifications between newly synthesized and old histones. Using this method in combination with top-down MS, it was found that newly synthesized histone H4 becomes progressively methylated at K20 during the G2, M, and G1 phases of the cell cycle, and the majority of K20 residues on new H4 are dimethylated within two to three cell cycles.[173]

For histone acetylation, dynamic incorporation of ^{13}C-labeled acetyl groups onto specific histone lysines was quantified, with three acetyl-CoA generating sources, $^{13}C_6$-glucose, $^{13}C_5$-glutamine, and $^{13}C_2$-acetate, showing that $^{13}C_6$-glucose contributed the most to histone acetylation in human cells.[174] In addition, $^{13}C_6$-glucose was also found to be incorporated to newly synthesized alanine residues through its metabolite, pyruvate, resulting in labeling of newly synthesized histones in the same experiment. The turnover rates of histone H3 and H4 acetylation sites indicated that most half-lives were in about an hour, and acetylation rates are much slower in quiescent fibroblasts than those in proliferating cells.[174] In another study, $^{13}C_6$-glucose was used to measure the steady-state turnover rates of 19 histone acetylation sites with quantitative target mass spectrometry via SRM. Seven long-lived histone acetylation sites with stability over 30 h were successfully identified.[175] These examples show that dynamic flux of histone variants and PTMs can be analyzed by quantitative proteomics and pulse-chased experiments.

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

4.6. Mass Spectrometry versus Western Blot for Detecting Histone PTMs

Two complementary methods, Western blotting (WB) analysis and MS, have been widely used for detection and quantification of histone marks. In the past few decades, WB has been the gold standard for measuring relative changes of protein abundance. WB offers advantages of simplicity, low cost, and convenience. It does not require expensive instruments and can be carried out in any biology lab. When a good antibody is available and an adequate experiment procedure is used, WB can be highly sensitive and specific for analysis of histone marks. Both pan-(sequence-independent) and sequence-specific anti-PTM antibodies have been used for analysis of histone marks.

Despite its wide applications, WB suffers from six shortcomings for analysis of histones and histone marks. (1) WB can detect only one histone mark at a time. In contrast, a histone protein can have dozens of simultaneously occurring modifications. (2) An antibody against a histone mark of interest can be interfered by other modification at its neighboring residues, or epitope occlusion. A good example is H3K9me3 and the neighboring S10 phosphorylation, present in the two adjacent histone residues.[176] In principle, the two amino acid residues can have four possible modification statuses: H3K9/H3S10, H3K9me3/H3S10, H3K9/H3S10ph, and H3K9me3/H3S10ph. An anti-H3K9me3 antibody may be blocked by phosphorylation at H3S10. Thus, three sequence-specific antibodies would be necessary to detect and quantify levels of the last three modification statuses. (3) Some histone marks have very subtle structure differences that require extreme antibody specificity. For example, a lysine residue can be modified by mono-, di-, and trimethylations. Many commercially available antibodies have cross-reactivity and cannot fully distinguish the three modifications at the same residues. (4) Many segments of the core histone proteins sequences are exceptional homologous. Thus, a sequence-specific antibody against one modification may have cross-reactivity to others. For example, histone H3K9S10 and K27S28 have similar sequence motifs (…ARKS…), and the lysine residues can be modified by mono-, di-, or trimethylation, whereas the Ser residue can be phosphorylated. A commercially available antibody for these histone marks has been found to possess cross-reactivity toward each other.[150f] (5) It is most challenging to raise antibodies against a specific histone variant. For instance, histone H3 has five variants in human cells. Only five amino acid residues are different among H3 variants H3.1, H3.2, and H3.3. At the highly modified N-terminal tail, there is only a one amino acid residue difference between H3.1/2 versus H3.3, A31 versus S31.[170] (6) WB can only recognize known PTMs. Thus, the ability to discover novel PTMs, as is often performed with MS, is not possible with WB. These caveats also exist when using antibodies in other immuno-based methods, such as in ChIP experiments.[138i,177]

Given the possible nonspecific binding caused by either antibody or experimental errors, cautions need to be taken to ensure that an antihistone PTM recognizing antibody has high specificity toward its intended target. To this end, a panel of adequate control experiments is necessary, including but not limited to (1) dot-spot assay with histone peptides, with or without the modification of interest, and with structurally similar modifications, (2) WB analysis, with or without competition of an antigen peptide bearing the modification of interest, and (3) WB analysis of a recombinant histone protein that is known to have no

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

histone modification. For example, for a sequence-specific anti-H3K4me3 antibody, its corresponding H3K4me2, H3K4me1, and H3K4ac peptides should be used as negative controls in the dot-spot assay to ensure high specificity of the antibody. Likewise, for an anti-H4K5bu (butyrylation) antibody, its corresponding H4K5ac, H4K5pr (propionylation), H4K5cr (crotonylation), and H4K5hib (2-hydroxyisobutyrylation) peptides should be used to test its specificity. Given the high number of histone marks and their structure similarity, this control experiment is not only important for WB analysis but also especially for ChIP-Seq experiments. Alternatively, immunoprecipitation of histone peptides using an anti-PTM antibody of interest followed by MS characterization of the peptides will be able to identify nonspecifically binding peptides.[138i]

All of these caveats, however, can be addressed by using MS-based approaches. MS can not only detect and quantify relative changes of histone PTMs, but also determine absolute concentration (or stoichiometry), when MRM assays and internal synthetic peptides are used. MS can effectively identify multiple histone PTMs simultaneously, regardless if they are present in adjacent sequence or on the same peptides. Accurate mass information enables one to distinguish among histone marks with similar structures and masses, for example, trimethylation (with mass shift of 42.0470 Da) versus acetylation (with mass shift of 42.0106 Da). Moreover, MS can be used to discover unknown PTMs as described in section 3.[6a,b] Given its quantitative nature and high specificity, MS has been suggested as an alternative method for WB. It has also even been recommended by the proteomic community that there is no need to perform both WB and MS, to deliver solid quantitative information for a protein or a protein modification.[178]

Nevertheless, an MS approach has its own problems. MS analysis is typically carried out in a professional proteomics laboratory, using expensive and sophisticated instruments. Operation of such equipment requires technical staff with high levels of training. MS cannot be easily used to distinguish between two protein modifications with the same mass shift, for example, structure isomers without creative methods. Last, good MS data interpretation requires a high level of experience, much more than WB readout.

Taken together, WB and MS-based techniques are complementary techniques and are of great value for analysis of histone marks.

4.7. Dynamic Analysis of Histone Variants Using Proteomics

Histone variants, especially those of H2A and H3 (Table 2), have different amino acid sequences and unique PTMs patterns, providing additional mechanisms to modulate chromatin structure.[2a,179] Some PTMs on histone variants play critical roles. For example, phosphorylation on H2A.X serine 139 is a marker for double-strand DNA breaks and is important for the repair response process (reviewed in ref 180). Conventional approaches, such as antibody-based assays, are not convenient to analyze histone variants and their PTMs due to their high sequence homology.[181] MS-based technologies offer advantages to identify and quantify histone variants, as well as the PTMs on specific variants, as long as there is a difference of one residue or one PTM among homologous proteins.[182]

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

In general, the methods for analyzing histone proteins and PTMs can also be used to analyze histone variants and their PTMs.[182] In addition, quantification of the histone variant of interest can be achieved by using one or a few variant-specific peptides generated from an enzymatic digestion. Using the SILAC technique, global turnover rates of a number of histones in Hela cells were determined, including canonical histones and H1.4, H2A.Z, H2B variants, and H3.3.[184] The analysis revealed that H1.4 and certain H2A variants, including H2A.Z, had a faster turnover rate than the canonical histones, which is in accord with the observation that histones turn over faster at open chromatin regions.[185]

Although bottom-up MS is powerful to analyze histone variants and their PTMs, it suffers from a couple of shortages. First, some histone variants differ by as little as one amino acid (H3.1 versus H3.2), and they cannot be distinguished if the peptide harboring this amino acid change between the variants is not detected. Second, if a PTM is on a peptide with shared sequence between the variants, it is impossible to distinguish the PTM from each individual variant. These issues can be overcome by using top-down MS, as it can distinguish all variants at the intact protein level, without the need to identify individual peptides. Many examples of top-down MS are reported and have been reviewed recently.[186] For example, Boyne et al. combined SILAC and top-down MS to confirm that most canonical H2A isoforms are replication-dependent.[187] In another study, the modifications on H3 variants were profiled using top-down MS.[186b] In asynchronous Hela cells, 5% of K4 was monomethylated and about 50% of K9 was dimethylated in H3.1. In addition, K14 and K23 were identified as the major acetylation sites.[186b] Nevertheless, a complex PTM pattern in histone variants may complicate this type of analysis.

5. COMPREHENSIVE LIST OF HISTONE PTMs

The dynamic changes of histone marks are closely associated with cellular physiology and diseases, such as cancers, neurodevelopmental disorders, and cardiovascular disease.[188] Thus, mass spectrometry approaches described above have been extensively used to detect and quantify histone marks. Here, we attempt to provide an up-to-date comprehensive list of reported histone marks (Tables 3–11). The primary literature associated with these studies is also listed. Sequences of mouse histone H3.1, H4, H2A type 1, H2B type 1-K, and H1.2 are used as templates to number the modified residues.

Analysis of histone marks described below was typically carried out in a model system of interest. Given the fact that the stoichiometry of histone marks is very different among cell types and cellular environments and that the earlier studies were carried out in mass spectrometers with much lower sensitivity, it is highly likely that many histone marks bearing the known types of PTMs are missing. Recently, identification of tyrosine phosphorylation at H4Y72 offers such an example.[189] In addition, for newly discovered lysine acylation marks, only one or a few cellular systems have been analyzed. Accordingly, additional histone sites bearing these types of modification should exist in other cells.

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

6. IDENTIFYING "READERS" FOR HISTONE PTMs

6.1. Introduction

Several mechanisms have been proposed for histone marks to exert their functions, including altering the physical properties of nucleosomes by neutralization of charge via acetylation, resulting in increasing nucleosome mobility and modulation of the higher order chromatin structure.[24b] In addition, histone marks often act through the recruitment of downstream molecules, referred to as "readers" or "effectors", which specifically recognize a particular modification in the context of the histone molecule or nucleosomes.[1c,291] Here, a few terms for the relevant proteins are defined. We would like to define a protein "reader" of a histone marks as a protein that directly interacts and recognizes a specific histone mark in a particular sequence context; a protein "binder" as a protein that either directly or indirectly associates with a histone, in a modification-dependent or -independent manner; and a "effector" as a protein that binds specifically to a post-translational modified histone substrate, and this binding event recruits other activities contained within the same polypeptide or complex. Thus, "effectors" translate histone marks into biological output. A "reader" is a direct "binder" of a histone mark.

Association of the downstream "effectors" then leads to changes in accessibility of the DNA template to the transcriptional machinery, recruitment of enzymatic activities, for example, ATP-dependent chromatin remodeling complexes, or changes in the higher order structure of chromatin, which, in turn, dictate specific regulatory outcomes. Recent studies demonstrated that specific histone PTM-"reader" molecule interactions play a role in the regulation of such diverse cellular processes as transcription, gene silencing, X-chromosome inactivation, DNA damage repair, V(D)J recombination, and maintenance of gene expression programs during development.[4e,291,292] In addition, recently PTM "reader" proteins have been used as a new class of proteins targeted for cancer therapeutics.[4e,293]

"Readers" often contain a structural domain that recognizes a histone PTM. The identified protein domains for histone marks include bromodomain for acetyllysine;[9a] chromodomain, Tudor domain, MBT-repeats, WD40-repeats, PHD finger, PWWP domain, ADD domain, zinc finger CW domain, BAH domain, and CHD domain for methyllysine;[8,291b] and BRCT domain and 14-3-3 domain for phosphorylated serine.[294] Adding to the complexity of epigenetic regulation, lysine residues can be mono-, di-, or trimethylated at the ε-amine in vivo. Underscoring the precision of translating epigenetic signals, recognition of histone PTMs by "reader" molecules is not only site specific, but can also be methyl-state specific. For example, PHD fingers of ING2 and BPTF can distinguish between different methyl states with high preference for binding to the H3K4me3.[8a,c]

Regulation of gene expression is intrinsically combinatorial, and recognition of histone PTMs can be seen as an additional variable in a combinatorial code that remains poorly understood. Recent epigenomic-level and mass spectrometry analyses strongly imply the coexistence of certain modifications such as H3K4me3 and H3K9/14/18/23ac at active genes, or H3K4me3 and H3K27me3 at the so-called "bivalent domains" associated with developmental genes in embryonic stem cells.[83b,295] Interestingly, histone "effector" complexes often contain multiple known or potential histone modification recognition

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

modules either within the same polypeptide or in different protein in the same complex.[7] For example, the human genome encodes 22 distinct proteins containing both PHD finger and a bromodomain, often adjacent to each other. On the basis of structural and biological evidence, we have recently proposed that such a dual PHD-Bromo module may function as combinatorial recognition motifs to nucleosomes containing both H3 methyl and H4 acetyl modifications.[8a,296] The extent to which multiple binding modules bind cooperatively to multiple covalent marks either on a single histone tail or on distinct histone tails remains an open question.

Identification of histone PTM "binders" is a challenge for several reasons. First, the binding constants measured for chromatin "effectors" are rather weak, in the mid to high micromolar range and roughly commensurate to those observed with phospho-dependent interacting partners in signal transduction cascades.[297] Second, histone peptides are typically very basic and as a result can bind cellular proteins containing an acidic domain, which increases the level of nonspecific binding. Third, multiple proteins bind histone tails specifically, albeit independently of the modification of interest, which additionally complicates the downstream analysis. Finally, histone PTM associated proteins (or protein complexes) may synergistically associate with multiple modifications. Thus, design of histone peptides bearing more than one PTM and recovery of intact protein complexes may prove critical for identifying such proteins.

6.2. General Strategy for Identifying "Readers" for Histone Marks

In the past several years, MS in combination with affinity purification has served as a powerful approach to characterize the interactions between histone marks and their "readers". In principle, a peptide, a protein, or a reconstituted nucleosome containing one or a few histone marks of interest can be used as bait to pull down its direct "binder" protein as well as its associated proteins in a protein complex. The pull-down experiment will easily lead to dozens of proteins that can be isolated specifically with the modified histone peptide of interest. Accordingly, additional biochemical experiment is necessary to pinpoint the direct "binders".

The current proteomic method for identifying direct "binders" and their associated proteins for histone marks typically involves five steps (Figure 16): (1) design and synthesis of biotinylated histone tail peptides bearing one or more PTMs of interest; (2) isolation of proteins associated with histone peptides in different cell types via optimized affinity purification method, followed by in-solution or in-gel tryptic digestion; (3) identification of associated proteins by nano-HPLC/MS/MS; (4) quantification of the identified proteins across the histone peptides, modified versus unmodified, to pinpoint the proteins specific to a histone mark of interest; and (5) identification of direct interactors via probing histone peptide(s) with recombinant candidate "binders".

Histone PTMs "readers" can recognize not only histone PTMs, but some of them can also recognize nonhistone PTMs. For example, the bromodomain of CBP is able to recognize acetylated nonhistone protein p53.[298] The principles of "reader" identification for both histone and nonhistone PTMs are the same. Therefore, the general strategy for identifying "readers" can also be applied to identify "readers" of nonhistone PTMs.

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

6.3. Identification of Binding Proteins Using Biotinylated Peptides Containing Histone PTM(s)

In this experiment, a biotinylated histone peptide is first chemically synthesized containing one or a few modified histone residues. When incubated with a protein extract, such as protein whole-cell lysate or nuclear extract, the peptide is used as a bait to isolate its "binder(s)". As a control, its corresponding unmodified peptide is used in a parallel experiment. The differentially binding proteins between the two parallel experiments suggest "binder" protein candidates.

The early studies spotted the differential proteins by visualizing the isolated proteins in SDS-PAGE. The protein(s) specific to the modified histone peptide were detected and further identified by MS. Using this approach, in 2006, Wysocka, Allis, and their colleagues identified a PHD finger BPTE protein as a direct "binder" protein to H3K4me3 and demonstrated that it is associated with chromatin remodeling.[8a] This study also revealed that PHD finger is a highly specialized methyl-lysine-binding domain. The same approach was used for a comprehensive, unbiased screen for proteins associated with H3K4me2, H3K9me2, and H3K9ac.[299] After pull-down and SDS-PAGE experiments, all of the proteins in each lane were identified and semiquantified by spectral counting method.[108a] The proteins specifically bound to modified histone H3 peptides then were distinguished. This study identified 86 proteins that directly or indirectly bind to the amino terminus of histone H3 containing one of the three histone marks. Many of them are known modification specific "binders", providing a good positive control. In addition, proteins containing the well-known PHD finger, WD40 repeats, and bromodomain were also identified as "binder" candidates.

SILAC and MS were also used in combination with pull-down experiments with histone peptides to identify "binders" for histone marks. This approach is getting more popular because it gives quantitative results that distinguish specific "binders" from nonspecific ones. In this experiment, the modified and unmodified histone peptides are used as baits and are incubated with "heavy" and "light" SILAC labeled protein extracts, respectively, to carry out the pull-down experiment. This is called "forward experiment". The isolated proteins from the two pull-down experiments are combined and subjected to protein identification and quantification by MS. In a parallel experiment, these bait peptides are incubated with inversely labeled nuclear extracts, which is called the "reverse experiment". Specific binding proteins to the modified histone peptides of interest are enriched in pull-down samples with modified peptides (versus its unmodified counterpart).[300] All of the identified proteins and their logarithmized abundance ratios from the two independent experiments are plotted by their SILAC ratios in the forward (*x* axis) and reverse (*y* axis) experiments. Background proteins are close to the grid origin, while the specific "binders" are far from the grid origin, showing significant ratios between "heavy" and "light" forms (Figure 16, right). By using this approach, TFIID was identified as a "reader" for H3K4me3.[301] In addition, an unbiased interaction screen was carried out for five major lysine trimethylation sites on histone H3 and H4, including H3K4me3, H3K9me3, H3K27me3, H3K36me3, and H4K20me3.[300] The results indicated that a double Tudor-

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

domain in the C-terminus of Sgf29 binds to H3K4me3, and PWWP domain is a putative H3K36me3 binding motif.

Although the SILAC in combination with the pull-down experiment is powerful to identify “binders”, this approach is difficult to be applied to identify “binder” proteins from animal tissues and used for a comparison across multiple pull-down experiments. To address this limitation, a label-free approach was used for quantifying protein abundance, across histone peptides containing different modification types or modification sites. In this experiment, the modified and unmodified histonetail peptides are incubated with protein extracts, respectively. The isolated proteins are digested and subjected to HPLC/MS/MS analysis; all of the identified proteins are quantified by a label-free quantification algorithm.[302] Each group of pull-down experiment (with a specific histone peptide, modified or unmodified) is usually carried out in multiple replicates (3–5 replicates) to improve the reproducibility. Typical statistic analysis will involve three steps to identify “histone mark”-specific binding proteins.[302] First, the input proteins are identified that can be detected among all of the replicates in at least one experimental group (either unmodified or modified group). Other proteins that do not meet the criteria will be removed. A fraction of the input proteins may not be detected among all of the groups. They will be given empty intensity values in those groups that do not detect the proteins. All of these empty intensity values will be imputed with random numbers at noise level. Second, all intensities will be logarithmized, and a t-test will be performed with a low FDR cutoff, for example, 0.01 or lower. The significantly enriched proteins among different groups can be distinguished. Finally, for each significantly enriched protein, its intensities in each replicate of all groups will be normalized by subtracting their average derived from all of the replicates. A heatmap of these normalized intensities can indicate enrichment of the proteins among different groups. Following this procedure, specific binding proteins to histone marks of interest can be identified on the basis of the quantified proteins across multiple pull-down experiments (Figure 16A). This approach was used to screen tissue-specific chromatin “binders” specific for H3K4me3 and H3K9me3.[302]

In addition to the histone peptide baits containing one histone mark, the peptides bearing more than one histone marks were also used for identifying protein “binders”. This type of experiments has a chance to identify direct “binders” that have low binding affinity when associated with one histone mark, but have greatly enhanced interaction in the context of multiple histone marks. Indeed, three such histone tail peptides were used as baits to identify their direct “binders”: a bis-acetylated H4K5ac/H4K12ac peptide, a triple-acetylated H4K8ac/H4K12ac/H4K16ac peptide, and a tetra-acetylated H4K5ac/H4K8ac/H4K12ac/H4K16ac peptide.[303] As compared to the specific “binders” of the mono- and bis-acetylated baits, more specific “binders” were isolated on the triple- and tetra-acetylated bait peptides. It seems that the synergistic action of multiple acetylation of H4 may contribute to the association of some “binders”.

6.4. Identification of Binding Proteins to Nucleosome-Associated Histone PTMs

The biotinylated peptide bait offers advantages of easy synthesis and low cost. Typical histone tail bait contains ~20 N-terminal amino acids, which cannot mimic the whole

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

histone tail. In addition, the linear peptide structure is very different from that of the whole histone proteins and that when the proteins are associated with nucleosomes. Accordingly, the pull-down experiment using histone peptides will be difficult to uncover those binding proteins, whose interaction with histone marks not only requires the histone mark(s) (either in the same histone protein or in different histone proteins), but also conformation of a nucleosome structure. To address this, the reconstituted nucleosomes containing DNA and recombinant histones, with or without histone marks of interest, are prepared and used as baits for pull-down experiment.[304] These modified nucleosomes are closer to true microenvironment than modified peptides, which enable one to reveal the "crosstalk" of histone marks and nucleosome-dependent binding proteins.

To identify both nucleosome- and histone mark-dependent "binders", the two types of nucleosome baits, with or without histone modification(s), are incubated with protein extract from SILAC-labeled cells. The isolated proteins are quantified to pinpoint the modification specific "binders". Using this approach, it was revealed that DNA and histone methylation cooperatively recruit the origin recognition complex (ORC), while DNA methylation disrupts the binding of Fbxl11/KDM2A to histone methylation.[304]

6.5. Validation of Interaction between a Direct "Binder" Protein and a Histone Mark

Recovery of protein complexes will provide important insights into the nature of histone modification recognition. However, this approach will not address which of the identified proteins directly interacts with modified histone tails. Accordingly, additional biochemical experiments are necessary to identify direct interaction (binding) proteins, or direct "binders", for the histone marks of interest, using either peptide or nucleosome containing histone mark(s) of interest. To this end, binding assay will be carried out using recombinantly expressed candidate proteins and modified peptides. Their unmodified peptides will be used as a control. Further, the direct "binder" protein candidates can be further validated by a conventional pull-down experiment, with protein lysate in cells that expresses the protein of interest, either in *E. coli* or in mammalian cells.

Once a direct "binder" is conclusively established for a histone mark of interest, several subsequent experiments can be pursued to characterize the interaction in more detail. First, recombinant deletion mutants of the direct "binder" protein are produced and used to test their interaction with the histone PTM peptide of interest. This study can be used to define a protein domain that is responsible for the interaction. Once a binding domain is defined biochemically, biophysical experiments can be used to confirm the interaction, including florescence polarization-based binding assay, surface plasmon resonance, and isothermal titration calorimetry-based binding assay.[305] To understand the molecular recognition between a histone mark and its binding domain (or protein), NMR and X-ray spectroscopy is used to reveal key insights into the mechanism by which the histone mark is recognized at the molecular level. Subsequent mutagenesis experiment can be used to confirm the molecular interactions. Additionally, ChIP in combination with DNA sequencing, with antibodies against both histone mark and its direct "binder", can be used to verify in vivo binding of histone binding protein for specific histone PTMs.[305a,306]

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

7. GLOBAL ANALYSIS OF NONHISTONE PTMs

7.1. Introduction to Experimental Strategy

It is highly likely that a PTM-regulatory enzyme, for example, the acetyltransferase p300/CBP, exerts its functions not only on histones but also on other proteins.[307] Thus, we can potentially expect known histone PTMs to be present in nonhistone proteins. Information on PTM substrates and their PTM sites is required to study the role of PTM in the function of these substrate proteins, and to reveal insights into the possible regulation of cellular physiology by the particular PTM. Protein lysine acetylation shows a good example in this regard. Lysine acetylation was initially identified in histones in the 1960s.[116] In 1997, identification of its first nonhistone substrate protein, p53, stimulated extensive studies of the roles of lysine acetylation in transcriptional regulation.[308] Identification and functional characterization of diverse substrates in both cytosolic and mitochondrial proteins promoted a breakthrough in our understanding of this modification in cellular metabolism and signaling.[129b]

Traditionally, biochemical approaches, such as in vitro PTM reaction assays using radioactive isotope-labeled substrates and Western blotting analysis, were preeminent techniques for the detection of the protein substrates containing a PTM of interest.[309] However, the radioactive isotopes are involved in health and environmental risks, and the protection and radioactive waste disposal are complex. Some radioactive isotopes (e.g., ^{32}P and ^{33}P) with relatively short half-lives necessitate that the material be freshly labeled for optimal efficiency.[310] In addition, ^{14}C or ^{3}H used for confirming protein methylation and acetylation has relative weak radio emitters, which makes it difficult to detect smaller quantities of proteins or the proteins only with a few PTMs.[311] Thus, it is difficult to use these isotopes to efficiently detect their corresponding modified proteins. Antibodies against a PTM of interest, either sequence-independent or sequence-dependent, are very valuable as validating tools, but antibodies can be costly and do not allow for direct screening of PTM substrates with high efficiency.

For well over a decade, MS-based proteomic approaches have emerged as the method of choice for system-wide analysis of PTM substrates and mapping of PTM sites.[311b] Two basic conditions should be met for these methods to be applied efficiently. First, a suitable enrichment method is available to separate the modified peptides of interest from the complex proteolytic digest including unmodified and other PTM-containing peptides. Only in such a way can the modified peptides be efficiently analyzed by nano-HPLC/MS/MS methods. Second, the peptide containing a PTM of interest is permissive for some level of fragmentation in peptide bonds, so that the resulting daughter fragment ions can be used for protein sequence database searching for identifying peptides and mapping PTM sites. Some PTMs, such as O-GlcNAc modification and ADP ribosylation, have very fragile chemical bonds either at the modification chemical moiety or at the modification linkage bond. The major fragmentation channels in MS/MS analysis for the peptides containing these types of PTMs happen at the chemical bond that links sugar to the modified amino acid residues or ADP-ribose moiety, respectively. To overcome the issues, MS/MS/MS and ETD approaches were used to generate daughter ions with fragmentation at the peptide bonds.[42c,312]

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

Alternatively, a chemical derivatization method can be used to label the modification residue of interest and facilitate efficient fragmentation.[42c,313] This approach, involving affinity enrichment of PTM peptides and HPLC/MS/MS analysis, has been extensively used for global detection and dynamic analysis of protein phosphorylation, lysine acetylation, lysine and arginine methylation, and protein ubiquitination.

The proteomic analysis of PTMs involves three steps (Figure 14): (1) the protein lysate of interest is prepared and digested with a proteolytic enzyme, typically trypsin; (2) the PTM peptides of interest are enriched from the proteolytic peptides using a suitable method; and (3) the isolated peptides are analyzed by nano-HPLC/MS/MS; and the resulting MS/MS data are searched against protein sequence database for identifying peptides and mapping PTM sites.

Four key factors are discussed here to improve sensitivity and accuracy for proteomics of PTMs: sample complexity, enrichment methods, sensitivity of HPLC/MS/MS system, and accuracy for identifying PTM peptides and mapping PTM sites.

7.2. Sample Complexity

There could be about 10 000 proteins in a protein lysate of interest that can have a difference of expression level up to 3–4 orders. Thus, a PTM proteomic experiment typically has bias toward the abundant proteins. In this type of experiment, a protein sample is first prepared and then digested with a proteolytic enzyme, such as trypsin. The resulting digest is subjected to enrichment of PTM peptides, followed by HPLC/MS/MS analysis for identification and quantification of PTM peptides. This strategy may suffer from high complexity of the peptide mixture, leading to compromised detection sensitivity.

To address this issue, a separation step can be included in the proteomic strategy to reduce the sample complexity. Three methods have been described in the past to separate proteins or proteolytic peptides before enrichment of PTM peptides, including: (1) The first is isolation of proteins from the cellular organelles of interest. Cytosolic and nuclear protein extracts have been widely made,[314] especially in the field of transcriptional regulation, for purifying proteins of interest. For protein extracts in other organelles, the organelle of interest is first isolated, typically by a centrifugation method. With the increased interest in organelle biology, for example, mitochondrion and ER, focused analysis for a specific organelle can improve sensitivity to identify PTMs. (2) Next is separation of either proteins (before the proteolytic digestion) or peptides (after the proteolytic digestion) into multiple fractions before affinity enrichment. Because it is difficult to reconstitute denatured proteins (which tend to precipitate unless a high concentration of SDS is used), a nondenaturing detergent, such as NP-40, is used for preparing protein lysates that are further resolved in ionic exchange column. Because some proteins cannot be retained in either a cation or an anion ion exchange column, a mixed-bed ion change column can address this issue with improved resolution that contains both cation and anion ion exchange beads.[129a,315] Each fraction can then be treated in a similar fashion as a protein lysate for the downstream experiments. (3) The final method is separation of proteolytic peptides using basic HPLC.[123c,316] High-pH reverse-phase HPLC can dramatically change the charge distribution within the peptide chain. Therefore, it is semiorthogonal to subsequent low-pH

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

reverse-phase HPLC coupled to MS via electrospray ionization, which greatly enhances the chromatographic resolving power.

All three methods have been used for proteomic analysis of protein modifications.[123c,129a,b] The fractionation experiment can not only reduce the sample complexity, but can also improve the yield for enrichment of peptides, for example, in the case of antibody affinity purification. The experimental design can vary depending on the goals of the experiment. Isolation of cellular organelles is an obvious choice if the focus is on a specific cellular compartment. Nevertheless, given the higher complexity of cytosolic and nuclear proteins than other organelles, they are usually further resolved into multiple fractions during proteomic studies. Separation of the non-denatured proteins is more complicated than that of peptides. With the more robust and improved separation with basic pH buffer, HPLC-based peptide separation is getting more popular and has been shown to deliver high sensitive proteomic data.[123c]

7.3. Enrichment Methods

The modified peptides of interest are present in a pool of other peptides. Accordingly, they need to be enriched for efficient HPLC/MS/MS analysis. In the past two decades, a variety of methods have been developed to enrich PTM peptides, depending on the chemical properties of PTM moieties: antibody-based affinity purification, noncovalent interaction, and chemical derivatization.

7.3.1. Antibody-Based Affinity Enrichment—Antibodies are valuable reagents for detecting, quantifying, and enriching PTMs. Pan, or sequence-independent, antibodies can recognize the modification, but not its surrounding sequence. They can be used to evaluate overall changes of the modification in a protein of interest. To detect a specific PTM site, a sequence-specific antibody is used, which can recognize not only the PTM moiety but also its surrounding sequence. Generation of pan- and sequence-specific antibodies should be similar to those for protein antibodies.[6b,317] However, methods for antigen design and antibody purification are different. In addition, the PTM antibody of interest should be carefully examined. A few strategies have been used for designing antigen for pan anti-PTM antibodies, including modified bovine serum albumin (BSA), ovalbumin (OVA) (e.g., acetylated BSA or OVA for generating antiacetyllysine antibodies), and randomized peptide libraries. The serum from immunized animal contains diverse antibodies and therefore should be subjected to immunoaffinity purification. The subset of antibodies against the PTM of interest in principle can be enriched by the antigen (e.g., acetylated lysine or acetylated BSA/OVA) and depleted by its unmodified counterpart (e.g., unmodified lysine or unmodified BSA/OVA). Likewise, the subset of antibody against a sequence-specific PTM can be enriched by the modified peptide and depleted by its unmodified counterpart. The resulting antibody should be carefully evaluated so that they will not cross react with other structure-similar PTMs. For example, when we made three structure-similar antibodies, anti-Ksucc, Kmal, and Kglu antibodies, we used dot-spot assay to examine the cross reactivity among several sets of peptide libraries containing a fixed unmodified lysine, malonylated lysine, succinylated lysine, glutarylated lysine, and acetylated lysine.[128b] Likewise, we did a similar characterization for anti-2-hydroxyisobutyryllysine antibody.[6a]

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

When a pan antibody against a PTM of interest is available, immunoprecipitation is usually the first candidate method for enriching PTM peptides. Ideally, the pan (sequence-independent) anti-PTM antibody has high specificity and affinity. The antibody can be linked to solid phase matrixes, such as agarose beads or magnetic beads, covalently or noncovalently (through protein-A or protein G conjugated beads). Both types of beads were used for immunoisolation of PTM peptides. This approach has been used for enriching peptides containing acetyllysine,[129a–c,318] phosphotyrosine,[319] malonyllysine,[150a] succinyllysine,[123c,268,320] glutaryllysine,[128b] methyllysine,[321] and glyglyllysine (for protein ubiquitination).[107,273,322] In the case of ubiquitinated proteins, tryptic digestion leads to generation of glyglylysyl peptides, which can be subsequently isolated by using antiglyglylysine antibodies.[323] Using antibody-based affinity enrichment and mass spectrometry, thousands of peptides containing these PTMs can be identified.

The quality of antibodies is critical for efficient isolation of PTM peptides. The pan anti-PTM antibody is targeted to the PTM chemical moiety, which, in many cases, is small in terms of its chemical structure. Thus, it could be challenging to develop this type of antibodies with high affinity. Indeed, antibodies against antiacetyllysine and phosphotyrosine typically have much lower affinity than their sequence-specific anti-PTM antibodies that have much bigger binding pockets. Accordingly, affinity enrichment of PTM peptides with the pan anti-PTM antibody proves to be a challenging experiment.[324] Caution needs to be taken to boost enrichment efficiency.

Both polyclonal and monoclonal anti-PTM antibodies have been used to isolate PTM peptides. The monoclonal antibody offers an advantage of high reproducibility among experiments. However, because only one sequence (or binding pocket) exists, monoclonal antibody typically has lower binding affinity and could be biased toward some unique peptide sequence, therefore compromising its recognition to the PTM peptides with diverse surrounding sequence. In contrast, polyclonal antibodies may be a mixture of several or even dozens of antibodies, each of which may have a unique primary sequence. They can likely recognize PTM peptides with much more broad surrounding sequences. Therefore, when the polyclonal antibodies are generated with high quality, they can enrich more PTM peptides that can be subsequently identified and quantified by MS. The most sensitive PTM proteomic studies from the authors' laboratory are those with polyclonal anti-PTM antibodies, including those for lysine acetylation, succinylation, and malonylation.[123c,129a,150a,268,320] For example, using antiacetyllysine polyclonal antibodies, about 5000 acetyllysine sites were identified, in MEF cells with or without expression of SIRT1, representing the most sensitivity proteomic analysis of lysine acetylation in a single experiment.[129a]

Nevertheless, polyclonal antibodies can potentially have other problems, for example, batch-to-batch variation and limited supply from an animal (e.g., rabbit) that can produce high-quality antibodies. Thus, in an ideal situation, a mixture of monoclonal and polyclonal antibodies should be used for the pull-down experiment, in which high reproducibility and enrichment of diverse PTM peptides can be achieved.

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

7.3.2. Affinity Enrichment Based on Noncovalent Interaction—It may not be an easy task to generate antibodies against some PTMs because of limited structure changes from their corresponding unmodified residues, such as phosphoserine, phosphothreonine, and methyllysine residues. Because of this, alternative affinity enrichment strategies were developed. One example is based on a noncovalent interaction between PTM moieties and a suitable solid matrix.

Peptides containing phosphoserine and phosphothreonine can be enriched by immobilized metal affinity chromatography (IMAC), by taking advantages of specific coordinate bonding between the immobilized metal ions (e.g., Fe^{3+}, Ga^{3+}, Zr^{4+}, and Al^{3+}) and a negatively charged phosphate group.[325] In addition, metal oxide affinity chromatography (MOAC) has also been widely used for specifically enriching phosphopeptides with metal oxide materials, such as titanium dioxide (TiO_2), zirconium dioxide (ZrO_2), and gallium(III) oxide (Ga_2O_3)[326]. Currently, TiO_2 represents the most common MOAC affinity medium, and various acids (e.g., 2,5-dihydroxybenzoic acid, glycolic acid, and lactic acid) were used to improve selectivity and capacity of TiO_2 toward phosphorylated peptides by competing with the binding of acidic background peptides.[327] Both IMAC and MOAC methods have been applied to efficient proteomic studies, leading to the identification and quantification of tens of thousands of phosphopeptides.[157,328] Integration of either IMAC or MOAC with orthogonal fractionation techniques, such as basic HPLC separation,[316,329] hydrophilic interaction chromatography (HILIC),[330] electrostatic repulsion–hydrophilic interaction chromatography (ERLIC),[331] strong cation exchange (SCX),[332] and peptide isoelectric focusing (IEF),[333] can further improve detection sensitivity of phosphoproteome analysis. Similarly, sequential elution from IMAC (SIMAC) strategy can sequentially separate mono- and poly phosphorylated peptides, thus reducing sample complexity and leading to better sensitivity.[334] In addition, Ca^{2+} and Ba^{2+} were used to enrich phosphopeptides by calcium- or barium-phosphate precipitation.[335] Mamone et al. further developed this concept. They used hydroxyapatite (HAP) chromatography to enrich phosphopeptides because multiphosphorylated peptides have higher affinity to the Ca^{2+}-containing HAP surfaces.[336]

Another type of affinity enrichment is using immobilized domains that can specifically recognize certain PTMs. For example, triple malignant brain tumor domains of L3MBTL1 (3xMBT) was used to enrich mono- and dimethylated lysine-containing peptides or proteins because it can specifically recognize such methylated lysine with minimal sequence specificity.[321c] Another example is tandem-repeated ubiquitin-binding entities (TUBEs). Single UBA domain can recognize lysine ubiquitination.[337] TUBEs were designed to enrich poly ubiquitinated proteins by using four tandem ubiquitin-associated (UBA) domains, which have markedly higher affinity than single UBA domain.[338] Lectins can bind specifically carbohydrate molecules. Therefore, lectin has been used to enrich glycopeptides or glycoproteins.[339] Macro domain, a protein module that can specifically recognize ADP-ribose, was used to enrich ADP-ribosylated proteins.[340]

7.3.3. Enrichment of PTM Peptides Based on Chemical Derivatization— Chemical derivatization has also been used to introduce an affinity handle, for example, biotin, for subsequent affinity enrichment. Because of its small size and bio-orthogonal nature, azide has been used for metabolic labeling of PTMs, such as sugars (e.g., O-GlcNAc

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

modification) and lipid-modified (e.g., farnesylation, myristoylation, myristoylation, and succinylation) residues.[341] The metabolically labeled PTM peptides can be easily conjugated to biotinylated molecules by Click chemistry,[342] enabling subsequent enrichment-based biotin–streptavidin interaction.[343] In the case of O-GlcNAc modification, a chemio-enzymatic method has been used to selectively conjugate O-GlcNAc moiety with a ketone containing galactose analogue, which enables a subsequent conjugation with an aminooxy biotin for affinity enrichment.[344]

Direct chemical derivatization, without metabolic labeling with an adizo precursor in cultured cells, has also been used for tagging a PTM of interest. Chemical derivatization was used to introduce an affinity tag for glycopeptides containing a complex carbohydrate, followed by a chemical reaction between the carbohydrate and hydrazide beads. The resulting beads-bound glycopeptides can then be released by enzymatic treatment with PNGase F.[345] In addition, *β*-elimination followed by Michael addition (BEMAD) with dithiothreitol (DTT) or biotin pentylamine (BAP) can replace the serine- or threonine-linked O-GlcNAc with a DTT or biotin tag for subsequent enrichment.[346] For a long period of time, poly ADP-ribosylation has been a challenge to analyze by mass spectrometry due to its labile nature. A method was recently developed to isolate Asp- and Glu-ADP-ribosylated peptides with boronate affinity enrichment, followed by treating the peptides with NH_2OH, generating a molecular signature, hydroxamic acid derivative with an addition of 15.0109 Da.[42c,313] Accordingly, the derivatization reaction not only removes the fragile ADP-ribosyl group but also marks the residue of ADP-ribosylation.

Boronate affinity enrichment is also used to capture proteins or peptides with O-GlcNAc modification. Boronic acids can covalently react with the *cis*-diol moiety of O-GlcNAc and form five- or six-membered borate esters. After enrichment, O-GlcNAc proteins or peptides can be released in acidic solutions and subjected to HPLC/MS/MS analysis.[347]

Another class of chemical reaction-based enrichment approach is often applied to the analysis of cystine modifications, such as *S*-nitrosylation, *S*-glutathionylation, and disulfide formation. In some cases, free cysteine residues are first blocked, and then oxidized cysteines were selectively reduced, generating a possibility to introduce an affinity tag (e.g., biotin) by oxidation (e.g., disulfidation) or nucleophilic (e.g., alkylation with 2-iodoacetamide derivatives) reactions.[348] Thus, the tagged peptides, which contain the information on oxidized cysteine, can be identified.

It has been challenging to develop good antibodies against lysine monomethylation due to its small size and little difference from its unmodified residue. Recently, Wu et al. developed a chemical proteomic method to address this technical challenge. This method involves derivatization of the monomethyl *ε*-amine group of the lysine residue by adding a propionyl group. The propionyl monomethylated lysine is much bigger than monomethyllysine, allowing generation of a pan antipropionyl monomethyllysine antibody with binding affinity. Affinity enrichment of propionyl monomethyllysine followed by HPLC/MS/MS analysis identified 448 mono-methylation sites on 401 proteins with high accuracy, the largest monomethyllysine data set ever reported.[321a]

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

7.4. Sensitivity of HPLC/MS/MS System

High sensitivity of HPLC/MS/MS analysis is another key factor for the overall sensitivity of PTM proteomics. Some PTM peptides may be isolated in very low amount from proteolytic digest, especially those from low-abundance substrate proteins. Accordingly, the HPLC/MS/MS system should be optimized to maximize its sensitivity.

In an ideal situation, HPLC should be tuned in such a way that the peptide peaks, either from early or from late eluted peptides, should have the highest possible concentration. The mass spectrometer's sensitivity is largely dependent on the concentrations of its analytes. Several parameters are critical to obtain narrow peak width and elevate concentration of analytes from HPLC column: high performance packing material, narrow-ID and long HPLC column, small particle size, and appropriate core size. The most popular packing material for peptides separation is C12 or C18 porous particles. Some particles, for example, core–shell particles, are created by new technologies and have less band broadening. This reduction in band broadening results in chromatographic separations with better resolution, higher sensitivity, and improved peak capacities.[349] The narrow HPLC column enables one to elute peptides in small solvent volume and thus high concentration. Typically, columns with 50–100 um inner diameters, packaged with either C12 or C18 beads, are currently used for HPLC/MS/MS analysis. Usually smaller diameter particles are advantageous to both efficiency and resolution, but may yield higher high backpressure. The pore size is also important because the molecules must fit into the porous structure to interact with the stationary phase. Smaller pore size (80–120 Å) is the best choice for small molecules with molecular weights up to 2000 Da. In addition to the column, a low flow rate is also critical to achieve high ionization efficiency and MS sensitivity. Several nanoflow HPLC pumps are commercially available that deliver solvents as low as 50–200 nL per minute, which include EASY-nLC 1000 from Thermo Fisher Scientific Inc., NanoLC-1D Plus and 2D Plus Systems from Eksigent Technologies, LLC., and nanoACQUITY UPLC from Waters Corp. The new models of mass spectrometer, for example, Q-Exactive Plus Hybrid Quadrupole-Orbitrap MS (Thermo Fisher Scientific Inc.) and TripleTOF 6600 (AB Sciex Pte. Ltd.), have acquisition speeds up to 18–100 Hz. A 30 min data acquisition time should enable generating, in principle, tens of thousands of MS/MS spectra. Almost any peptide with a peak width within several seconds should be detected. In a typical analysis, fewer than 5000 PTM-containing peptides are isolated in a protein digest (or peptide fraction). Accordingly, a long gradient, for example, 2 h or longer, is likely to increase the peak width of a PTM peptide and therefore compromise detection sensitivity.

In addition to HPLC conditions, the mass spectrometer should also be optimized in such a way that maximized sensitivity can be achieved. To this end, standard calibration molecules from MS venders (e.g., caffeine, peptide Met-Arg-Phe-Ala, ultramark 1621, and *n*-butylamine) are typically used to test the instrument system so that maximized sensitivity for both MS and MS/MS analysis will be achieved.

7.5. Accuracy for Identifying PTM Peptides and Mapping PTM Sites

Accuracy of PTM data sets from proteomic screening is critical to the research community. The PTM data sets are likely used not only by proteomic researchers but also by the

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

biomedical community. The misidentification of a PTM site may cause months of nonproductive efforts for a biologist who attempts to pursue its biological functions. Despite improved mass accuracy of mass spectrometers and tremendous efforts to improve algorithms for protein sequence alignment of MS/MS data, the false identification of PTM peptides and incorrect mapping of sites likely remains an issue for discovery efforts in the close future.

The gold standard for verifying identification of a PTM peptide is to compare the MS/MS spectra of the cell-derived peptides and its synthetic counterpart. We consider this as a very valuable approach when it is determined to take significant effort to investigate the biological functions of this modification event. This is important when the quality of a MS/MS spectrum is not ideal. Given the low cost of peptide synthesis, this experiment seems to be a logical step for a biologist, if the researcher is not sure about the quality of peptide identification. However, this process is slow and expensive, when applied into a large data set. Therefore, the proteomics researchers, who carry out the experiment and deposit the PTM data sets to the relevant database, should take the responsibility to ensure the accuracy of PTM data sets. To evaluate the quality of the PTM data sets, we would suggest that authors manually check the quality of the MS/MS spectra for a randomly selected set of peptides (e.g., 10–20 peptides). Such manual verification of peptide identification has previously been described,[350] which should provide some guidance for the manual inspection experiments. It is realized that the fragmentation patterns are different among mass spectrometers and mass spectrometric methods (e.g., collision induced-fragmentation methods vs high-energy collision fragmentation, or ion trap mass spectrometer vs quadrupole time-of-flight mass spectrometers). To address this problem, a training data set of MS/MS spectra derived from known proteins, for example, standard proteins, will be useful.

7.6. Quantitative Proteomics for Identifying Substrates of PTM-Regulatory Enzymes

The PTM enzymes, such as kinases and phosphatases, are frequently important knots of complex cellular signaling networks. The enzymes typically exert their functions through their PTM substrates. Accordingly, identification of substrates and their PTM sites for the PTM enzymes can not only reveal their functions, but also map complex cellular regulatory mechanism. Traditionally, PTM substrates were identified one at a time, by a candidate approach. For example, many PTM enzymes can interact with their own substrates. Accordingly, identification of enzymes' interaction partners followed by an assay, either in vitro or in vivo, represents a practical approach for identifying PTM substrates. This approach has been used for detecting substrates for kinases, phosphatases, and ubiquitination enzymes. While useful, this classic approach is laborious.

Quantitative proteomics of PTM substrates makes it possible to identify PTM substrates with high efficiency. This approach typically involves three steps (Figure 15): (1) Cells of interest, with or without an expression of a PTM-regulatory enzyme, are grown with SILAC media for more than 6 doubling, respectively. Equal numbers of cells are combined; the resulting protein extracts are made and subjected to tryptic digestion. (2) Next is affinity enrichment of PTM peptides. This can be carried out with an adequate enrichment method

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

as described above. (3) The enriched PTM peptides are analyzed by HPLC/MS/MS followed by protein sequence database searching for identifying PTM peptides, mapping PTM sites, and quantifying their changes between two pools of samples. The detected PTM level is dependent on not only PTM stoichiometry but also protein expression level. Accordingly, protein expression levels are also quantified to normalize the changes of PTM levels, giving the PTM changes caused by dynamic change of stoichiometry instead of protein expression. The differentially changed PTM peptides are substrate candidates for a PTM enzyme.

For example, protein expression, lysine succinylation, and lysine acetylation were quantified in MEF cells, wide-type, and SIRT5 knockout cells, by SILAC and mass spectrometry.[123c] The changes of K_{succ} and K_{ac} were normalized by cognate changes of protein expression. This quantitative analysis showed that about 12% of K_{succ} substrates have more than 4- fold increase in abundance in SIRT5 knockout cells, including 26 histone K_{succ} sites and 108 K_{succ} substrates only identified in SIRT5 knockout cells, suggesting that SIRT5 is the master regulator of lysine succinylation. In contrast, only 0.015% of lysine acetylation substrates were found increased more than 4-fold in SIRT5 knockout cells, indicating that SIRT5, most likely, is not an enzyme for lysine acetylation, while it has been annotated as a deacetylation enzyme based on sequence alignment. Likewise, similar quantitative proteomic experiments have been used to identify K_{ac} substrate candidates for SIRT1.[320b,324]

To study the dynamic response of phosphotyrosine-based signaling events under the stimulation of EGF, Blagoev et al. used SILAC and MS to identify 81 signaling proteins, whose tyrosine phosphorylation statuses were stimulated.[351] A similar approach was used to comprehensively compare the tyrosine phosphorylation substrates and associated partners upon EGF and platelet derived growth factor (PDGF) stimulation, respectively.[352] Using this strategy, more than 900 phosphorylation sites recognized by ataxia telangiectasia mutated (ATM) and ATM- and RAD3-related kinases were identified in response to DNA damage.[353] Similarly, quantitative phospho-proteomics in combination with in vitro kinase assay identified growth factor receptor-bound protein 10 (Grb10) as an mTORC1 substrate, and this substrate negatively regulates insulin signaling.[354] Quantitative proteomics has also been used to determine changes of protein ubiquitination to identify substrates for ubiquitin-regulatory enzymes, such as cullin-RING ubiquitin ligases (CRL),[270] and RING-HECT hybrid E3 ubiquitin ligase PARKIN.[322b] To find substrate sites associated with histones, the experiment is simpler. The experiment can be focused on core histones, which can be extracted by an acid-extraction method.[121] Quantitative analysis of histone marks in cell lines, with or without an enzyme, can identify multiple histone marks that can be regulated by a PTM enzyme.

It is necessary to mention that proteins can also be nonenzymatically modified under a high concentration of corresponding CoAs and elevated pH conditions in mitochondria, including but not limited to acetylation and succinylation.[131,355]

This experimental strategy can be modified to fit into an experimental system of interest. For example, transient knockdown or overexpression of a PTM enzyme in combination with SILAC can also be used for quantifying PTM substrates. This approach likely avoids potential compensation caused by gene knockout. In addition, quantitative proteomics can

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

also be used for analysis of PTM changes in knockout animals. An in vitro-labeled or label-free quantification method can be used to label proteins from mice, with or without a genetic knockout of a gene. The changes of PTM peptides in two types of mouse tissues can be quantified accordingly. As soon as a PTM substrate candidate is identified, it can be further confirmed by biochemical approaches that have been routinely used in the biology laboratories.

7.7. A List of Proteomic Studies of PTM Substrates

With the rapidly developed mass spectrometry technology and enrichment methods for PTM peptides, many PTM proteomic studies were carried out. Here, we attempt to summarize those publications that described the well-established methods for PTM proteomic studies (Table 12), and that used these methods for PTM proteomics (Table 13).

8. CONCLUSIONS AND OUTLOOK

Epigenetic changes are critical to diverse cellular processes and diseases. The epigenome and epigenetic mechanisms will not be fully described until we have a comprehensive inventory of epigenetic marks, their dynamics, and binding partners. In the past decades, MS-based proteomics becomes the method of choice to reveal these key biochemical events. The information generated from these studies would provide a stepping stone to investigate the roles of the histone marks in chromatin structure and function.

In the past two decades, MS has been the primary tool for detection of histone marks. Equipped with a more sensitive mass spectrometer, MS technology is nowadays able to detect those histone marks with stoichiometry of 0.01% or higher. Enrichment with an anti-PTM antibody followed by MS analysis can further enhance overall detection sensitivity. Given the facts that MS analysis has been carried out in limited types of cellular systems in the past, and that new histone marks have been recently described,[6a,b,122,123b,c,128b,189,266,267] analysis of core histones from different tissues under diverse physiopathological conditions likely discovers undescribed histone marks, either the known types or the new types of histone PTMs.

The MS-based quantification technologies, using labeling (in vitro or in vivo) or label-free approaches, enable quantifying dynamic changes of histone marks. Quantification of histone PTMs that are present only in one peptide is straightforward. Nevertheless, dynamic changes of histone marks in most epigenetic processes have not yet been characterized, which, we anticipate, will be carried out in the future.

While detection and quantification of histone marks is getting easier, crosstalks among histone marks, either in *cis* or in *trans*, have not yet been carefully examined.[373] While in principle the top-down approach can be used to detect such a PTM crosstalk in histones, this approach has low detection sensitivity, when compared to bottom-up techniques. Accordingly, it is still difficult to analyze crosstalk among those PTMs with low stoichiometry. Hopefully, future instruments and methods can address this technological challenge. In addition, complication can arise, when the top-down approach is used to analyze a protein that has the same set of PTMs, but with different PTM localizations. As a

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

pretended example, a H3 protein bearing H3K*x*me2/H3K*y*ac modification would be difficult to distinguish from that containing H3K*x*ac/H3K*y*me2 modifications, where *x* and *y* represent residue numbers in the H3 protein sequence. The two isobaric molecules have the same parent masses, but with different modification sites. In principle, an antibody against a histone mark of interest can be used to pull down the histone isoform. Mapping of other PTMs in the isolated histone can generate other PTMs that coexist in the same histone molecule. Nevertheless, only limited studies have been carried out in the past.[267]

Many direct "binder" proteins as well as their binding domains have been identified in the past decade.[4e,294e,373] Nevertheless, direct "binder" proteins for the majority of histone marks have not yet been identified or analyzed. Identification of direct "binder" proteins for combinatorial histone marks (e.g., two histone marks or more) remains to be a challenging task. The experiment is more difficult when the histone marks are present in *trans* (on two or more histone proteins). In principle, this can be carried out using reconstituted nucleosome with the recombinant histones bearing histone marks of interest. New protein chemistry technologies have been developed in the past decades for the synthesis of histone proteins bearing one or a few histone marks of interest, either through chemical synthesis or through recombinant gene expression.[374] However, it takes a significant effort to make this type of unique histone proteins. In addition, given the fact that there are more than 300 hundred histone marks known,[5] it is a huge amount of work to carry out proteomics studies to identify "binders" for nucleosome-dependent, combinatorial histone marks. Therefore, such experiments have not been carried out in the past.

In the past decades, diverse enrichment methods have been developed for PTM peptides, making efficient proteomics of PTMs possible. As long as the PTM peptides can be isolated, either by an antibody or by a chemical derivatization method, and are amenable to peptide-bond fragmentation in MS, global analysis of PTM substrates is possible. This approach, in combination with a suitable quantification technique, enables quantifying dynamic changes of PTMs under diverse cellular environments. Importantly, quantitative analysis in cells, with or without expression of a PTM-regulatory enzyme, can define the key PTM substrate sites, not only in histones but also in nonhistone proteins.[129a,270] It is highly likely that for most enzymes, it is the concerted action in both histone and nonhistone protein substrates that is responsible for the phenotype of a PTM regulatory enzyme. Identification of protein substrates for a PTM enzyme has been traditionally carried out one at a time, and is laborious. Only a small number of PTM substrates are known for most PTM-regulatory enzymes (e.g., those for phosphorylation, methylation, and ubiquitination). Thus, the MS-based approach will have a key role in mapping cellular networks by defining substrates for PTM-regulatory enzymes.

In sum, tremendous progress has been made in the past decades in developing MS-based proteomic technologies and applying them to the analysis of histone marks, qualitatively and quantitatively. Many of these methods are currently being rapidly further refined. Efforts to apply the methods, either existing or optimized ones, to generate new types of information represent exciting opportunities to reveal novel epigenetic mechanisms, and to dissect complex cellular networks associated with normal physiology and disease.

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

Acknowledgments

We want to thank Dr. Casey Frankenberger, Dr. Yue Chen, and Dr. Minjia Tan for helpful discussions and suggestions. We want to thank Dr. Simone Sidoli for providing example middle-down data. This work has been supported by NIH grants GM105933, CA160036 to Y.Z., an NIH Innovator grant (DP2OD007447) from the Office of the Director, and an NSF Faculty CAREER award to B.A.G. More information about the authors' research can be found at https://benmay.uchicago.edu/faculty/zhao and http://www.med.upenn.edu/biocbiop/faculty/garcia/.

ABBREVIATIONS

Aha	azidohomoalanine
AQUA	absolute quantification
ATM	ataxia telangiectasia mutated
AUC	area under curve
BAP	biotin pentylamine
BEMAD	β-elimination followed by Michael addition
BSA	bovine serum albumin
CapNMR	capillary microcoil NMR
ChIP	chromatin immunoprecipitation
CID	collision-induced dissociation
CRL	cullin-RING ubiquitin ligases
DDA	data-dependent acquisition
DIA	data-independent acquisition
DiLeu	*N*,*N*-dimethyl leucine
DTT	dithiothreitol
ECD	electron capture dissociation
ERLIC	electrostatic repulsion–hydrophilic interaction chromatography
ES cells	embryonic stem cells
ESI	electrospray ionization
ETD	electron transfer dissociation
FDR	false discovery rate
FT	Fourier transform
Ga_2O_3	gallium(III) oxide
Grb10	growth factor receptor-bound protein 10
HAP	hydroxyapatite
HCD	higher-energy collision dissociation
HILIC	hydrophilic interaction chromatography

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

hmSILAC	“heavy” methyl SILAC
HPLC	high-performance liquid chromatography
iDiLeu	isotopic *N*,*N*-dimethyl leucine
IEF	isoelectric focusing
IMAC	immobilized metal affinity chromatography
IPI	International Protein Index
iPSCs	induced pluripotent stem cells
ITRAQ	isobaric tags for relative and absolute quantification
K_{2ohbu}	lysine 2-hydroxybutyrylation
K_{3ohbu}	lysine 3-hydroxybutyrylation
K_{4ohbu}	lysine 4-hydroxybutyrylation
K_{ac}	lysine acetylation
K_{cr}	lysine crotonylation
K_{hib}	lysine 2-hydroxyisobutyrylation
K_{me}	lysine methylation
KO	knockout
K_{pr}	lysine propionylation
K_{succ}	lysine succinylation
K_{ub}	lysine ubiquitination
LWAC	lectin weak affinity chromatography
m/z	mass-to-charge
MALDI	matrix-assisted laser desorption/ionization
Δ**mass**	mass shift caused by a post-translational modification
mES cells	mouse embryonic stem cells
MOAC	metal oxide affinity chromatography
MRM	multiple reaction monitoring
MS	mass spectrometry
MS/MS	mass spectrometry/mass spectrometry or tandem mass spectrometry
NPC	neural precursor cells
ORC	origin recognition complex
OVA	ovalbumin
PDGF	platelet derived growth factor

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

PTM	post-translational modification
SAHA	suberoylanilide hydroxamic acid
SCX	strong cation exchange
seq	sequencing
SILAC	stable isotope labeling by amino acids in cell culture
SIMAC	sequential elution from IMAC
SRM	selective reaction monitoring
SWATH	sequential window acquisition of all theoretical spectra
TiO_2	titanium dioxide
TMT	tandem mass tag
TOF	time-of-flight
TUBEs	tandem-repeated ubiquitin-binding entities
UBA	ubiquitin-associated
WB	Western blotting
WT	wild-type
ZrO_2	zirconium dioxide

References

1. (a) Gardner KE, Allis CD, Strahl BD. J Mol Biol. 2011; 409:36. [PubMed: 21272588] (b) Bannister AJ, Kouzarides T. Cell Res. 2011; 21:381. [PubMed: 21321607] (c) Jenuwein T, Allis CD. Science. 2001; 293:1074. [PubMed: 11498575] (d) Lee JS, Smith E, Shilatifard A. Cell. 2010; 142:682. [PubMed: 20813257]
2. (a) Maze I, Noh KM, Soshnev AA, Allis CD. Nat Rev Genet. 2014; 15:259. [PubMed: 24614311] (b) Banaszynski LA, Allis CD, Lewis PW. Dev Cell. 2010; 19:662. [PubMed: 21074717] (c) Schwartzentruber J, Korshunov A, Liu XY, Jones DT, Pfaff E, Jacob K, Sturm D, Fontebasso AM, Quang DA, Tonjes M, Hovestadt V, Albrecht S, Kool M, Nantel A, Konermann C, Lindroth A, Jager N, Rausch T, Ryzhova M, Korbel JO, Hielscher T, Hauser P, Garami M, Klekner A, Bognar L, Ebinger M, Schuhmann MU, Scheurlen W, Pekrun A, Fruhwald MC, Roggendorf W, Kramm C, Durken M, Atkinson J, Lepage P, Montpetit A, Zakrzewska M, Zakrzewski K, Liberski PP, Dong Z, Siegel P, Kulozik AE, Zapatka M, Guha A, Malkin D, Felsberg J, Reifenberger G, von Deimling A, Ichimura K, Collins VP, Witt H, Milde T, Witt O, Zhang C, Castelo-Branco P, Lichter P, Faury D, Tabori U, Plass C, Majewski J, Pfister SM, Jabado N. Nature. 2012; 482:226. [PubMed: 22286061] (d) Yuen BTK, Knoepfler PS. Cancer Cell. 2013; 24:567. [PubMed: 24229707]
3. Khoury GA, Baliban RC, Floudas CA. Sci Rep. 2011; 1
4. (a) Dawson MA, Kouzarides T. Cell. 2012; 150:12. [PubMed: 22770212] (b) Feinberg AP. J Int Med. 2014; 276:5.(c) Helin K, Dhanak D. Nature. 2013; 502:480. [PubMed: 24153301] (d) Baylin SB, Jones PA. Nat Rev Cancer. 2011; 11:726. [PubMed: 21941284] (e) Chi P, Allis CD, Wang GG. Nat Rev Cancer. 2010; 10:457. [PubMed: 20574448]
5. Huang H, Sabari BR, Garcia BA, Allis CD, Zhao Y. Cell. 2014; 159:458. [PubMed: 25303536]
6. (a) Dai LZ, Peng C, Montellier E, Lu ZK, Chen Y, Ishii H, Debernardi A, Buchou T, Rousseaux S, Jin FL, Sabari BR, Deng ZY, Allis CD, Ren B, Khochbin S, Zhao YM. Nat Chem Biol. 2014; 10:365. [PubMed: 24681537] (b) Tan MJ, Luo H, Lee S, Jin FL, Yang JS, Montellier E, Buchou T,

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

Cheng ZY, Rousseaux S, Rajagopal N, Lu ZK, Ye Z, Zhu Q, Wysocka J, Ye Y, Khochbin S, Ren B, Zhao YM. Cell. 2011; 146:1015.(c) Garcia BA, Pesavento JJ, Mizzen CA, Kelleher NL. Nat Methods. 2007; 4:487. [PubMed: 17529979] (d) Wisniewski JR, Zougman A, Kruger S, Mann M. Mol Cell Proteomics. 2007; 6:72. [PubMed: 17043054]

7. Ruthenburg AJ, Li H, Patel DJ, Allis CD. Nat Rev Mol Cell Biol. 2007; 8:983. [PubMed: 18037899]
8. (a) Wysocka J, Swigut T, Xiao H, Milne TA, Kwon SY, Landry J, Kauer M, Tackett AJ, Chait BT, Badenhorst P, Wu C, Allis CD. Nature. 2006; 442:86. [PubMed: 16728976] (b) Pena PV, Davrazou F, Shi X, Walter KL, Verkhusha VV, Gozani O, Zhao R, Kutateladze TG. Nature. 2006; 442:100. [PubMed: 16728977] (c) Shi X, Hong T, Walter KL, Ewalt M, Michishita E, Hung T, Carney D, Pena P, Lan F, Kaadige MR, Lacoste N, Cayrou C, Davrazou F, Saha A, Cairns BR, Ayer DE, Kutateladze TG, Shi Y, Cote J, Chua KF, Gozani O. Nature. 2006; 442:96. [PubMed: 16728974] (d) Huang Y, Fang J, Bedford MT, Zhang Y, Xu RM. Science. 2006; 312:748. [PubMed: 16601153] (e) Pray-Grant MG, Daniel JA, Schieltz D, Yates JR III, Grant PA. Nature. 2005; 433:434. [PubMed: 15647753] (f) Sims RJ III, Chen CF, Santos-Rosa H, Kouzarides T, Patel SS, Reinberg D. J Biol Chem. 2005; 280:41789. [PubMed: 16263726] (g) Wysocka J, Swigut T, Milne TA, Dou Y, Zhang X, Burlingame AL, Roeder RG, Brivanlou AH, Allis CD. Cell. 2005; 121:859. [PubMed: 15960974] (h) Iwase S, Xiang B, Ghosh S, Ren T, Lewis PW, Cochrane JC, Allis CD, Picketts DJ, Patel DJ, Li HT, Shi Y. Nat Struct Mol Biol. 2011; 18:769. [PubMed: 21666679] (i) He FH, Umehara T, Saito K, Harada T, Watanabe S, Yabuki T, Kigawa T, Takahashi M, Kuwasako K, Tsuda K, Matsuda T, Aoki M, Seki E, Kobayashi N, Guntert P, Yokoyama S, Muto Y. Structure. 2010; 18:1127. [PubMed: 20826339] (j) Hoppmann V, Thorstensen T, Kristiansen PE, Veiseth SV, Rahman MA, Finne K, Aalen RB, Aasland R. EMBO J. 2011; 30:1939. [PubMed: 21522130] (k) Chambers AL, Pearl LH, Oliver AW, Downs JA. Nucleic Acids Res. 2013; 41:9168. [PubMed: 23907388] (l) Yang N, Xu RM. Crit Rev Biochem Mol Biol. 2013; 48:211. [PubMed: 23181513] (m) Kuo AJ, Song JK, Cheung P, Ishibe-Murakami S, Yamazoe S, Chen JK, Patel DJ, Gozani O. Nature. 2012; 484:115. [PubMed: 22398447]
9. (a) Dhalluin C, Carlson JE, Zeng L, He C, Aggarwal AK, Zhou MM. Nature. 1999; 399:491. [PubMed: 10365964] (b) Sanchez R, Meslamani J, Zhou MM. Biochim Biophys Acta. 2014; 1839:676. [PubMed: 24686119]
10. Stadler SC, Allis CD. Semin Cancer Biol. 2012; 22:404. [PubMed: 22706095]
11. (a) Hu G, Zhao K. Methods Mol Biol. 2014; 1150:175. [PubMed: 24743998] (b) Hon GC, Hawkins RD, Ren B. Hum Mol Genet. 2009; 18:R195. [PubMed: 19808796]
12. Martin C, Zhang Y. Nat Rev Mol Cell Biol. 2005; 6:838. [PubMed: 16261189]
13. Barski A, Cuddapah S, Cui KR, Roh TY, Schones DE, Wang ZB, Wei G, Chepelev I, Zhao KJ. Cell. 2007; 129:823. [PubMed: 17512414]
14. (a) Downs JA, Lowndes NF, Jackson SP. Nature. 2000; 408:1001. [PubMed: 11140636] (b) Cheung WL, Turner FB, Krishnamoorthy T, Wolner B, Ahn SH, Foley M, Dorsey JA, Peterson CL, Berger SL, Allis CD. Curr Biol. 2005; 15:656. [PubMed: 15823538]
15. Doyon Y, Cayrou C, Ullah M, Landry AJ, Cote V, Selleck W, Lane WS, Tan S, Yang XJ, Cote J. Mol Cell. 2006; 21:51. [PubMed: 16387653]
16. Tsukiyama T, Wu C. Curr Opin Genet Dev. 1997; 7:182. [PubMed: 9115421]
17. (a) Narlikar GJ, Sundaramoorthy R, Owen-Hughes T. Cell. 2013; 154:490. [PubMed: 23911317] (b) Swygert SG, Peterson CL. Biochim Biophys Acta, Gene Regul Mech. 2014; 1839:728.
18. Tong JK, Hassig CA, Schnitzler GR, Kingston RE, Schreiber SL. Nature. 1998; 395:917. [PubMed: 9804427]
19. (a) Martin C, Zhang Y. Curr Opin Cell Biol. 2007; 19:266. [PubMed: 17466502] (b) Goldberg AD, Allis CD, Bernstein E. Cell. 2007; 128:635. [PubMed: 17320500]
20. (a) Waddington CH. Endeavour. 1942; 1:18.(b) Bird A. Nature. 2007; 447:396. [PubMed: 17522671]
21. Egger G, Liang G, Aparicio A, Jones PA. Nature. 2004; 429:457. [PubMed: 15164071]
22. (a) Fuks F, Hurd PJ, Wolf D, Nan X, Bird AP, Kouzarides T. J Biol Chem. 2003; 278:4035. [PubMed: 12427740] (b) Johnson L, Cao X, Jacobsen S. Curr Biol. 2002; 12:1360. [PubMed: 12194816]

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

23. (a) Tamaru H, Selker EU. Nature. 2001; 414:277. [PubMed: 11713521] (b) Jackson JP, Lindroth AM, Cao X, Jacobsen SE. Nature. 2002; 416:556. [PubMed: 11898023]

24. (a) Trojer P, Reinberg D. Cell. 2006; 125:213. [PubMed: 16630806] (b) Kouzarides T. Cell. 2007; 128:693. [PubMed: 17320507]

25. Kohli RM, Zhang Y. Nature. 2013; 502:472. [PubMed: 24153300]

26. Aebersold R, Goodlett DR. Chem Rev. 2001; 101:269. [PubMed: 11712248]

27. (a) Fenn JB, Mann M, Meng CK, Wong SF, Whitehouse CM. Science. 1989; 246:64. [PubMed: 2675315] (b) Smith RD, Loo JA, Edmonds CG, Barinaga CJ, Udseth HR. Anal Chem. 1990; 62:882. [PubMed: 2194402]

28. (a) Tanaka K, Waki H, Ido Y, Akita S, Yoshida Y, Yoshida T. Rapid Commun Mass Spectrom. 1988; 2:151.(b) Karas M, Hillenkamp F. Anal Chem. 1988; 60:2299. [PubMed: 3239801] (c) Beavis RC, Chait BT. Rapid Commun Mass Spectrom. 1989; 3:436. [PubMed: 2520224] (d) Hillenkamp F, Karas M, Beavis RC, Chait BT. Anal Chem. 1991; 63:1193A. [PubMed: 1897719]

29. Ens W, Standing KG. Methods Enzymol. 2005; 402:49. [PubMed: 16401506]

30. Yost RA, Boyd RK. Methods Enzymol. 1990; 193:154. [PubMed: 2074819]

31. Payne AH, Glish GL. Methods Enzymol. 2005; 402:109. [PubMed: 16401508]

32. Hu Q, Noll RJ, Li H, Makarov A, Hardman M, Graham Cooks R. J Mass Spectrom. 2005; 40:430. [PubMed: 15838939]

33. Marshall AG, Hendrickson CL, Jackson GS. Mass Spectrom Rev. 1998; 17:1. [PubMed: 9768511]

34. (a) Williams DH, Bradley CV, Santikarn S, Bojesen G. Biochem J. 1982; 201:105. [PubMed: 6282257] (b) Hunt DF, Henderson RA, Shabanowitz J, Sakaguchi K, Michel H, Sevilir N, Cox AL, Appella E, Engelhard VH. Science. 1992; 255:1261. [PubMed: 1546328] (c) Wilm M, Shevchenko A, Houthaeve T, Breit S, Schweigerer L, Fotsis T, Mann M. Nature. 1996; 379:466. [PubMed: 8559255] (d) Biemann K. J Proteomics. 2014; 107C:62. [PubMed: 24448399]

35. Wells JM, McLuckey SA. Methods Enzymol. 2005; 402:148. [PubMed: 16401509]

36. Olsen JV, Macek B, Lange O, Makarov A, Horning S, Mann M. Nat Methods. 2007; 4:709. [PubMed: 17721543]

37. Zubarev RA, Kelleher NL, McLafferty FW. J Am Chem Soc. 1998; 120:3265.

38. Syka JEP, Coon JJ, Schroeder MJ, Shabanowitz J, Hunt DF. Proc Natl Acad Sci USA. 2004; 101:9528. [PubMed: 15210983]

39. (a) McAlister GC, Phanstiel DH, Brumbaugh J, Westphall MS, Coon JJ. Mol Cell Proteomics. 2011; 10:O111.009456. [PubMed: 21393638] (b) Lau KW, Hart SR, Lynch JA, Wong SC, Hubbard SJ, Gaskell SJ. Rapid Commun Mass Spectrom. 2009; 23:1508. [PubMed: 19370712]

40. Jedrychowski MP, Huttlin EL, Haas W, Sowa ME, Rad R, Gygi SP. Mol Cell Proteomics. 2011; 10:M111.009910. [PubMed: 21917720]

41. Huang TY, McLuckey SA. Annu Rev Anal Chem. 2010; 3:365.

42. (a) Good DM, Wirtala M, McAlister GC, Coon JJ. Mol Cell Proteomics. 2007; 6:1942. [PubMed: 17673454] (b) Kelleher RL, Zubarev RA, Bush K, Furie B, Furie BC, McLafferty FW, Walsh CT. Anal Chem. 1999; 71:4250. [PubMed: 10517147] (c) Wang Z, Udeshi ND, O'Malley M, Shabanowitz J, Hunt DF, Hart GW. Mol Cell Proteomics. 2010; 9:153. [PubMed: 19692427]

43. Zubarev RA. Curr Opin Biotechnol. 2004; 15:12. [PubMed: 15102460]

44. (a) Roepstorff P, Fohlman J. Biomed Mass Spectrom. 1984; 11:601. [PubMed: 6525415] (b) Steen H, Mann M. Nat Rev Mol Cell Biol. 2004; 5:699. [PubMed: 15340378]

45. O'Farrell PH. J Biol Chem. 1975; 250:4007. [PubMed: 236308]

46. (a) Domon B, Aebersold R. Science. 2006; 312:212. [PubMed: 16614208] (b) Mann M, Jensen ON. Nat Biotechnol. 2003; 21:255. [PubMed: 12610572] (c) Yates JR III. Annu Rev Biophys Biomol Struct. 2004; 33:297. [PubMed: 15139815]

47. (a) Zhang Y, Fonslow BR, Shan B, Baek MC, Yates JR III. Chem Rev. 2013; 113:2343. [PubMed: 23438204] (b) Cox J, Mann M. Annu Rev Biochem. 2011; 80:273. [PubMed: 21548781]

48. Eng JK, Searle BC, Clauser KR, Tabb DL. Mol Cell Proteomics. 2011; 10:R111.009522. [PubMed: 21876205]

49. (a) Hughes C, Ma B, Lajoie GA. Proteome Bioinformatics. 2010; 604:105.(b) Seidler J, Zinn N, Boehm ME, Lehmann WD. Proteomics. 2010; 10:634. [PubMed: 19953542] (c) Kim S, Bandeira

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

N, Pevzner PA. Mol Cell Proteomics. 2009; 8:1391. [PubMed: 19254948] (d) Bandeira N, Pham V, Pevzner P, Arnott D, Lill JR. Nat Biotechnol. 2008; 26:1336. [PubMed: 19060866]

50. (a) Lam H, Aebersold R. Methods. 2011; 54:424. [PubMed: 21277371] (b) Lam H. Mol Cell Proteomics. 2011; 10:R111.008565. [PubMed: 21900153]
51. Eng JK, Mccormack AL, Yates JR. J Am Soc Mass Spectrom. 1994; 5:976. [PubMed: 24226387]
52. Perkins DN, Pappin DJ, Creasy DM, Cottrell JS. Electrophoresis. 1999; 20:3551. [PubMed: 10612281]
53. Craig R, Beavis RC. Bioinformatics. 2004; 20:1466. [PubMed: 14976030]
54. Fu Y, Yang Q, Sun R, Li D, Zeng R, Ling CX, Gao W. Bioinformatics. 2004; 20:1948. [PubMed: 15044235]
55. MacLean B, Tomazela DM, Shulman N, Chambers M, Finney GL, Frewen B, Kern R, Tabb DL, Liebler DC, MacCoss MJ. Bioinformatics. 2010; 26:966. [PubMed: 20147306]
56. Field HI, Fenyo D, Beavis RC. Proteomics. 2002; 2:36. [PubMed: 11788990]
57. Zhang N, Aebersold R, Schwikowski B. Proteomics. 2002; 2:1406. [PubMed: 12422357]
58. Hernandez P, Gras R, Frey J, Appel RD. Proteomics. 2003; 3:870. [PubMed: 12833510]
59. (a) Cox J, Mann M. Nat Biotechnol. 2008; 26:1367. [PubMed: 19029910] (b) Cox J, Neuhauser N, Michalski A, Scheltema RA, Olsen JV, Mann M. J Proteome Res. 2011; 10:1794. [PubMed: 21254760]
60. (a) Shteynberg D, Nesvizhskii AI, Moritz RL, Deutsch EW. Mol Cell Proteomics. 2013; 12:2383. [PubMed: 23720762] (b) Perez-Riverol Y, Wang R, Hermjakob H, Muller M, Vesada V, Vizcaino JA. Biochim Biophys Acta, Proteins Proteomics. 2014; 1844:63.
61. Elias JE, Haas W, Faherty BK, Gygi SP. Nat Methods. 2005; 2:667. [PubMed: 16118637]
62. Kapp EA, Schutz F, Connolly LM, Chakel JA, Meza JE, Miller CA, Fenyo D, Eng JK, Adkins JN, Omenn GS, Simpson RJ. Proteomics. 2005; 5:3475. [PubMed: 16047398]
63. Chen Y, Chen W, Cobb MH, Zhao YM. Proc Natl Acad Sci USA. 2009; 106:761. [PubMed: 19136633]
64. Creasy DM, Cottrell JS. Proteomics. 2004; 4:1534. [PubMed: 15174123]
65. Annan RS, Huddleston MJ, Verma R, Deshaies RJ, Carr SA. Anal Chem. 2001; 73:393. [PubMed: 11217738]
66. (a) Cerny M, Skalak J, Cerna H, Brzobohaty B. J Proteomics. 2013; 92:2. [PubMed: 23777897] (b) Molden RC, Goya J, Khan Z, Garcia BA. Mol Cell Proteomics. 2014; 13:1106. [PubMed: 24532841] (c) Yang C, Zhong X, Li L. Electrophoresis. 201410.1002/elps.201400017
67. (a) Lin S, Garcia BA. Methods Enzymol. 2012; 512:3. [PubMed: 22910200] (b) Filosa G, Barabino S, Bachi A. NeuroMol Med. 2013; 15:661.(c) Zhang K, Tian SS, Fan EG. Analyst. 2013; 138:1628. [PubMed: 23361263]
68. Bowie JH, Brinkworth CS, Dua S. Mass Spectrom Rev. 2002; 21:87. [PubMed: 12373746]
69. Zhang MX, Kaltashov IA. Anal Chem. 2006; 78:4820. [PubMed: 16841900]
70. Chalkley RJ, Clauser KR. Mol Cell Proteomics. 2012; 11:3. [PubMed: 22328712]
71. Beausoleil SA, Villen J, Gerber SA, Rush J, Gygi SP. Nat Biotechnol. 2006; 24:1285. [PubMed: 16964243]
72. Savitski MM, Lemeer S, Boesche M, Lang M, Mathieson T, Bantscheff M, Kuster B. Mol Cell Proteomics. 2011; 10:M110.003830. [PubMed: 21057138]
73. Ruttenberg BE, Pisitkun T, Knepper MA, Hoffert JD. J Proteome Res. 2008; 7:3054. [PubMed: 18543960]
74. Albuquerque CP, Smolka MB, Payne SH, Bafna V, Eng J, Zhou HL. Mol Cell Proteomics. 2008; 7:1389. [PubMed: 18407956]
75. Bailey CM, Sweet SMM, Cunningham DL, Zeller M, Heath JK, Cooper HJ. J Proteome Res. 2009; 8:1965. [PubMed: 19275241]
76. Taus T, Kocher T, Pichler P, Paschke C, Schmidt A, Henrich C, Mechtler K. J Proteome Res. 2011; 10:5354. [PubMed: 22073976]

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

77. Phanstiel DH, Brumbaugh J, Wenger CD, Tian SL, Probasco MD, Bailey DJ, Swaney DL, Tervo MA, Bolin JM, Ruotti V, Stewart R, Thomson JA, Coon JJ. Nat Methods. 2011; 8:821. [PubMed: 21983960]
78. Baker PR, Trinidad JC, Chalkley RJ. Mol Cell Proteomics. 2011; 10:M111.008078. [PubMed: 21490164]
79. Vaudel M, Breiter D, Beck F, Rahnenfuhrer J, Martens L, Zahedi RP. Proteomics. 2013; 13:1036. [PubMed: 23307401]
80. (a) Chait BT. Science. 2006; 314:65. [PubMed: 17023639] (b) Ahlf DR, Thomas PM, Kelleher NL. Curr Opin Chem Biol. 2013; 17:787. [PubMed: 23988518]
81. (a) Gstaiger M, Aebersold R. Nat Rev Genet. 2009; 10:617. [PubMed: 19687803] (b) Aebersold R, Mann M. Nature. 2003; 422:198. [PubMed: 12634793] (c) Olsen JV, Mann M. Mol Cell Proteomics. 2013; 12:3444. [PubMed: 24187339]
82. (a) Tipton JD, Tran JC, Catherman AD, Ahlf DR, Durbin KR, Kelleher NL. J Biol Chem. 2011; 286:25451. [PubMed: 21632550] (b) Catherman AD, Skinner OS, Kelleher NL. Biochem Biophys Res Commun. 2014; 445:683. [PubMed: 24556311]
83. (a) Young NL, DiMaggio PA, Plazas-Mayorca MD, Baliban RC, Floudas CA, Garcia BA. Mol Cell Proteomics. 2009; 8:2266. [PubMed: 19654425] (b) Taverna SD, Ueberheide BM, Liu Y, Tackett AJ, Diaz RL, Shabanowitz J, Chait BT, Hunt DF, Allis CD. Proc Natl Acad Sci USA. 2007; 104:2086. [PubMed: 17284592]
84. (a) Patrie SM, Charlebois JP, Whipple D, Kelleher NL, Hendrickson CL, Quinn JP, Marshall AG, Mukhopadhyay B. J Am Soc Mass Spectrom. 2004; 15:1099. [PubMed: 15234368] (b) Rhoads TW, Rose CM, Bailey DJ, Riley NM, Molden RC, Nestler AJ, Merrill AE, Smith LM, Hebert AS, Westphall MS, Pagliarini DJ, Garcia BA, Coon JJ. Anal Chem. 2014; 86:2314. [PubMed: 24475910] (c) Han XM, Jin M, Breuker K, McLafferty FW. Science. 2006; 314:109. [PubMed: 17023655]
85. Mechref Y. Curr Protoc Protein Sci. 2012; 12(Unit 12.11.1)
86. Bernstein BE, Mikkelsen TS, Xie X, Kamal M, Huebert DJ, Cuff J, Fry B, Meissner A, Wernig M, Plath K, Jaenisch R, Wagschal A, Feil R, Schreiber SL, Lander ES. Cell. 2006; 125:315. [PubMed: 16630819]
87. Sadygov RG, Cociorva D, Yates JR III. Nat Methods. 2004; 1:195. [PubMed: 15789030]
88. (a) Tanner S, Shu H, Frank A, Wang LC, Zandi E, Mumby M, Pevzner PA, Bafna V. Anal Chem. 2005; 77:4626. [PubMed: 16013882] (b) Tsur D, Tanner S, Zandi E, Bafna V, Pevzner PA. Nat Biotechnol. 2005; 23:1562. [PubMed: 16311586] (c) Hansen BT, Davey SW, Ham AJL, Liebler DC. J Proteome Res. 2005; 4:358. [PubMed: 15822911] (d) Tanner S, Payne SH, Dasari S, Shen Z, Wilmarth PA, David LL, Loomis WF, Briggs SP, Bafna V. J Proteome Res. 2008; 7:170. [PubMed: 18034453]
89. (a) Baliban RC, DiMaggio PA, Plazas-Mayorca MD, Young NL, Garcia BA, Floudas CA. Mol Cell Proteomics. 2010; 9:764. [PubMed: 20103568] (b) Na S, Bandeira N, Paek E. Mol Cell Proteomics. 2012; 11:M111.010199. [PubMed: 22186716] (c) Baumgartner C, Rejtar T, Kullolli M, Akella LM, Karger BL. J Proteome Res. 2008; 7:4199. [PubMed: 18686985] (d) Havilio M, Wool A. Anal Chem. 2007; 79:1362. [PubMed: 17297935] (e) Kim S, Na S, Sim JW, Park H, Jeong J, Kim H, Seo Y, Seo J, Lee KJ, Paek E. Nucleic Acids Res. 2006; 34:W258. [PubMed: 16845006] (f) Fu Y, Xiu LY, Jia W, Ye D, Sun RX, Qian XH, He SM. Mol Cell Proteomics. 2011; 10:M110.000455.(g) Han Y, Ma B, Zhang K. J Bioinform Comput Biol. 2005; 3:697. [PubMed: 16108090] (h) Liu C, Yan B, Song Y, Xu Y, Cai L. Bioinformatics. 2006; 22:E307. [PubMed: 16873487] (i) Na S, Paek E. J Proteome Res. 2009; 8:4418. [PubMed: 19658439] (j) Savitski MM, Nielsen ML, Zubarev RA. Mol Cell Proteomics. 2006; 5:935. [PubMed: 16439352] (k) Searle BC, Dasari S, Wilmarth PA, Turner M, Reddy AP, David LL, Nagalla SR. J Proteome Res. 2005; 4:546. [PubMed: 15822933]
90. (a) Oda Y, Huang K, Cross FR, Cowburn D, Chait BT. Proc Natl Acad Sci USA. 1999; 96:6591. [PubMed: 10359756] (b) MacCoss MJ, McDonald WH, Saraf A, Sadygov R, Clark JM, Tasto JJ, Gould KL, Wolters D, Washburn M, Weiss A, Clark JI, Yates JR. Proc Natl Acad Sci USA. 2002; 99:7900. [PubMed: 12060738]
91. Wu CC, MacCoss MJ, Howell KE, Matthews DE, Yates JR. Anal Chem. 2004; 76:4951. [PubMed: 15373428]

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

92. (a) Ong SE, Blagoev B, Kratchmarova I, Kristensen DB, Steen H, Pandey A, Mann M. Mol Cell Proteomics. 2002; 1:376. [PubMed: 12118079] (b) Zhu HN, Pan SQ, Gu S, Bradbury EM, Chen X. Rapid Commun Mass Spectrom. 2002; 16:2115. [PubMed: 12415544]

93. (a) Zhang RJ, Sioma CS, Thompson RA, Xiong L, Regnier FE. Anal Chem. 2002; 74:3662. [PubMed: 12175151] (b) Zhang RJ, Sioma CS, Wang SH, Regnier FE. Anal Chem. 2001; 73:5142. [PubMed: 11721911]

94. Doherty MK, Whitehead C, McCormack H, Gaskell SJ, Beynon RJ. Proteomics. 2005; 5:522. [PubMed: 15627957]

95. Savas JN, Toyama BH, Xu T, Yates JR III, Hetzer MW. Science. 2012; 335:942. [PubMed: 22300851]

96. Kruger M, Moser M, Ussar S, Thievessen I, Luber CA, Forner F, Schmidt S, Zanivan S, Fassler R, Mann M. Cell. 2008; 134:353. [PubMed: 18662549]

97. Hebert AS, Merrill AE, Bailey DJ, Still AJ, Westphall MS, Strieter ER, Pagliarini DJ, Coon JJ. Nat Methods. 2013; 10:332. [PubMed: 23435260]

98. Merrill AE, Hebert AS, MacGilvray ME, Rose CM, Bailey DJ, Bradley JC, Wood WW, ElMasri M, Westphall MS, Gasch AP, Coon JJ. Mol Cell Proteomics. 2014; 13:2503. [PubMed: 24938287]

99. (a) Thompson A, Schafer J, Kuhn K, Kienle S, Schwarz J, Schmidt G, Neumann T, Hamon C. Anal Chem. 2003; 75:1895. [PubMed: 12713048] (b) Dayon L, Hainard A, Licker V, Turck N, Kuhn K, Hochstrasser DF, Burkhard PR, Sanchez JC. Anal Chem. 2008; 80:2921. [PubMed: 18312001] (c) Werner T, Sweetman G, Savitski MF, Mathieson T, Bantscheff M, Savitski MM. Anal Chem. 2014; 86:3594. [PubMed: 24579773] (d) Werner T, Becher I, Sweetman G, Doce C, Savitski MM, Bantscheff M. Anal Chem. 2012; 84:7188. [PubMed: 22881393] (e) McAlister GC, Huttlin EL, Haas W, Ting L, Jedrychowski MP, Rogers JC, Kuhn K, Pike I, Grothe RA, Blethrow JD, Gygi SP. Anal Chem. 2012; 84:7469. [PubMed: 22880955]

100. (a) Ross PL, Huang YN, Marchese JN, 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, Pappin DJ. Mol Cell Proteomics. 2004; 3:1154. [PubMed: 15385600] (b) Unwin RD, Pierce A, Watson RB, Sternberg DW, Whetton AD. Mol Cell Proteomics. 2005; 4:924. [PubMed: 15849271] (c) Pierce A, Unwin RD, Evans CA, Griffiths S, Carney L, Zhang L, Jaworska E, Lee CF, Blinco D, Okoniewski MJ, Miller CJ, Bitton DA, Spooncer E, Whetton AD. Mol Cell Proteomics. 2008; 7:853. [PubMed: 17951628]

101. (a) Xiang F, Ye H, Chen RB, Fu Q, Li LJ. Anal Chem. 2010; 82:2817. [PubMed: 20218596] (b) Greer T, Lietz CB, Xiang F, Li L. J Am Soc Mass Spectrom. 201410.1007/s13361

102. Nahnsen S, Bielow C, Reinert K, Kohlbacher O. Mol Cell Proteomics. 2013; 12:549. [PubMed: 23250051]

103. Dephoure N, Gygi SP. Sci Signaling. 2012; 5:rs2.

104. (a) Bantscheff M, Boesche M, Eberhard D, Matthieson T, Sweetman G, Kuster B. Mol Cell Proteomics. 2008; 7:1702. [PubMed: 18511480] (b) Ow SY, Salim M, Noirel J, Evans C, Rehman I, Wright PC. J Proteome Res. 2009; 8:5347. [PubMed: 19754192] (c) Karp NA, Huber W, Sadowski PG, Charles PD, Hester SV, Lilley KS. Mol Cell Proteomics. 2010; 9:1885. [PubMed: 20382981] (d) Shirran SL, Botting CH. J Proteomics. 2010; 73:1391. [PubMed: 20230925]

105. (a) Ting L, Rad R, Gygi SP, Haas W. Nat Methods. 2011; 8:937. [PubMed: 21963607] (b) Koehler CJ, Arntzen MO, Treumann A, Thiede B. Methods Mol Biol. 2012; 893:129. [PubMed: 22665299] (c) Vincent CE, Rensvold JW, Westphall MS, Pagliarini DJ, Coon JJ. Anal Chem. 2013; 85:2079. [PubMed: 23046161] (d) Bamberger C, Pankow S, Park SK, Yates JR III. J Proteome Res. 2014; 13:1494. [PubMed: 24417624]

106. (a) Picotti P, Aebersold R. Nat Methods. 2012; 9:555. [PubMed: 22669653] (b) Maiolica A, Junger MA, Ezkurdia I, Aebersold R. J Proteomics. 2012; 75:3495. [PubMed: 22579752] (c) James A, Jorgensen C. Methods Mol Biol. 2010; 658:167. [PubMed: 20839104] (d) Abbatiello SE, Mani DR, Schilling B, Maclean B, Zimmerman LJ, Feng X, Cusack MP, Sedransk N, Hall SC, Addona T, Allen S, Dodder NG, Ghosh M, Held JM, Hedrick V, Inerowicz HD, Jackson A, Keshishian H, Kim JW, Lyssand JS, Riley CP, Rudnick P, Sadowski P, Shaddox K, Smith D, Tomazela D, Wahlander A, Waldemarson S, Whitwell CA, You J, Zhang S, Kinsinger CR, Mesri M, Rodriguez H, Borchers CH, Buck C, Fisher SJ, Gibson BW, Liebler D, Maccoss M, Neubert

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

TA, Paulovich A, Regnier F, Skates SJ, Tempst P, Wang M, Carr SA. Mol Cell Proteomics. 2013; 12:2623. [PubMed: 23689285]

107. Kirkpatrick DS, Denison C, Gygi SP. Nat Cell Biol. 2005; 7:750. [PubMed: 16056266]

108. (a) Carvalho PC, Hewel J, Barbosa VC, Yates JR III. Genet Mol Res. 2008; 7:342. [PubMed: 18551400] (b) Lundgren DH, Hwang SI, Wu LF, Han DK. Expert Rev Proteomics. 2010; 7:39. [PubMed: 20121475] (c) Usaite R, Wohlschlegel J, Venable JD, Park SK, Nielsen J, Olsson L, Yates JR III. J Proteome Res. 2008; 7:266. [PubMed: 18173223]

109. Neilson KA, Ali NA, Muralidharan S, Mirzaei M, Mariani M, Assadourian G, Lee A, van Sluyter SC, Haynes PA. Proteomics. 2011; 11:535. [PubMed: 21243637]

110. (a) Gillet LC, Navarro P, Tate S, Rost H, Selevsek N, Reiter L, Bonner R, Aebersold R. Mol Cell Proteomics. 2012; 11:O111.016717. [PubMed: 22261725] (b) Collins BC, Gillet LC, Rosenberger G, Rost HL, Vichalkovski A, Gstaiger M, Aebersold R. Nat Methods. 2013; 10:1246. [PubMed: 24162925] (c) Geromanos SJ, Vissers JPC, Silva JC, Dorschel CA, Li GZ, Gorenstein MV, Bateman RH, Langridge JI. Proteomics. 2009; 9:1683. [PubMed: 19294628] (d) Blackburn K, Mbeunkui F, Mitra SK, Mentzel T, Goshe MB. J Proteome Res. 2010; 9:3621. [PubMed: 20450226] (e) Liu Y, Huttenhain R, Surinova S, Gillet LC, Mouritsen J, Brunner R, Navarro P, Aebersold R. Proteomics. 2013; 13:1247. [PubMed: 23322582]

111. Rost HL, Rosenberger G, Navarro P, Gillet L, Miladinovic SM, Schubert OT, Wolski W, Collins BC, Malmstrom J, Malmstrom L, Aebersold R. Nat Biotechnol. 2014; 32:219. [PubMed: 24727770]

112. Craft GE, Chen A, Nairn AC. Methods. 2013; 61:186. [PubMed: 23623823]

113. (a) Waanders LF, Hanke S, Mann M. J Am Soc Mass Spectrom. 2007; 18:2058. [PubMed: 17920290] (b) Collier TS, Hawkridge AM, Georgianna DR, Payne GA, Muddiman DC. Anal Chem. 2008; 80:4994. [PubMed: 18512951]

114. (a) Hung CW, Tholey A. Anal Chem. 2012; 84:161. [PubMed: 22103715] (b) Jia CX, Yu Q, Wang JX, Li LJ. Proteomics. 2014; 14:1185. [PubMed: 24532205]

115. (a) Mazur MT, Cardasis HL, Spellman DS, Liaw A, Yates NA, Hendrickson RC. Proc Natl Acad Sci USA. 2010; 107:7728. [PubMed: 20388904] (b) Ntai I, Kim K, Fellers RT, Skinner OS, Smith AD, Early BP, Savaryn JP, LeDuc RD, Thomas PM, Kelleher NL. Anal Chem. 2014; 86:4961. [PubMed: 24807621]

116. Allfrey VG, Faulkner R, Mirsky AE. Proc Natl Acad Sci USA. 1964; 51:786. [PubMed: 14172992]

117. Gershey EL, Vidali G, Allfrey VG. J Biol Chem. 1968; 243:5018. [PubMed: 5679978]

118. Holt GD, Hart GW. J Biol Chem. 1986; 261:8049. [PubMed: 3086323]

119. (a) Sakabe K, Hart GW. J Biol Chem. 2010; 285:34460. [PubMed: 20805223] (b) Sakabe K, Wang ZH, Hart GW. Proc Natl Acad Sci USA. 2010; 107:19915. [PubMed: 21045127] (c) Zhang SS, Roche K, Nasheuer HP, Lowndes NF. J Biol Chem. 2011; 286:37483. [PubMed: 21896475]

120. (a) Xiong L, Ping LY, Yuan BF, Wang YS. J Am Soc Mass Spectrom. 2009; 20:1172. [PubMed: 19303795] (b) Zhang JM, Chen Y, Zhang ZH, Xing G, Wysocka J, Zhao YM. J Proteome Res. 2010; 9:585. [PubMed: 19877717]

121. Shechter D, Dormann HL, Allis CD, Hake SB. Nat Protoc. 2007; 2:1445. [PubMed: 17545981]

122. Chen Y, Sprung R, Tang Y, Ball H, Sangras B, Kim SC, Falck JR, Peng JM, Gu W, Zhao YM. Mol Cell Proteomics. 2007; 6:812. [PubMed: 17267393]

123. (a) Zhang Z, Tan M, Xie Z, Dai L, Chen Y, Zhao Y. Nat Chem Biol. 2011; 7:58. [PubMed: 21151122] (b) Xie ZY, Dai JBA, Dai LZ, Tan MJ, Cheng ZY, Wu YM, Boeke JD, Zhao YM. Mol Cell Proteomics. 2012; 11:100. [PubMed: 22389435] (c) Park J, Chen Y, Tishkoff DX, Peng C, Tan M, Dai L, Xie Z, Zhang Y, Zwaans BM, Skinner ME, Lombard DB, Zhao Y. Mol Cell. 2013; 50:919. [PubMed: 23806337]

124. Li X, Fekete A, Englmann M, Frommberger M, Lv S, Chen G, Schmitt-Kopplin P. Anal Bioanal Chem. 2007; 389:1439. [PubMed: 17849105]

125. Senko MW, Remes PM, Canterbury JD, Mathur R, Song Q, Eliuk SM, Mullen C, Earley L, Hardman M, Blethrow JD, Bui H, Specht A, Lange O, Denisov E, Makarov A, Horning S, Zabrouskov V. Anal Chem. 2013; 85:11710. [PubMed: 24251866]

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

126. Kind T, Fiehn O. BMC Bioinf. 2007; 8:105.
127. Schroeder FC, Gronquist M. Angew Chem, Int Ed. 2006; 45:7122.
128. (a) Peng C, Lu ZK, Xie ZY, Cheng ZY, Chen Y, Tan MJ, Luo H, Zhang Y, He W, Yang K, Zwaans BMM, Tishkoff D, Ho L, Lombard D, He TC, Dai JB, Verdin E, Ye Y, Zhao YM. Mol Cell Proteomics. 2011; 10:M111.012658. [PubMed: 21908771] (b) Tan MJ, Peng C, Anderson KA, Chhoy P, Xie ZY, Dai LZ, Park J, Chen Y, Huang H, Zhang Y, Ro J, Wagner GR, Green MF, Madsen AS, Schmiesing J, Peterson BS, Xu GF, Ilkayeva OR, Muehlbauer MJ, Braulke T, Muhlhausen C, Backos DS, Olsen CA, McGuire PJ, Pletcher SD, Lombard DB, Hirschey MD, Zhao YM. Cell Metab. 2014; 19:605. [PubMed: 24703693]
129. (a) Chen Y, Zhao WH, Yang JS, Cheng ZY, Luo H, Lu ZK, Tan MJ, Gu W, Zhao YM. Mol Cell Proteomics. 2012; 11:1048. [PubMed: 22826441] (b) Kim SC, Sprung R, Chen Y, Xu YD, Ball H, Pei JM, Cheng TL, Kho Y, Xiao H, Xiao L, Grishin NV, White M, Yang XJ, Zhao YM. Mol Cell. 2006; 23:607. [PubMed: 16916647] (c) Choudhary C, Kumar C, Gnad F, Nielsen ML, Rehman M, Walther TC, Olsen JV, Mann M. Science. 2009; 325:834. [PubMed: 19608861] (d) Zhao SM, Xu W, Jiang WQ, Yu W, Lin Y, Zhang TF, Yao J, Zhou L, Zeng YX, Li H, Li YX, Shi J, An WL, Hancock SM, He FC, Qin LX, Chin J, Yang PY, Chen X, Lei QY, Xiong Y, Guan KL. Science. 2010; 327:1000. [PubMed: 20167786]
130. Frenkel EP, Kitchens RL. J Biol Chem. 1977; 252:504. [PubMed: 13070]
131. Weinert BT, Iesmantavicius V, Wagner SA, Scholz C, Gummesson B, Beli P, Nystrom T, Choudhary C. Mol Cell. 2013; 51:265. [PubMed: 23830618]
132. (a) Brownell JE, Zhou J, Ranalli T, Kobayashi R, Edmondson DG, Roth SY, Allis CD. Cell. 1996; 84:843. [PubMed: 8601308] (b) Carmen AA, Rundlett SE, Grunstein M. J Biol Chem. 1996; 271:15837. [PubMed: 8663039] (c) Taunton J, Hassig CA, Schreiber SL. Science. 1996; 272:408. [PubMed: 8602529] (d) Rea S, Eisenhaber F, O'Carroll N, Strahl BD, Sun ZW, Schmid M, Opravil S, Mechtler K, Ponting CP, Allis CD, Jenuwein T. Nature. 2000; 406:593. [PubMed: 10949293]
133. (a) Bonenfant D, Coulot M, Towbin H, Schindler P, van Oostrum J. Mol Cell Proteomics. 2006; 5:541. [PubMed: 16319397] (b) Unnikrishnan A, Gafken PR, Tsukiyama T. Nat Struct Mol Biol. 2010; 17:430. [PubMed: 20228802] (c) Jufvas A, Stralfors P, Vener AV. PLoS One. 2011; 6:e15960. [PubMed: 21249133] (d) Tropberger P, Pott S, Keller C, Kamieniarz-Gdula K, Caron M, Richter F, Li GH, Mittler G, Liu ET, Buhler M, Margueron R, Schneider R. Cell. 2013; 152:859. [PubMed: 23415232]
134. Bernstein E, Muratore-Schroeder TL, Diaz RL, Chow JC, Changolkar LN, Shabanowitz J, Heard E, Pehrson JR, Hunt DF, Allis CD. Proc Natl Acad Sci USA. 2008; 105:1533. [PubMed: 18227505]
135. Duncan EM, Muratore-Schroeder TL, Cook RG, Garcia BA, Shabanowitz J, Hunt DF, Allis CD. Cell. 2008; 135:284. [PubMed: 18957203]
136. (a) Smith CM, Haimberger ZW, Johnson CO, Wolf AJ, Gafken PR, Zhang ZL, Parthun MR, Gottschling DE. Proc Natl Acad Sci USA. 2002; 99:16454. [PubMed: 12196632] (b) Mackeen MM, Kramer HB, Chang KH, Coleman ML, Hopkinson RJ, Schofield CJ, Kessler BM. J Proteome Res. 2010; 9:4082. [PubMed: 20583823]
137. (a) Peters AHFM, Kubicek S, Mechtler K, O'Sullivan RJ, Derijck AAHA, Perez-Burgos L, Kohimaier A, Opravil S, Tachibana M, Shinkai Y, Martens JHA, Jenuwein T. Mol Cell. 2003; 12:1577. [PubMed: 14690609] (b) Syka JEP, Marto JA, Bai DL, Horning S, Senko MW, Schwartz JC, Ueberheide B, Garcia B, Busby S, Muratore T, Shabanowitz J, Hunt DF. J Proteome Res. 2004; 3:621. [PubMed: 15253445] (c) Garcia BA, Joshi S, Thomas CE, Chitta RK, Diaz RL, Busby SA, Andrews PC, Loo RRO, Shabanowitz J, Kelleher NL, Mizzen CA, Allis CD, Hunt DF. Mol Cell Proteomics. 2006; 5:1593. [PubMed: 16835217] (d) Garcia BA, Shabanowitz J, Hunt DF. Curr Opin Chem Biol. 2007; 11:66. [PubMed: 17157550]
138. (a) Mandava V, Fernandez JP, Deng H, Janzen CJ, Hake SB, Cross GAM. Mol Biochem Parasitol. 2007; 156:41. [PubMed: 17714803] (b) Ouvry-Patat SA, Schey KL. J Mass Spectrom. 2007; 42:664. [PubMed: 17405180] (c) Robin P, Fritsch L, Philipot O, Svinarchuk F, Ait-Si-Ali S. Genome Biol. 2007; 8:R270. [PubMed: 18096052] (d) Drogaris P, Wurtele H, Masumoto H, Verreault A, Thibault P. Anal Chem. 2008; 80:6698. [PubMed: 18671409] (e) Darwanto A, Curtis MP, Schrag M, Kirsch W, Liu P, Xu GL, Neidigh JW, Zhang KL. J Biol Chem. 2010;

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

285:21868. [PubMed: 20442396] (f) Anderson DC, Green GR, Smith K, Selker EU. Biochemistry. 2010; 49:5244. [PubMed: 20462202] (g) Wilczek C, Chitta R, Woo E, Shabanowitz J, Chait BT, Hunt DF, Shechter D. J Biol Chem. 2011; 286:42221. [PubMed: 22009756] (h) Yuan W, Xu M, Huang C, Liu N, Chen S, Zhu B. J Biol Chem. 2011; 286:7983. [PubMed: 21239496] (i) Peach SE, Rudomin EL, Udeshi ND, Carr SA, Jaffe JD. Mol Cell Proteomics. 2012; 11:128. [PubMed: 22442256] (j) Sidoli S, Cheng L, Jensen ON. J Proteomics. 2012; 75:3419. [PubMed: 22234360] (k) Zheng YP, Sweet SMM, Popovic R, Martinez-Garcia E, Tipton JD, Thomas PM, Licht JD, Kelleher NL. Proc Natl Acad Sci USA. 2012; 109:13549. [PubMed: 22869745]

139. Britton LMP, Gonzales-Cope M, Zee BM, Garcia BA. Expert Rev Proteomics. 2011; 8:631. [PubMed: 21999833]
140. (a) Jaffe JD, Wang Y, Chan HM, Zhang JH, Huether R, Kryukov GV, Bhang HEC, Taylor JE, Hu M, Englund NP, Yan F, Wang ZF, McDonald ER, Wei L, Ma J, Easton J, Yu ZT, deBeaumount R, Gibaja V, Venkatesan K, Schlegel R, Sellers WR, Keen N, Liu J, Caponigro G, Barretina J, Cooke VG, Mullighan C, Carr SA, Downing JR, Garraway LA, Stegmeier F. Nat Genet. 2013; 45:1386. [PubMed: 24076604] (b) Liao RJ, Wu HP, Deng HB, Yu YY, Hu M, Zhai HL, Yang PY, Zhou SL, Yi W. Anal Chem. 2013; 85:2253. [PubMed: 23339652]
141. Tweedie-Cullen RY, Brunner AM, Grossmann J, Mohanna S, Sichau D, Nanni P, Panse C, Mansuy IM. PLoS One. 2012; 7:e36980. [PubMed: 22693562]
142. LeRoy G, DiMaggio PA, Chan EY, Zee BM, Blanco MA, Bryant B, Flaniken IZ, Liu S, Kang YB, Trojer P, Garcia BA. Epigenet Chromatin. 2013; 6
143. (a) MacCoss MJ, Wu CC, Yates JR III. Anal Chem. 2002; 74:5593. [PubMed: 12433093] (b) Eng JK, Fischer B, Grossmann J, MacCoss MJ. J Proteome Res. 2008; 7:4598. [PubMed: 18774840]
144. (a) Liu XW, Hengel S, Wu S, Tolic N, Pasa-Tolic L, Pevzner PA. J Proteome Res. 2013; 12:5830. [PubMed: 24188097] (b) Guan S, Burlingame AL. Mol Cell Proteomics. 2010; 9:804. [PubMed: 19955076]
145. (a) Zheng YP, John S, Pesavento JJ, Schultz-Norton JR, Schiltz RL, Baek S, Nardulli AM, Hager GL, Kelleher NL, Mizzen CA. J Cell Biol. 2010; 189:407. [PubMed: 20439994] (b) Gardner KE, Zhou L, Parra MA, Chen XA, Strahl BD. PLoS One. 2011; 6:e16244. [PubMed: 21249157] (c) Jung HR, Sidoli S, Haldbo S, Sprenger RR, Schwammle V, Pasini D, Helin K, Jensen ON. Anal Chem. 2013; 85:8232. [PubMed: 23889513]
146. Sidoli S, Schwammle V, Ruminowicz C, Hansen TA, Wu X, Helin K, Jensen ON. Proteomics. 2014:2200. [PubMed: 25073878]
147. (a) Hunter T. Mol Cell. 2007; 28:730. [PubMed: 18082598] (b) Yang XJ, Seto E. Mol Cell. 2008; 31:449. [PubMed: 18722172]
148. (a) Mikesh LM, Ueberheide B, Chi A, Coon JJ, Syka JEP, Shabanowitz J, Hunt DF. Biochim Biophys Acta, Proteins Proteomics. 2006; 1764:1811.(b) Shvartsburg AA, Zheng YP, Smith RD, Kelleher NL. Anal Chem. 2012; 84:6317. [PubMed: 22812477] (c) Webb K, Bennett EJ. Nat Methods. 2013; 10:620. [PubMed: 23807195] (d) Schwammle V, Aspalter CM, Sidoli S, Jensen ON. Mol Cell Proteomics. 2014; 13:1855. [PubMed: 24741113]
149. Mertins P, Qiao JW, Patel J, Udeshi ND, Clauser KR, Mani DR, Burgess MW, Gillette MA, Jaffe JD, Carr SA. Nat Methods. 2013; 10:634. [PubMed: 23749302]
150. (a) Colak G, Pougovkina O, Dai L, Tan M, te Brinke H, Huang H, Cheng Z, Park J, Wan X, Wanders RJ, Lombard DB, de Boer VC, Zhao Y. submitted. (b) Bonenfant D, Towbin H, Coulot M, Schindler P, Mueller DR, van Oostrum J. Mol Cell Proteomics. 2007; 6:1917. [PubMed: 17644761] (c) Jung HR, Pasini D, Helin K, Jensen ON. Mol Cell Proteomics. 2010; 9:838. [PubMed: 20150217] (d) Cuomo A, Moretti S, Minucci S, Bonaldi T. Amino Acids. 2011; 41:387. [PubMed: 20617350] (e) Guan XY, Rastogi N, Parthun MR, Freitas MA. Mol Cell Proteomics. 2013; 12:2048. [PubMed: 23592332] (f) Hammond SL, Byrum SD, Namjoshi S, Graves HK, Dennehey BK, Tackett AJ, Tyler JK. Cell Cycle. 2014; 13:440. [PubMed: 24275038] (g) Zhao YL, Fang XL, Wang Y, Zhang JM, Jiang S, Liu Z, Ma ZY, Xu LY, Li EM, Zhang K. Curr Pharm Des. 2014; 20:1866. [PubMed: 23888955] (h) Zhang KL, Schrag M, Crofton A, Trivedi R, Vinters H, Kirsch W. Proteomics. 2012; 12:1261. [PubMed: 22577027]

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

151. Sridharan R, Gonzales-Cope M, Chronis C, Bonora G, McKee R, Huang C, Patel S, Lopez D, Mishra N, Pellegrini M, Carey M, Garcia BA, Plath K. Nat Cell Biol. 2013; 15:872. [PubMed: 23748610]
152. Beck HC, Nielsen EC, Matthiesen R, Jensen LH, Sehested M, Finn P, Grauslund M, Hansen AM, Jensen ON. Mol Cell Proteomics. 2006; 5:1314. [PubMed: 16627869]
153. Lin S, Wein S, Gonzales-Cope M, Otte GL, Yuan ZF, Afjehi-Sadat L, Maile T, Berger SL, Rush J, Lill JR, Arnott D, Garcia BA. Mol Cell Proteomics. 2014; 13:2450. [PubMed: 25000943]
154. Gerber SA, Rush J, Stemman O, Kirschner MW, Gygi SP. Proc Natl Acad Sci USA. 2003; 100:6940. [PubMed: 12771378]
155. Steen H, Jebanathirajah JA, Springer M, Kirschner MW. Proc Natl Acad Sci USA. 2005; 102:3948. [PubMed: 15741271]
156. Weinert BT, Iesmantavicius V, Moustafa T, Scholz C, Wagner SA, Magnes C, Zechner R, Choudhary C. Mol Syst Biol. 2014; 10:716. [PubMed: 24489116]
157. Olsen JV, Vermeulen M, Santamaria A, Kumar C, Miller ML, Jensen LJ, Gnad F, Cox J, Jensen TS, Nigg EA, Brunak S, Mann M. Sci Signaling. 2010; 3:ra3.
158. Wu RH, Haas W, Dephoure N, Huttlin EL, Zhai B, Sowa ME, Gygi SP. Nat Methods. 2011; 8:677. [PubMed: 21725298]
159. (a) Smith CM, Gafken PR, Zhang ZL, Gottschling DE, Smith JB, Smith DL. Anal Biochem. 2003; 316:23. [PubMed: 12694723] (b) Hersman E, Nelson DM, Griffith WP, Jelinek C, Cotter RJ. Int J Mass Spectrom. 2012; 312:5. [PubMed: 22389584] (c) Smith CM. Methods. 2005; 36:395. [PubMed: 16098763]
160. Baeza J, Dowell JA, Smallegan MJ, Fan J, Amador-Noguez D, Khan Z, Denu JM. J Biol Chem. 2014
161. Cieniewicz AM, Moreland L, Ringel AE, Mackintosh SG, Raman A, Gilbert TM, Wolberger C, Tackett AJ, Taverna SD. Mol Cell Proteomics. 2014; 13:2896. [PubMed: 25106422]
162. Rexach JE, Rogers CJ, Yu SH, Tao JF, Sun YE, Hsieh-Wilson LC. Nat Chem Biol. 2010; 6:645. [PubMed: 20657584]
163. Balasubramaniam D, Eissler CL, Stauffachei CV, Hall MC. Proteomics. 2010; 10:4301. [PubMed: 21046619]
164. Drogaris P, Villeneuve V, Pomies C, Lee EH, Bourdeau V, Bonneil E, Ferbeyre G, Verreault A, Thibault P. Sci Rep. 2012; 210.1038/srep00220
165. Zhang CC, Liu YF, Andrews PC. Methods. 2013; 61:236. [PubMed: 23454290]
166. Deal RB, Henikoff JG, Henikoff S. Science. 2010; 328:1161. [PubMed: 20508129]
167. Eichelbaum K, Krijgsveld J. Methods Mol Biol. 2014; 1174:101. [PubMed: 24947377]
168. Xu M, Long C, Chen X, Huang C, Chen S, Zhu B. Science. 2010; 328:94. [PubMed: 20360108]
169. (a) Ong SE, Mittler G, Mann M. Nat Methods. 2004; 1:119. [PubMed: 15782174] (b) Fodor BD, Kubicek S, Yonezawa M, O'Sullivan RJ, Sengupta R, Perez-Burgos L, Opravil S, Mechtler K, Schotta G, Jenuwein T. Genes Dev. 2006; 20:1557. [PubMed: 16738407]
170. Young NL, DiMaggio PA, Garcia BA. Cell Mol Life Sci. 2010; 67:3983. [PubMed: 20683756]
171. (a) Zee BM, Britton LMP, Wolle D, Haberman DM, Garcia BA. Mol Cell Biol. 2012; 32:2503. [PubMed: 22547680] (b) Zee BM, Levin RS, Xu B, LeRoy G, Wingreen NS, Garcia BA. J Biol Chem. 2010; 285:3341. [PubMed: 19940157]
172. Osley MA. Annu Rev Biochem. 1991; 60:827. [PubMed: 1883210]
173. Pesavento JJ, Yang H, Kelleher NL, Mizzen CA. Mol Cell Biol. 2008; 28:468. [PubMed: 17967882]
174. Evertts AG, Zee BM, Dimaggio PA, Gonzales-Cope M, Coller HA, Garcia BA. J Biol Chem. 2013; 288:12142. [PubMed: 23482559]
175. Zheng Y, Thomas PM, Kelleher NL. Nat Commun. 2013; 4:2203. [PubMed: 23892279]
176. Duan Q, Chen HB, Costa M, Dai W. J Biol Chem. 2008; 283:33585. [PubMed: 18835819]
177. Oneill LP, Turner BM. EMBO J. 1995; 14:3946. [PubMed: 7664735]

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript

178. (a) Arnott D, Kishiyama A, Luis EA, Ludlum SG, Marsters JC, Stults JT. Mol Cell Proteomics. 2002; 1:148. [PubMed: 12096133] (b) Aebersold R, Burlingame AL, Bradshaw RA. Mol Cell Proteomics. 2013; 12:2381. [PubMed: 23756428]
179. Weber CM, Henikoff S. Genes Dev. 2014; 28:672. [PubMed: 24696452]
180. Redon C, Pilch D, Rogakou E, Sedelnikova O, Newrock K, Bonner W. Curr Opin Genet Dev. 2002; 12:162. [PubMed: 11893489]
181. Soldi M, Cuomo A, Bremang M, Bonaldi T. Int J Mol Sci. 2013; 14:5402. [PubMed: 23466885]
182. (a) Garcia BA, Busby SA, Barber CM, Shabanowitz J, Allis CD, Hunt DF. J Proteome Res. 2004; 3:1219. [PubMed: 15595731] (b) Karaczyn AA, Golebiowski F, Kasprzak KS. Chem Res Toxicol. 2005; 18:1934. [PubMed: 16359184] (c) Chu FX, Nusinow DA, Chalkely RJ, Plath K, Panning B, Burlingame AL. Mol Cell Proteomics. 2006; 5:194. [PubMed: 16210244] (d) Miao J, Fan Q, Cui L, Li JS, Li JY, Cui LW. Gene. 2006; 369:53. [PubMed: 16410041] (e) Bergmuller E, Gehrig PM, Gruissem W. J Proteome Res. 2007; 6:3655. [PubMed: 17691833] (f) Villar-Garea A, Imhof A. PLoS One. 2008; 3:e1553. [PubMed: 18253500] (g) Bonet-Costa C, Vilaseca M, Diema C, Vujatovic O, Vaquero A, Omenaca N, Castejon L, Bernues J, Giralt E, Azorin F. J Proteomics. 2012; 75:4124. [PubMed: 22647927] (h) Draker R, Ng MK, Sarcinella E, Ignatchenko V, Kislinger T, Cheung P. PLoS Genet. 2012; 8:e1003047. [PubMed: 23144632]
183. (a) Talbert PB, Henikoff S. Nat Rev Mol Cell Biol. 2010; 11:264. [PubMed: 20197778] (b) Arnaudo AM, Molden RC, Garcia BA. Crit Rev Biochem Mol Biol. 2011; 46:284. [PubMed: 21526979]
184. Zee BM, Levin RS, DiMaggio PA, Garcia BA. Epigenet Chromatin. 2010; 310.1186/1756
185. Dion MF, Kaplan T, Kim M, Buratowski S, Friedman N, Rando OJ. Science. 2007; 315:1405. [PubMed: 17347438]
186. (a) Siuti N, Roth MJ, Mizzen CA, Kelleher NL, Pesavento JJ. J Proteome Res. 2006; 5:233. [PubMed: 16457587] (b) Thomas CE, Kelleher NL, Mizzen CA. J Proteome Res. 2006; 5:240. [PubMed: 16457588] (c) Karch KR, Denizio JE, Black BE, Garcia BA. Front Genet. 2013; 4:264. [PubMed: 24391660] (d) Yuan ZF, Arnaudo AM, Garcia BA. Annu Rev Anal Chem. 2014; 7:113.
187. Boyne MT, Pesavento JJ, Mizzen CA, Kelleher NL. J Proteome Res. 2006; 5:248. [PubMed: 16457589]
188. (a) Portela A, Esteller M. Nat Biotechnol. 2010; 28:1057. [PubMed: 20944598] (b) Ordovas JM, Smith CE. Nat Rev Cardiol. 2010; 7:510. [PubMed: 20603647]
189. Chou RH, Wang YN, Hsieh YH, Li LY, Xia W, Chang WC, Chang LC, Cheng CC, Lai CC, Hsu JL, Chang WJ, Chiang SY, Lee HJ, Liao HW, Chuang PH, Chen HY, Wang HL, Kuo SC, Chen CH, Yu YL, Hung MC. Dev Cell. 2014; 30:224. [PubMed: 25073158]
190. (a) Polioudaki H, Markaki Y, Kourmouli N, Dialynas G, Theodoropoulos PA, Singh PB, Georgatos SD. FEBS Lett. 2004; 560:39. [PubMed: 14987995] (b) Casas-Mollano JA, Jeong BR, Xu JP, Moriyama H, Cerutti H. Proc Natl Acad Sci USA. 2008; 105:6486. [PubMed: 18420823]
191. Metzger E, Imhof A, Patel D, Kahl P, Hoffmeyer K, Friedrichs N, Muller JM, Greschik H, Kirfel J, Ji S, Kunowska N, Beisenherz-Huss C, Gunther T, Buettner R, Schule R. Nature. 2010; 464:792. [PubMed: 20228790]
192. (a) Besant PG, Attwood PV. Mol Cell Biochem. 2009; 329:93. [PubMed: 19387796] (b) Besant PG, Attwood PV. Biochem Soc Trans. 2012; 40:290. [PubMed: 22260708]
193. Tweedie-Cullen RY, Reck JM, Mansuy IM. J Proteome Res. 2009; 8:4966. [PubMed: 19737024]
194. (a) Sassone-Corsi P, Mizzen CA, Cheung P, Crosio C, Monaco L, Jacquot S, Hanauer A, Allis CD. Science. 1999; 285:886. [PubMed: 10436156] (b) Johansen KM, Johansen J. Chromosome Res. 2006; 14:393. [PubMed: 16821135]
195. Preuss U, Landsberg G, Scheidtmann KH. Nucleic Acids Res. 2003; 31:878. [PubMed: 12560483]
196. Pantazis P, Bonner WM. J Biol Chem. 1981; 256:4669. [PubMed: 7217105]
197. (a) Goto H, Tomono Y, Ajiro K, Kosako H, Fujita M, Sakurai M, Okawa K, Iwamatsu A, Okigaki T, Takahashi T, Inagaki M. J Biol Chem. 1999; 274:25543. [PubMed: 10464286] (b) Metzger E, Yin N, Wissmann M, Kunowska N, Fischer K, Friedrichs N, Patnaik D, Higgins

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

JMG, Potier N, Scheidtmann KH, Buettner R, Schule R. Nat Cell Biol. 2008; 10:53. [PubMed: 18066052]

198. (a) Dawson MA, Foster SD, Bannister AJ, Robson SC, Hannah R, Wang XN, Xhemalce B, Wood AD, Green AR, Gottgens B, Kouzarides T. Cell Rep. 2012; 2:470. [PubMed: 22999934] (b) Dawson MA, Bannister AJ, Gottgens B, Foster SD, Bartke T, Green AR, Kouzarides T. Nature. 2009; 461:819. [PubMed: 19783980]
199. Hurd PJ, Bannister AJ, Halls K, Dawson MA, Vermeulen M, Olsen JV, Ismail H, Somers J, Mann M, Owen-Hughes T, Gout I, Kouzarides T. J Biol Chem. 2009; 284:16575. [PubMed: 19363025]
200. Aslam A, Logie C. PLoS One. 2010; 5:e10851. [PubMed: 20520775]
201. Kim K, Kim JM, Kim JS, Choi J, Lee YS, Neamati N, Song JS, Heo K, An W. J Mol Cell. 2013; 52:459.
202. Ajiro K. J Biol Chem. 2000; 275:439. [PubMed: 10617636]
203. Freitas MA, Sklenar AR, Parthun MR. J Cell Biochem. 2004; 92:691. [PubMed: 15211567]
204. Lau ATY, Lee SY, Xu YM, Zheng D, Cho YY, Zhu F, Kim HG, Li SQ, Zhang ZG, Bode AM, Dong ZG. J Biol Chem. 2011; 286:26628. [PubMed: 21646345]
205. (a) Ruizcarrillo A, Wangh LJ, Allfrey VG. Science. 1975; 190:117. [PubMed: 1166303] (b) Barber CM, Turner FB, Wang YM, Hagstrom K, Taverna SD, Mollah S, Ueberheide B, Meyer BJ, Hunt DF, Cheung P, Allis CD. Chromosoma. 2004; 112:360. [PubMed: 15133681]
206. Bungard D, Fuerth BJ, Zeng PY, Faubert B, Maas NL, Viollet B, Carling D, Thompson CB, Jones RG, Berger SL. Science. 2010; 329:1201. [PubMed: 20647423]
207. Zhang H, Wang Z, Zhang Z. Nucleic Acids Res. 2013; 41:8085. [PubMed: 23828041]
208. (a) Garcia BA, Hake SB, Diaz RL, Kauer M, Morris SA, Recht J, Shabanowitz J, Mishra N, Strahl BD, Allis CD, Hunt DF. J Biol Chem. 2007; 282:7641. [PubMed: 17194708] (b) Guillemette B, Drogaris P, Lin HHS, Armstrong H, Hiragami-Hamada K, Imhof A, Bonneil E, Thibault P, Verreault A, Festenstein RJ. PLoS Genet. 2011; 7:e1001354. [PubMed: 21483810]
209. Zhang LW, Eugeni EE, Parthun MR, Freitas MA. Chromosoma. 2003; 112:77. [PubMed: 12937907]
210. (a) Spencer TE, Jenster G, Burcin MM, Allis CD, Zhou JX, Mizzen CA, McKenna NJ, Onate SA, Tsai SY, Tsai MJ, OMalley BW. Nature. 1997; 389:194. [PubMed: 9296499] (b) Strasak L, Bartova E, Harnicarova A, Galiova G, Krejci J, Kozubek S. J Cell Physiol. 2009; 220:91. [PubMed: 19248079]
211. Britton LMP, Newhart A, Bhanu NV, Sridharan R, Gonzales-Cope M, Plath K, Janicki SM, Garcia BA. Epigenetics. 2013; 8:1101. [PubMed: 23949383]
212. Gatta R, Dolfini D, Zambelli F, Imbriano C, Pavesi G, Mantovani R. Epigenetics. 2011; 6:630. [PubMed: 21739721]
213. (a) Schiltz RL, Mizzen CA, Vassilev A, Cook RG, Allis CD, Nakatani Y. J Biol Chem. 1999; 274:1189. [PubMed: 9880483] (b) Mishima Y, Miyagi S, Saraya A, Negishi M, Endoh M, Endo TA, Toyoda T, Shinga J, Katsumoto T, Chiba T, Yamaguchi N, Kitabayashi I, Koseki H, Iwama A. Blood. 2011; 118:2443. [PubMed: 21753189]
214. (a) Yamamoto T, Horikoshi M. J Biol Chem. 1997; 272:30595. [PubMed: 9388189] (b) Poplawski A, Hu KF, Lee W, Natesan S, Peng DN, Carlson S, Shi XB, Balaz S, Markley JL, Glass KC. J Mol Biol. 2014; 426:1661. [PubMed: 24333487]
215. Jin QH, Yu LR, Wang LF, Zhang ZJ, Kasper LH, Lee JE, Wang CC, Brindle PK, Dent SYR, Ge K. EMBO J. 2011; 30:249. [PubMed: 21131905]
216. Basu A, Rose KL, Zhang JM, Beavis RC, Ueberheide B, Garcia BA, Chait B, Zhao YM, Hunt DF, Segal E, Allis CD, Hake SB. Proc Natl Acad Sci USA. 2009; 106:13785. [PubMed: 19666589]
217. Kamieniarz K, Izzo A, Dundr M, Tropberger P, Ozretic L, Kirfel J, Scheer E, Tropel P, Wisniewski JR, Tora L, Viville S, Buettner R, Schneider R. Genes Dev. 2012; 26:797. [PubMed: 22465951]
218. Daujat S, Bauer UM, Shah V, Turner B, Berger S, Kouzarides T. Curr Biol. 2002; 12:2090. [PubMed: 12498683]

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

219. Creyghton MP, Cheng AW, Welstead GG, Kooistra T, Carey BW, Steine EJ, Hanna J, Lodato MA, Frampton GM, Sharp PA, Boyer LA, Young RA, Jaenisch R. Proc Natl Acad Sci USA. 2010; 107:21931. [PubMed: 21106759]
220. Morris SA, Rao B, Garcia BA, Hake SB, Diaz RL, Shabanowitz J, Hil DF, Allis CD, Lieb JD, Strahl BD. J Biol Chem. 2007; 282:7632. [PubMed: 17189264]
221. Martinez-Redondo P, Vaquero A. Genes Cancer. 2013; 4:148. [PubMed: 24020006]
222. Singh SK, Maeda K, Eid MMA, Almofty SA, Ono M, Pham P, Goodman MF, Sakaguchi N. Nat Commun. 2013; 410.1038/ncomms2823
223. Tjeertes JV, Miller KM, Jackson SP. EMBO J. 2009; 28:1878. [PubMed: 19407812]
224. Puerta C, Hernandez F, Lopezalarcon L, Palacian E. Biochem Biophys Res Commun. 1995; 210:409. [PubMed: 7755616]
225. Di Cerbo V, Mohn F, Ryan DP, Montellier E, Kacem S, Tropberger P, Kallis E, Holzner M, Hoerner L, Feldmann A, Richter FM, Bannister AJ, Mittler G, Michaelis J, Khochbin S, Feil R, Schuebeler D, Owen-Hughes T, Daujat S, Schneider R. eLife. 2014; 3
226. van Bavel CC, Dieker J, Muller S, Briand JP, Monestier M, Berden JH, van der Vlag J. Mol Immunol. 2009; 47:511. [PubMed: 19747733]
227. Gatta R, Mantovani R. Epigenetics. 2011; 6:527.
228. (a) Zhao R, Nakamura T, Fu Y, Lazar Z, Spector DL. Nat Cell Biol. 2011; 13:1295. [PubMed: 21983563] (b) Bhaskara S, Knutson SK, Jiang GC, Chandrasekharan MB, Wilson AJ, Zheng SY, Yenamandra A, Locke K, Yuan JL, Bonine-Summers AR, Wells CE, Kaiser JF, Washington MK, Zhao ZM, Wagner FF, Sun ZW, Xia F, Holson EB, Khabele D, Hiebert SW. Cancer Cell. 2010; 18:436. [PubMed: 21075309]
229. (a) Jeppesen P, Turner BM. Cell. 1993; 74:281. [PubMed: 8343956] (b) Elsheikh SE, Green AR, Rakha EA, Powe DG, Ahmed RA, Collins HM, Soria D, Garibaldi JM, Paish CE, Ammar AA, Grainge MJ, Ball GR, Abdelghany MK, Martinez-Pomares L, Heery DM, Ellis IO. Cancer Res. 2009; 69:3802. [PubMed: 19366799]
230. (a) Song OK, Wang XR, Waterborg JH, Sternglanz R. J Biol Chem. 2003; 278:38109. [PubMed: 12915400] (b) Evjenth R, Hole K, Karlsen OA, Ziegler M, Arnesen T, Lillehaug JR. J Biol Chem. 2009; 284:31122. [PubMed: 19744929] (c) Hole K, Van Damme P, Dalva M, Aksnes H, Glomnes N, Varhaug JE, Lillehaug JR, Gevaert K, Arnesen T. PLoS One. 2011; 6:e24713. [PubMed: 21935442]
231. Kapoor-Vazirani P, Kagey JD, Powell DR, Vertino PM. Cancer Res. 2008; 68:6810. [PubMed: 18701507]
232. (a) Wang HB, Cao R, Xia L, Erdjument-Bromage H, Borchers C, Tempst P, Zhang Y. Mol Cell. 2001; 8:1207. [PubMed: 11779497] (b) Nishioka K, Chuikov S, Sarma K, Erdjument-Bromage H, Allis CD, Tempst P, Reinberg D. Genes Dev. 2002; 16:479. [PubMed: 11850410]
233. Popova EY, Xu XM, DeWan AT, Salzberg AC, Berg A, Hoh J, Zhang SS, Barnstable CJ. PLoS One. 2012; 7:e46867. [PubMed: 23056497]
234. Okitsu CY, Hsieh JCF, Hsieh CL. Mol Cell Biol. 2010; 30:2933. [PubMed: 20404096]
235. Van Aller GS, Reynoird N, Barbash O, Huddleston M, Liu SC, Zmoos AF, McDevitt P, Sinnamon R, Le BC, Mas G, Annan R, Sage J, Garcia BA, Tummino PJ, Gozani O, Kruger RG. Epigenetics. 2012; 7:340. [PubMed: 22419068]
236. Tachibana M, Sugimoto K, Fukushima T, Shinkai Y. J Biol Chem. 2001; 276:25309. [PubMed: 11316813]
237. Wen B, Wu H, Shinkai Y, Irizarry RA, Feinberg AP. Nat Genet. 2009; 41:246. [PubMed: 19151716]
238. (a) Nielsen SJ, Schneider R, Bauer UM, Bannister AJ, Morrison A, O'Carroll D, Firestein R, Cleary M, Jenuwein T, Herrera RE, Kouzarides T. Nature. 2001; 412:561. [PubMed: 11484059] (b) O'Carroll D, Scherthan H, Peters AHFM, Opravil S, Haynes AR, Laible G, Rea S, Schmid M, Lebersorger A, Jerratsch M, Sattler L, Mattei MG, Denny P, Brown SDM, Schweizer D, Jenuwein T. Mol Cell Biol. 2000; 20:9423. [PubMed: 11094092]
239. Schotta G, Lachner M, Sarma K, Ebert A, Sengupta R, Reuter G, Reinberg D, Jenuwein T. Genes Dev. 2004; 18:1251. [PubMed: 15145825]

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

240. Garcia BA, Busby SA, Shabanowitz J, Hunt DF, Mishra N. J Proteome Res. 2005; 4:2032. [PubMed: 16335948]

241. Rogenhofer S, Kahl P, Mertens C, Hauser S, Hartmann W, Buttner R, Muller SC, von Ruecker A, Ellinger J. BJU Int. 2012; 109:459. [PubMed: 21810159]

242. Daujat S, Zeissler U, Waldmann T, Happel N, Schneider R. J Biol Chem. 2005; 280:38090. [PubMed: 16127177]

243. Gnesutta N, Nardini M, Mantovani R. Transcription. 2013; 4:114. [PubMed: 23756340]

244. Chavez JD, Weisbrod CR, Zheng CX, Eng JK, Bruce JE. Mol Cell Proteomics. 2013; 12:1451. [PubMed: 23354917]

245. Yu YX, Song CY, Zhang QY, DiMaggio PA, Garcia BA, York A, Carey MF, Grunstein M. Mol Cell. 2012; 46:7. [PubMed: 22387026]

246. Jack APM, Bussemer S, Hahn M, Punzeler S, Snyder M, Wells M, Csankovszki G, Solovei I, Schotta G, Hake SB. PLoS One. 2013; 8:e51765. [PubMed: 23451023]

247. Hake SB, Garcia BA, Duncan EM, Kauer M, Dellaire G, Shabanowitz J, Bazett-Jones DP, Allis CD, Hunt DF. J Biol Chem. 2006; 281:559. [PubMed: 16267050]

248. Daujat S, Weiss T, Mohn F, Lange UC, Ziegler-Birling C, Zeissler U, Lappe M, Schubeler D, Torres-Padilla ME, Schneider R. Nat Struct Mol Biol. 2009; 16:777. [PubMed: 19561610]

249. Frederiks F, Tzouros M, Oudgenoeg G, van Welsem T, Fornerod M, Krijgsveld J, van Leeuwen F. Nat Struct Mol Biol. 2008; 15:550. [PubMed: 18511943]

250. Weiss T, Hergeth S, Zeissler U, Izzo A, Tropberger P, Zee BM, Dundr M, Garcia BA, Daujat S, Schneider R. Epigenet Chromatin. 2010; 3

251. Lacoste N, Utley RT, Hunter JM, Poirier GG, Cote J. J Biol Chem. 2002; 277:30421. [PubMed: 12097318]

252. Tessarz P, Santos-Rosa H, Robson S, Sylvestersen KB, Nelson CJ, Nielsen ML, Kouzarides T. Nature. 2014; 505:564. [PubMed: 24352239]

253. Kirmizis A, Santos-Rosa H, Penkett CJ, Singer MA, Green RD, Kouzarides T. Nat Struct Mol Biol. 2009; 16:449. [PubMed: 19270702]

254. (a) Guccione E, Bassi C, Casadio F, Martinato F, Cesaroni M, Schuchlautz H, Luscher B, Amati B. Nature. 2007; 449:933. [PubMed: 17898714] (b) Guccione E, Martinato F, Finocchiaro G, Luzi L, Tizzoni L, Dall'Olio V, Zardo G, Nervi C, Bernard L, Amati B. Nat Cell Biol. 2006; 8:764. [PubMed: 16767079]

255. Hyllus D, Stein C, Schnabel K, Schiltz E, Imhof A, Dou YL, Hsieh J, Bauer UM. Genes Dev. 2007; 21:3369. [PubMed: 18079182]

256. Waldmann T, Izzo A, Kamieniarz K, Richter F, Vogler C, Sarg B, Lindner H, Young NL, Mittler G, Garcia BA, Schneider R. Epigenet Chromatin. 2011; 410.1186/1756

257. Migliori V, Muller J, Phalke S, Low D, Bezzi M, Mok WC, Sahu SK, Gunaratne J, Capasso P, Bassi C, Cecatiello V, De Marco A, Blackstock W, Kuznetsov V, Amati B, Mapelli M, Guccione E. Nat Struct Mol Biol. 2012; 19:136. [PubMed: 22231400]

258. Pal S, Vishwanath SN, Erdjument-Bromage H, Tempst P, Sif S. Mol Cell Biol. 2004; 24:9630. [PubMed: 15485929]

259. Guo A, Gu H, Zhou J, Mulhern D, Wang Y, Lee KA, Yang V, Aguiar M, Kornhauser J, Jia X, Ren J, Beausoleil SA, Silva JC, Vemulapalli V, Bedford MT, Comb MJ. Mol Cell Proteomics. 2014; 13:372. [PubMed: 24129315]

260. Blythe SA, Cha SW, Tadjuidje E, Heasman J, Klein PS. Dev Cell. 2010; 19:220. [PubMed: 20708585]

261. (a) Yang YZ, Lu Y, Espejo A, Wu JC, Xu W, Liang SD, Bedford MT. Mol Cell. 2010; 40:1016. [PubMed: 21172665] (b) Schurter BT, Koh SS, Chen D, Bunick GJ, Harp JM, Hanson BL, Henschen-Edman A, Mackay DR, Stallcup MR, Aswad DW. Biochemistry. 2001; 40:5747. [PubMed: 11341840]

262. Sylvestersen KB, Horn H, Jungmichel S, Jensen LJ, Nielsen ML. Mol Cell Proteomics. 2014:2072. [PubMed: 24563534]

263. Casadio F, Lu XD, Pollock SB, LeRoy G, Garcia BA, Muir TW, Roeder RG, Allis CD. Proc Natl Acad Sci USA. 2013; 110:14894. [PubMed: 23980157]

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

264. Bremang M, Cuomo A, Agresta AM, Stugiewicz M, Spadotto V, Bonaldi T. Mol BioSyst. 2013; 9:2231. [PubMed: 23748837]
265. Liu B, Lin YH, Darwanto A, Song XH, Xu GL, Zhang KL. J Biol Chem. 2009; 284:32288. [PubMed: 19801601]
266. Zhang K, Chen Y, Mang ZH, Zhao YM. J Proteome Res. 2009; 8:900. [PubMed: 19113941]
267. Montellier E, Boussouar F, Rousseaux S, Zhang K, Buchou T, Fenaille F, Shiota H, Debernardi A, Hery P, Curtet S, Jamshidikia M, Barral S, Holota H, Bergon A, Lopez F, Guardiola P, Pernet K, Imbert J, Petosa C, Tan MJ, Zhao YM, Gerard M, Khochbin S. Genes Dev. 2013; 27:1680. [PubMed: 23884607]
268. Weinert BT, Scholz C, Wagner SA, Iesmantavicius V, Su D, Daniel JA, Choudhary C. Cell Rep. 2013; 4:842. [PubMed: 23954790]
269. Wisniewski JR, Zougman A, Mann M. Nucleic Acids Res. 2008; 36:570. [PubMed: 18056081]
270. Kim W, Bennett EJ, Huttlin EL, Guo A, Li J, Possemato A, Sowa ME, Rad R, Rush J, Comb MJ, Harper JW, Gygi SP. Mol Cell. 2011; 44:325. [PubMed: 21906983]
271. Rose NR, Klose RJ. Biochim Biophys Acta. 2014:1362. [PubMed: 24560929]
272. (a) Nickel BE, Davie JR. Biochemistry. 1989; 28:964. [PubMed: 2540826] (b) Kallin EM, Cao R, Jothi R, Xia K, Cui KR, Zhao KJ, Zhang Y. PLoS Genet. 2009; 5:e1000506. [PubMed: 19503595]
273. Udeshi ND, Svinkina T, Mertins P, Kuhn E, Mani DR, Qiao JW, Carr SA. Mol Cell Proteomics. 2013; 12:825. [PubMed: 23266961]
274. Wu LP, Zee BM, Wang YM, Garcia BA, Dou YL. Mol Cell. 2011; 43:132. [PubMed: 21726816]
275. Han J, Zhang H, Zhang H, Wang Z, Zhou H, Zhang Z. Cell. 2013; 155:817. [PubMed: 24209620]
276. (a) Henry KW, Wyce A, Lo WS, Duggan LJ, Emre NCT, Kao CF, Pillus L, Shilatifard A, Osley MA, Berger SL. Genes Dev. 2003; 17:2648. [PubMed: 14563679] (b) Chen S, Li J, Wang DL, Sun FL. Cell Res. 2012; 22:1402. [PubMed: 22847742]
277. (a) Gatti M, Pinato S, Maspero E, Soffientini P, Polo S, Penengo L. Cell Cycle. 2012; 11:2538. [PubMed: 22713238] (b) Fradet-Turcotte A, Canny MD, Escribano-Diaz C, Orthwein A, Leung CCY, Huang H, Landry MC, Kitevski-LeBlanc J, Noordermeer SM, Sicheri F, Durocher D. Nature. 2013; 499:50. [PubMed: 23760478]
278. (a) Wang Y, Wysocka J, Sayegh J, Lee YH, Perlin JR, Leonelli L, Sonbuchner LS, McDonald CH, Cook RG, Dou Y, Roeder RG, Clarke S, Stallcup MR, Allis CD, Coonrod SA. Science. 2004; 306:279. [PubMed: 15345777] (b) Cuthbert GL, Daujat S, Snowden AW, Erdjument-Bromage H, Hagiwara T, Yamada M, Schneider R, Gregory PD, Tempst P, Bannister AJ, Kouzarides T. Cell. 2004; 118:545. [PubMed: 15339660]
279. Zhang XS, Bolt M, Guertin MJ, Chen W, Zhang S, Cherrington BD, Slade DJ, Dreyton CJ, Subramanian V, Bicker KL, Thompson PR, Mancini MA, Lis JT, Coonrod SA. Proc Natl Acad Sci USA. 2012; 109:13331. [PubMed: 22853951]
280. Wysocka J, Allis CD, Coonrod S. Front Biosci. 2006; 11:344. [PubMed: 16146736]
281. Hagiwara T, Hidaka Y, Yamada M. Biochemistry. 2005; 44:5827. [PubMed: 15823041]
282. Christophorou MA, Castelo-Branco G, Halley-Stott RP, Oliveira CS, Loos R, Radzisheuskaya A, Mowen KA, Bertone P, Silva JCR, Zernicka-Goetz M, Nielsen ML, Gurdon JB, Kouzarides T. Nature. 2014; 507:104. [PubMed: 24463520]
283. Fong JJ, Nguyen BL, Bridger R, Medrano EE, Wells L, Pan SJ, Sifers RN. J Biol Chem. 2012; 287:12195. [PubMed: 22371497]
284. Fujiki R, Hashiba W, Sekine H, Yokoyama A, Chikanishi T, Ito S, Imai Y, Kim J, He HH, Igarashi K, Kanno J, Ohtake F, Kitagawa H, Roeder RG, Brown M, Kato S. Nature. 2011; 480:557. [PubMed: 22121020]
285. Messner S, Altmeyer M, Zhao HT, Pozivil A, Roschitzki B, Gehrig P, Rutishauser D, Huang DZ, Caflisch A, Hottiger MO. Nucleic Acids Res. 2010; 38:6350. [PubMed: 20525793]
286. Ogata N, Ueda K, Hayaishi O. J Biol Chem. 1980; 255:7610. [PubMed: 7400135]
287. (a) Monneau YR, Soufari H, Nelson CJ, Mackereth CD. J Biol Chem. 2013; 288:25826. [PubMed: 23888048] (b) Howe FS, Boubriak I, Sale MJ, Nair A, Clynes D, Grijzenhout A, Murray SC, Woloszczuk R, Mellor J. Mol Cell. 2014; 55:733. [PubMed: 25127513]

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

288. Nelson CJ, Santos-Rosa H, Kouzarides T. Cell. 2006; 126:905. [PubMed: 16959570]
289. Hendriks IA, D'Souza RCJ, Yang B, Verlaan-de Vries M, Mann M, Vertegaal ACO. Nat Struct Mol Biol. 2014; 21:927. [PubMed: 25218447]
290. Nathan D, Ingvarsdottir K, Sterner DE, Bylebyl GR, Dokmanovic M, Dorsey LA, Whelan KA, Krsmanovic M, Lane WS, Meluh PB, Johnson ES, Berger SL. Genes Dev. 2006; 20:966. [PubMed: 16598039]
291. (a) Ruthenburg AJ, Allis CD, Wysocka J. Mol Cell. 2007; 25:15. [PubMed: 17218268] (b) Taverna SD, Li H, Ruthenburg AJ, Allis CD, Patel DJ. Nat Struct Mol Biol. 2007; 14:1025. [PubMed: 17984965]
292. Baker LA, Allis CD, Wang GG. Mutat Res. 2008; 647:3. [PubMed: 18682256]
293. James LI, Frye SV. Clin Pharmacol Ther. 2013; 93:312. [PubMed: 23403847]
294. (a) Yu XC, Chini CCS, He M, Mer G, Chen JJ. Science. 2003; 302:639. [PubMed: 14576433] (b) Muslin AJ, Tanner JW, Allen PM, Shaw AS. Cell. 1996; 84:889. [PubMed: 8601312] (c) Bottomley MJ. EMBO Rep. 2004; 5:464. [PubMed: 15184976] (d) Yun MY, Wu J, Workman JL, Li B. Cell Res. 2011; 21:564. [PubMed: 21423274] (e) Patel DJ, Wang ZX. Annu Rev Biochem. 2013; 82:81. [PubMed: 23642229]
295. (a) Mikkelsen TS, Ku M, Jaffe DB, Issac B, Lieberman E, Giannoukos G, Alvarez P, Brockman W, Kim TK, Koche RP, Lee W, Mendenhall E, O'Donovan A, Presser A, Russ C, Xie X, Meissner A, Wernig M, Jaenisch R, Nusbaum C, Lander ES, Bernstein BE. Nature. 2007; 448:553. [PubMed: 17603471] (b) Jiang L, Smith JN, Anderson SL, Ma P, Mizzen CA, Kelleher NL. J Biol Chem. 2007; 282:27923. [PubMed: 17652096] (c) Bernstein BE, Meissner A, Lander ES. Cell. 2007; 128:669. [PubMed: 17320505]
296. Li H, Ilin S, Wang W, Duncan EM, Wysocka J, Allis CD, Patel DJ. Nature. 2006; 442:91. [PubMed: 16728978]
297. Schreiber SL, Bernstein BE. Cell. 2002; 111:771. [PubMed: 12526804]
298. Mujtaba S, He Y, Zeng L, Yan S, Plotnikova O, Sachchidanand Sanchez R, Zeleznik-Le NJ, Ronai Z, Zhou MM. Mol Cell. 2004; 13:251. [PubMed: 14759370]
299. Chan DW, Wang Y, Wu M, Wong JM, Qin J, Zhao YM. Proteomics. 2009; 9:2343. [PubMed: 19337993]
300. Vermeulen M, Eberl HC, Matarese F, Marks H, Denissov S, Butter F, Lee KK, Olsen JV, Hyman AA, Stunnenberg HG, Mann M. Cell. 2010; 142:967. [PubMed: 20850016]
301. Vermeulen M, Mulder KW, Denissov S, Pijnappel WWMP, van Schaik FMA, Varier RA, Baltissen MPA, Stunnenberg HG, Mann M, Timmers HTM. Cell. 2007; 131:58. [PubMed: 17884155]
302. Eberl HC, Spruijt CG, Kelstrup CD, Vermeulen M, Mann M. Mol Cell. 2013; 49:368. [PubMed: 23201125]
303. Lang DN, Schumann M, Gelato K, Fischle W, Schwarzer D, Krause E. Proteomics. 2013; 13:2989. [PubMed: 23970329]
304. Bartke T, Vermeulen M, Xhemalce B, Robson SC, Mann M, Kouzarides T. Cell. 2010; 143:470. [PubMed: 21029866]
305. (a) Ruthenburg AJ, Li H, Milne TA, Dewell S, McGinty RK, Yuen M, Ueberheide B, Dou Y, Muir TW, Patel DJ, Allis CD. Cell. 2011; 145:692. [PubMed: 21596426] (b) Wang Z, Song J, Milne TA, Wang GG, Li H, Allis CD, Patel DJ. Cell. 2010; 141:1183. [PubMed: 20541251]
306. LeRoy G, Chepelev I, DiMaggio PA, Blanco MA, Zee BM, Zhao K, Garcia BA. Genome Biol. 2012; 13:R68. [PubMed: 22897906]
307. (a) Ogryzko VV, Schiltz RL, Russanova V, Howard BH, Nakatani Y. Cell. 1996; 87:953. [PubMed: 8945521] (b) Liu X, Wang L, Zhao KH, Thompson PR, Hwang Y, Marmorstein R, Cole PA. Nature. 2008; 451:846. [PubMed: 18273021] (c) Das C, Kundu TK. IUBMB Life. 2005; 57:137. [PubMed: 16036576]
308. Gu W, Shi XL, Roeder RG. Nature. 1997; 387:819. [PubMed: 9194564]
309. (a) Zhu H, Klemic JF, Chang S, Bertone P, Casamayor A, Klemic KG, Smith D, Gerstein M, Reed MA, Snyder M. Nat Genet. 2000; 26:283. [PubMed: 11062466] (b) Reimer U, Reineke U, Schneider-Mergener J. Curr Opin Biotechnol. 2002; 13:315. [PubMed: 12323352]

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

310. Rees B. Health Phys. 1996; 70:639. [PubMed: 8690573]
311. (a) Erce MA, Pang CN, Hart-Smith G, Wilkins MR. Proteomics. 2012; 12:564. [PubMed: 22246820] (b) Zhao Y, Jensen ON. Proteomics. 2009; 9:4632. [PubMed: 19743430]
312. Zee BM, Garcia BA. Anal Chem. 2010; 82:28. [PubMed: 19928949]
313. Zhang YJ, Wang JQ, Ding M, Yu YH. Nat Methods. 2013; 10:981. [PubMed: 23955771]
314. Dignam JD, Lebovitz RM, Roeder RG. Nucleic Acids Res. 1983; 11:1475. [PubMed: 6828386]
315. Motoyama A, Xu T, Ruse CI, Wohlschlegel JA, Yates JR. Anal Chem. 2007; 79:3623. [PubMed: 17411013]
316. Wang YX, Yang F, Gritsenko MA, Wang YC, Clauss T, Liu T, Shen YF, Monroe ME, Lopez-Ferrer D, Reno T, Moore RJ, Klemke RL, Camp DG, Smith RD. Proteomics. 2011; 11:2019. [PubMed: 21500348]
317. (a) Guan KL, Yu W, Lin Y, Xiong Y, Zhao SM. Nat Protoc. 2010; 5:1583. [PubMed: 21085124] (b) White DA, Belyaev ND, Turner BM. Methods-a Companion To Methods In Enzymology. 1999; 19:417.(c) Xu W, Zhao S. Methods Mol Biol. 2013; 981:137. [PubMed: 23381859]
318. Hebert AS, Dittenhafer-Reed KE, Yu W, Bailey DJ, Selen ES, Boersma MD, Carson JJ, Tonelli M, Balloon AJ, Higbee AJ, Westphall MS, Pagliarini DJ, Prolla TA, Assadi-Porter F, Roy S, Denu JM, Coon JJ. Mol Cell. 2013; 49:186. [PubMed: 23201123]
319. (a) Rush J, Moritz A, Lee KA, Guo A, Goss VL, Spek EJ, Zhang H, Zha XM, Polakiewicz RD, Comb MJ. Nat Biotechnol. 2005; 23:94. [PubMed: 15592455] (b) Ballif BA, Carey GR, Sunyaev SR, Gygi SP. J Proteome Res. 2008; 7:311. [PubMed: 18034455]
320. (a) Rardin MJ, He W, Nishida Y, Newman JC, Carrico C, Danielson SR, Guo A, Gut P, Sahu AK, Li B, Uppala R, Fitch M, Riiff T, Zhu L, Zhou J, Mulhern D, Stevens RD, Ilkayeva OR, Newgard CB, Jacobson MP, Hellerstein M, Goetzman ES, Gibson BW, Verdin E. Cell Metab. 2013; 18:920. [PubMed: 24315375] (b) Colak G, Xie ZY, Zhu AY, Dai LZ, Lu ZK, Zhang Y, Wan XL, Chen Y, Cha YH, Lin HN, Zhao YM, Tan MJ. Mol Cell Proteomics. 2013; 12:3509. [PubMed: 24176774]
321. (a) Wu Z, Cheng Z, Sun M, Wan X, Liu P, Huang M, He T, Geng M, Ding J, Tan M, Zhao Y. Mol Cell Proteomics. 2014 accepted. (b) Cao XJ, Arnaudo AM, Garcia BA. Epigenetics. 2013; 8:477. [PubMed: 23644510] (c) Moore KE, Carlson SM, Camp ND, Cheung P, James RG, Chua KF, Wolf-Yadlin A, Gozani O. Mol Cell. 2013; 50:444. [PubMed: 23583077]
322. (a) Udeshi ND, Mertins P, Svinkina T, Carr SA. Nat Protoc. 2013; 8:1950. [PubMed: 24051958] (b) Sarraf SA, Raman M, Guarani-Pereira V, Sowa ME, Huttlin EL, Gygi SP, Harper JW. Nature. 2013; 496:372. [PubMed: 23503661]
323. Xu GQ, Paige JS, Jaffrey SR. Nat Biotechnol. 2010; 28:868. [PubMed: 20639865]
324. Chen Y, Colak G, Zhao YM. Sirtuins: Methods Protoc. 2013; 1077:105.
325. (a) Ficarro SB, McCleland ML, Stukenberg PT, Burke DJ, Ross MM, Shabanowitz J, Hunt DF, White FM. Nat Biotechnol. 2002; 20:301. [PubMed: 11875433] (b) Posewitz MC, Tempst P. Anal Chem. 1999; 71:2883. [PubMed: 10424175] (c) Wolschin F, Wienkoop S, Weckwerth W. Proteomics. 2005; 5:4389. [PubMed: 16222723] (d) Feng S, Ye M, Zhou H, Jiang X, Jiang X, Zou H, Gong B. Mol Cell Proteomics. 2007; 6:1656. [PubMed: 17575324]
326. (a) Pinkse MW, Uitto PM, Hilhorst MJ, Ooms B, Heck AJ. Anal Chem. 2004; 76:3935. [PubMed: 15253627] (b) Thingholm TE, Larsen MR, Ingrell CR, Kassem M, Jensen ON. J Proteome Res. 2008; 7:3304. [PubMed: 18578522] (c) Beckers GJ, Hoehenwarter W, Rohrig H, Conrath U, Weckwerth W. Methods Mol Biol. 2014; 1072:621. [PubMed: 24136551] (d) Kweon HK, Hakansson K. Anal Chem. 2006; 78:1743. [PubMed: 16536406] (e) Li Y, Lin HQ, Deng CH, Yang PY, Zhang XM. Proteomics. 2008; 8:238. [PubMed: 18081192]
327. (a) Larsen MR, Thingholm TE, Jensen ON, Roepstorff P, Jorgensen TJ. Mol Cell Proteomics. 2005; 4:873. [PubMed: 15858219] (b) Sugiyama N, Masuda T, Shinoda K, Nakamura A, Tomita M, Ishihama Y. Mol Cell Proteomics. 2007; 6:1103. [PubMed: 17322306] (c) Aryal UK, Ross ARS. Rapid Commun Mass Spectrom. 2010; 24:219. [PubMed: 20014058]
328. Villen J, Gygi SP. Nat Protoc. 2008; 3:1630. [PubMed: 18833199]
329. (a) Song CX, Ye ML, Han GH, Jiang XN, Wang FJ, Yu ZY, Chen R, Zou HF. Anal Chem. 2010; 82:53. [PubMed: 19950968] (b) Ficarro SB, Zhang Y, Carrasco-Alfonso MJ, Garg B, Adelmant

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

G, Webber JT, Luckey CJ, Marto JA. Mol Cell Proteomics. 2011; 10:O111.011064. [PubMed: 21788404]

330. (a) McNulty DE, Annan RS. Mol Cell Proteomics. 2008; 7:971. [PubMed: 18212344] (b) McNulty, DE.; Huddleston, MJ.; Annan, RS. Sample Preparation in Biological Mass Spectrometry. Springer; Netherlands: 2011.
331. Gan CS, Guo T, Zhang H, Lim SK, Sze SK. J Proteome Res. 2008; 7:4869. [PubMed: 18828627]
332. (a) Beausoleil SA, Jedrychowski M, Schwartz D, Elias JE, Villen J, Li JX, Cohn MA, Cantley LC, Gygi SP. Proc Natl Acad Sci USA. 2004; 101:12130. [PubMed: 15302935] (b) Huttlin EL, Jedrychowski MP, Elias JE, Goswami T, Rad R, Beausoleil SA, Villen J, Haas W, Sowa ME, Gygi SP. Cell. 2010; 143:1174. [PubMed: 21183079]
333. (a) Xu Y, Sprung R, Kwon SW, Kim SC, Zhao Y. J Proteome Res. 2007; 6:1153. [PubMed: 17330948] (b) Cantin GT, Venable JD, Cociorva D, Yates JR III. J Proteome Res. 2006; 5:127. [PubMed: 16396503] (c) Hung CW, Kubler D, Lehmann WD. Electrophoresis. 2007; 28:2044. [PubMed: 17523138] (d) Xu CF, Wang HB, Li DM, Kong XP, Neubert TA. Anal Chem. 2007; 79:2007. [PubMed: 17249638]
334. Thingholm TE, Jensen ON, Robinson PJ, Larsen MR. Mol Cell Proteomics. 2008; 7:661. [PubMed: 18039691]
335. (a) Xia Q, Cheng D, Duong DM, Gearing M, Lah JJ, Levey AI, Peng J. J Proteome Res. 2008; 7:2845. [PubMed: 18510355] (b) Ruse CI, McClatchy DB, Lu B, Cociorva D, Motoyama A, Park SK, Yates JR III. J Proteome Res. 2008; 7:2140. [PubMed: 18452278]
336. Mamone G, Picariello G, Ferranti P, Addeo F. Proteomics. 2010; 10:380. [PubMed: 19953538]
337. Raasi S, Varadan R, Fushman D, Pickart CM. Nat Struct Mol Biol. 2005; 12:708. [PubMed: 16007098]
338. Hjerpe R, Aillet F, Lopitz-Otsoa F, Lang V, England P, Rodriguez MS. EMBO Rep. 2009; 10:1250. [PubMed: 19798103]
339. (a) Vosseller K, Trinidad JC, Chalkley RJ, Specht CG, Thalhammer A, Lynn AJ, Snedecor JO, Guan SH, Medzihradszky KF, Maltby DA, Schoepfer R, Burlingame AL. Mol Cell Proteomics. 2006; 5:923. [PubMed: 16452088] (b) Zielinska DF, Gnad F, Wisniewski JR, Mann M. Cell. 2010; 141:897. [PubMed: 20510933] (c) Kim JY, Kim SK, Kang D, Moon MH. Anal Chem. 2012; 84:5343. [PubMed: 22616828] (d) Ito S, Hayama K, Hirabayashi J. Methods Mol Biol. 2009; 534:195. [PubMed: 19277551]
340. (a) Dania N, Stilla A, Marchegiani A, Tamburro A, Till S, Ladurner AG, Corda D, Di Girolamo M. Proc Natl Acad Sci USA. 2009; 106:4243. [PubMed: 19246377] (b) Kaji H, Saito H, Yamauchi Y, Shinkawa T, Taoka M, Hirabayashi J, Kasai K, Takahashi N, Isobe T. Nat Biotechnol. 2003; 21:667. [PubMed: 12754521]
341. (a) Vocadlo DJ, Hang HC, Kim EJ, Hanover JA, Bertozzi CR. Proc Natl Acad Sci USA. 2003; 100:9116. [PubMed: 12874386] (b) Baskin JM, Prescher JA, Laughlin ST, Agard NJ, Chang PV, Miller IA, Lo A, Codelli JA, Bertozzi CR. Proc Natl Acad Sci USA. 2007; 104:16793. [PubMed: 17942682] (c) Kho Y, Kim SC, Jiang C, Barma D, Kwon SW, Cheng J, Jaunbergs J, Weinbaum C, Tamanoi F, Falck J, Zhao Y. Proc Natl Acad Sci USA. 2004; 101:12479. [PubMed: 15308774] (d) Sprung R, Nandi A, Chen Y, Kim SC, Barma D, Falck JR, Zhao YM. J Proteome Res. 2005; 4:950. [PubMed: 15952742] (e) Grammel M, Hang HC. Nat Chem Biol. 2013; 9:475. [PubMed: 23868317] (f) Yang YY, Grammel M, Raghavan AS, Charron G, Hang HC. Chem Biol. 2010; 17:1212. [PubMed: 21095571]
342. Demko ZP, Sharpless KB. Angew Chem, Int Ed. 2002; 41:2110.
343. Bayer EA, Wilchek M. J Chromatogr. 1990; 510:3. [PubMed: 2205618]
344. (a) Khidekel N, Ficarro SB, Peters EC, Hsieh-Wilson LC. Proc Natl Acad Sci USA. 2004; 101:13132. [PubMed: 15340146] (b) Rexach JE, Clark PM, Hsieh-Wilson LC. Nat Chem Biol. 2008; 4:97. [PubMed: 18202679] (c) Khidekel N, Ficarro SB, Clark PM, Bryan MC, Swaney DL, Rexach JE, Sun YE, Coon JJ, Peters EC, Hsieh-Wilson LC. Nat Chem Biol. 2007; 3:339. [PubMed: 17496889]
345. (a) Tian Y, Zhou Y, Elliott S, Aebersold R, Zhang H. Nat Protoc. 2007; 2:334. [PubMed: 17406594] (b) Zhang H, Li XJ, Martin DB, Aebersold R. Nat Biotechnol. 2003; 21:660.

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

[PubMed: 12754519] (c) Ramachandran P, Boontheung P, Xie Y, Sondej M, Wong DT, Loo JA. J Proteome Res. 2006; 5:1493. [PubMed: 16740002]

346. Wells L, Vosseller K, Cole RN, Cronshaw JM, Matunis MJ, Hart GW. Mol Cell Proteomics. 2002; 1:791. [PubMed: 12438562]
347. Chen WX, Smeekens JM, Wu RH. Mol Cell Proteomics. 2014; 13:1563. [PubMed: 24692641]
348. (a) Guo J, Gaffrey MJ, Su D, Liu T, Camp DG, Smith RD, Qian WJ. Nat Protoc. 2014; 9:64. [PubMed: 24336471] (b) Forrester MT, Thompson JW, Foster MW, Nogueira L, Moseley MA, Stamler JS. Nat Biotechnol. 2009; 27:557. [PubMed: 19483679]
349. Fekete S, Olah E, Fekete J. J Chromatogr, A. 2012; 1228:57. [PubMed: 21982449]
350. Chen Y, Kwon SW, Kim SC, Zhao Y. J Proteome Res. 2005; 4:998. [PubMed: 15952748]
351. Blagoev B, Ong SE, Kratchmarova I, Mann M. Nat Biotechnol. 2004; 22:1139. [PubMed: 15314609]
352. Kratchmarova I, Blagoev B, Haack-Sorensen M, Kassem M, Mann M. Science. 2005; 308:1472. [PubMed: 15933201]
353. Matsuoka S, Ballif BA, Smogorzewska A, McDonald ER 3rd, Hurov KE, Luo J, Bakalarski CE, Zhao Z, Solimini N, Lerenthal Y, Shiloh Y, Gygi SP, Elledge SJ. Science. 2007; 316:1160. [PubMed: 17525332]
354. (a) Hsu PP, Kang SA, Rameseder J, Zhang Y, Ottina KA, Lim D, Peterson TR, Choi Y, Gray NS, Yaffe MB, Marto JA, Sabatini DM. Science. 2011; 332:1317. [PubMed: 21659604] (b) Yu YH, Yoon SO, Poulogiannis G, Yang Q, Ma XJM, Villen J, Kubica N, Hoffman GR, Cantley LC, Gygi SP, Blenis J. Science. 2011; 332:1322. [PubMed: 21659605]
355. Wagner GR, Payne RM. J Biol Chem. 2013; 288:29036. [PubMed: 23946487]
356. Teo CF, Ingale S, Wolfert MA, Elsayed GA, Not LG, Chatham JC, Wells L, Boons GJ. Nat Chem Biol. 2010; 6:338. [PubMed: 20305658]
357. Sparbier K, Wenzel T, Kostrzewa M. J Chromatogr, B. 2006; 840:29.
358. Mayya V, Lundgren DH, Hwang SI, Rezaul K, Wu LF, Eng JK, Rodionov V, Han DK. Sci Signaling. 2009; 2:ra46.
359. Villen J, Beausoleil SA, Gerber SA, Gygi SP. Proc Natl Acad Sci USA. 2007; 104:1488. [PubMed: 17242355]
360. Iesmantavicius V, Weinert BT, Choudhary C. Mol Cell Proteomics. 2014:1979. [PubMed: 24961812]
361. Mertins P, Yang F, Liu T, Mani DR, Petyuk VA, Gillette MA, Clauser KR, Qiao JW, Gritsenko MA, Moore RJ, Levine DA, Townsend R, Erdmann-Gilmore P, Snider JE, Davies SR, Ruggles KV, Fenyo D, Kitchens RT, Li S, Olvera N, Dao F, Rodriguez H, Chan DW, Liebler D, White F, Rodland KD, Mills GB, Smith RD, Paulovich AG, Ellis M, Carr SA. Mol Cell Proteomics. 2014; 13:1690. [PubMed: 24719451]
362. Zawadzka AM, Schilling B, Cusack MP, Sahu AK, Drake P, Fisher SJ, Benz CC, Gibson BW. Mol Cell Proteomics. 2014; 13:1034. [PubMed: 24505115]
363. Trinidad JC, Barkan DT, Gulledge BF, Thalhammer A, Sali A, Schoepfer R, Burlingame AL. Mol Cell Proteomics. 2012; 11:215. [PubMed: 22645316]
364. Weinert BT, Wagner SA, Horn H, Henriksen P, Liu WSR, Olsen JV, Jensen LJ, Choudhary C. Sci Signaling. 2011; 4:ra48.
365. Rardin MJ, Newman JC, Held JM, Cusack MP, Sorensen DJ, Li BA, Schilling B, Mooney SD, Kahn CR, Verdin E, Gibson BW. Proc Natl Acad Sci USA. 2013; 110:6601. [PubMed: 23576753]
366. Henriksen P, Wagner SA, Weinert BT, Sharma S, Bacinskaja G, Rehman M, Juffer AH, Walther TC, Lisby M, Choudhary C. Mol Cell Proteomics. 2012; 11:1510. [PubMed: 22865919]
367. Still AJ, Floyd BJ, Hebert AS, Bingman CA, Carson JJ, Gunderson DR, Dolan BK, Grimsrud PA, Dittenhafer-Reed KE, Stapleton DS, Keller MP, Westphall MS, Denu JM, Attie AD, Coon JJ, Pagliarini DJ. J Biol Chem. 2013; 288:26209. [PubMed: 23864654]
368. Lundby A, Lage K, Weinert BT, Bekker-Jensen DB, Secher A, Skovgaard T, Kelstrup CD, Dmytriyev A, Choudhary C, Lundby C, Olsen JV. Cell Rep. 2012; 2:419. [PubMed: 22902405]

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

369. Hahne H, Sobotzki N, Nyberg T, Helm D, Borodkin VS, van Aalten DM, Agnew B, Kuster B. J Proteome Res. 2013; 12:927. [PubMed: 23301498]

370. Alfaro JF, Gong CX, Monroe ME, Aldrich JT, Clauss TR, Purvine SO, Wang Z, Camp DG II, Shabanowitz J, Stanley P, Hart GW, Hunt DF, Yang F, Smith RD. Proc Natl Acad Sci USA. 2012; 109:7280. [PubMed: 22517741]

371. Wang Z, Gucek M, Hart GW. Proc Natl Acad Sci USA. 2008; 105:13793. [PubMed: 18779572]

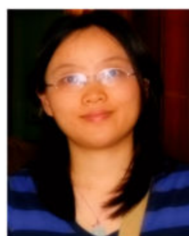
372. Nandi A, Sprung R, Barma DK, Zhao YX, Kim SC, Falck JR, Zhao YM. Anal Chem. 2006; 78:452. [PubMed: 16408927]

373. Du J, Patel DJ. Biochim Biophys Acta. 2014; 1839:719. [PubMed: 24747177]

374. (a) Fierz B, Muir TW. Nat Chem Biol. 2012; 8:417. [PubMed: 22510649] (b) Chatterjee C, Muir TW. J Biol Chem. 2010; 285:11045. [PubMed: 20147749] (c) Kim CH, Kang M, Kim HJ, Chatterjee A, Schultz PG. Angew Chem, Int Ed. 2012; 51:7246.

Biographies

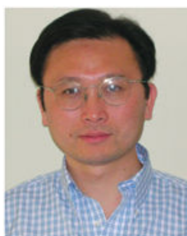
He Huang studied chemistry at Tongji University (Shanghai, China). He then received his Ph.D. in medicinal chemistry from Shanghai Institute of Materia Medica, Chinese Academy of Sciences, under the supervision of Professors Hong Liu and Hualiang Jiang. He was a postdoctoral associate with Professor Richard B. Silverman at Northwestern University, where he was involved in discovering novel selective inhibitors of neuronal nitric oxide synthase as potential therapeutic agents for a variety of neurodegenerative problems. Since 2012, he joined Professor Yingming Zhao's group at The University of Chicago as a postdoctoral researcher. His current research is focused on the proteomics studies of lysine modifications and identification of substrates for acetylation-regulatory enzymes.



Shu Lin graduated in 2004 from Peking University (Beijing, China) with a Bachelor's degree in Biology. She obtained a Ph.D. degree from University of Pennsylvania in 2011, with Dr. Marisa S. Bartolomei. She studied epigenetic controls of genomic imprinting in the Bartolomei laboratory. In 2011, Dr. Lin joined Dr. Benjamin A. Garcia's laboratory in Princeton University as a postdoctoral fellow. She moved with the Garcia laboratory to University of Pennsylvania in 2012. In the Garcia laboratory, she is involved in a number of studies focusing on dynamic regulation of histone PTMs. She has recently fully

characterized a synthetic peptide library of histone PTMs and variants for bottom-up mass spectrometry.

Benjamin A. Garcia obtained his B.S. in Chemistry at UC Davis in 2000, where he worked as an undergraduate researcher in Prof. Carlito Lebrilla's laboratory. He then received his Ph.D. in Chemistry in 2005 at the University of Virginia under Prof. Donald Hunt and then was an NIH NRSA Postdoctoral Fellow at the University of Illinois under Prof. Neil Kelleher from 2005–2008. From there Ben was appointed as an Assistant Professor in the Molecular Biology Department at Princeton University from 2008–2012, until his recruitment as the Presidential Associate Professor of Biochemistry and Biophysics at the University of Pennsylvania Perelman School of Medicine in 2012. The Garcia lab has been developing and applying novel proteomic approaches and bioinformatics for interrogating protein modifications, especially those involved in epigenetic mechanisms such as histones, publishing over 125 publications. Dr. Garcia is on the editorial boards for the *BMC Genomics* and *Molecular and Cellular Proteomics* journals, and serves on the Board of Directors for the U.S. Human Proteome Organization. He has also been recognized with many honors and awards for his mass spectrometry research including the American Society for Mass Spectrometry Research Award, a National Science Foundation early faculty CAREER award, an NIH Director's New Innovator Award, the Presidential Early Career Award for Scientists and Engineers (PECASE), an Alfred P. Sloan Fellowship, an AB Sciex Young Investigator award, a Biomed Central Research award in Molecular and Cellular Science, the PITTCON Achievement Award, and the American Chemical Society Arthur F. Findeis Award for Achievements by a Young Analytical Scientist.



Yingming Zhao received his Ph.D. degree from the Rockefeller University under Professor Brian Chait in 1997. He is a Professor in the Ben May Department for Cancer Research at the University of Chicago. His main research interests for the past several years lie in developing and applying MS-based proteomics technologies in the discovery of new protein post-translational modification (PTM) pathways. He also uses an integrated approach, involving proteomics, biochemistry, molecular biology, and cell biology, to decode PTM networks that have implications for human health and are not amenable to conventional

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

techniques. He and his colleagues carried out the first lysine acetylation proteomic studies, demonstrating the role of this modification outside nuclei. His lab recently discovered seven types of new lysine acylation pathways: propionylation, butyrylation, crotonylation, malonylation, succinylation, glutarylation, and 2-hydroxyisobutyrylation. They identified about 200 new histone marks as of today, which more than doubles the number of the histone marks discovered during the first 40 years of chromatin biology. His laboratory identified thousands of protein substrates and revealed numerous enzymes for the new PTM pathways and their lysine acylation substrates. His laboratory demonstrates that the new PTM pathways have critical roles in epigenetic regulation and cellular metabolism. These pathways also can explain cellular dysfunctions associated with diverse inborn metabolic diseases, therefore offering new avenues for therapeutic intervention.

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript

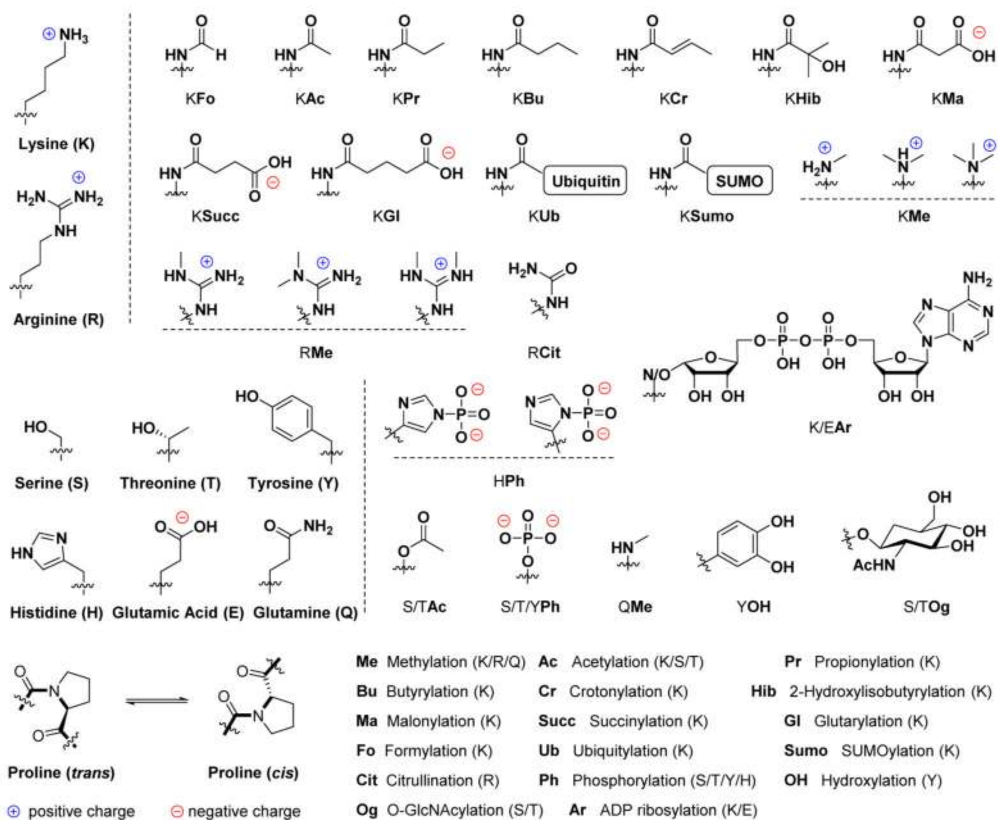


Figure 1.

Structures of histone post-translational modifications.

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

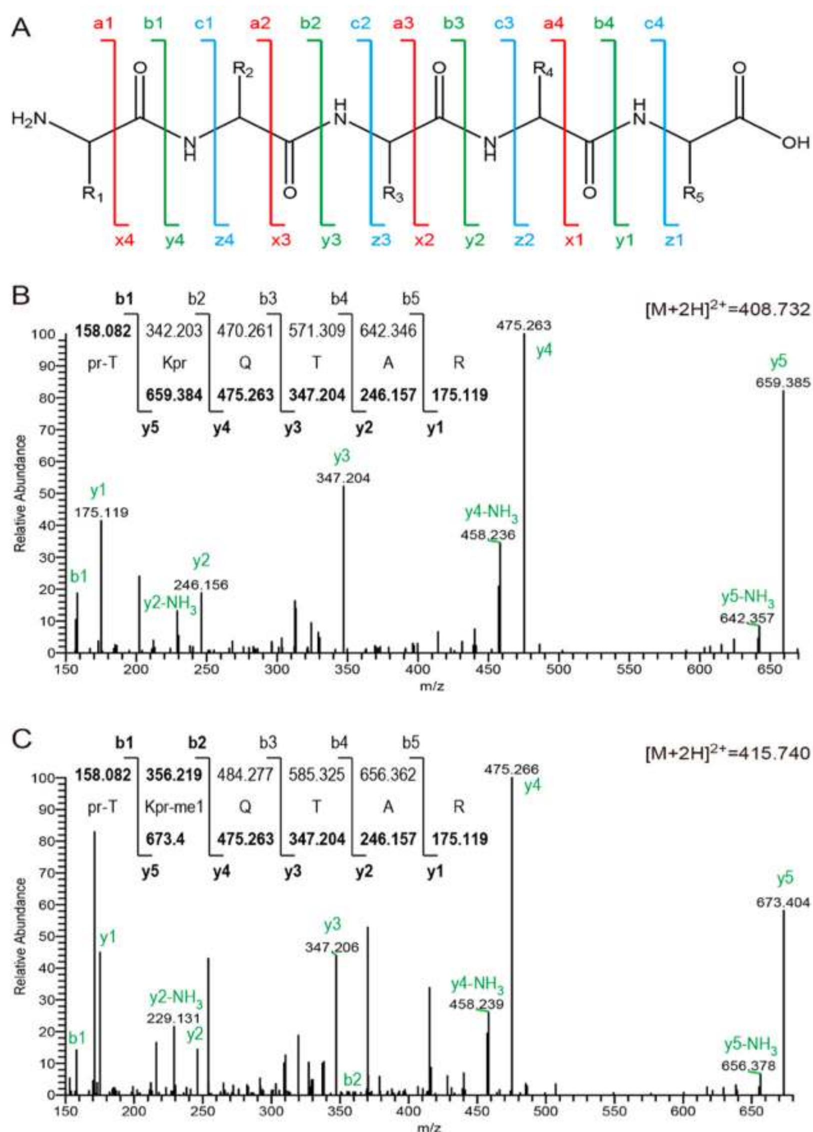


Figure 2.

MS/MS for peptide sequencing and PTM detection. (A) Nomenclature for fragment ions in mass spectra for peptides (modified from ref 45). Schematic showing a five residue peptide. The vertical lines show the bond cleavage, and the horizontal lines show the paired-product ions formed. Red, a and x ions; green, b and y ions; blue, c and z ions. (B) A MS/MS spectrum example obtained by HCD. The N-terminal of this peptide and the lysine residue is propionylated, adding 56.026 Da. The parent ion is charge +2 with *m/z* equals to 408.732. The detected b and y ions are highlighted and labeled. (C) MS/MS spectra showing the same peptide that is monomethylated on the lysine residue, which adds 14.016 Da.

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

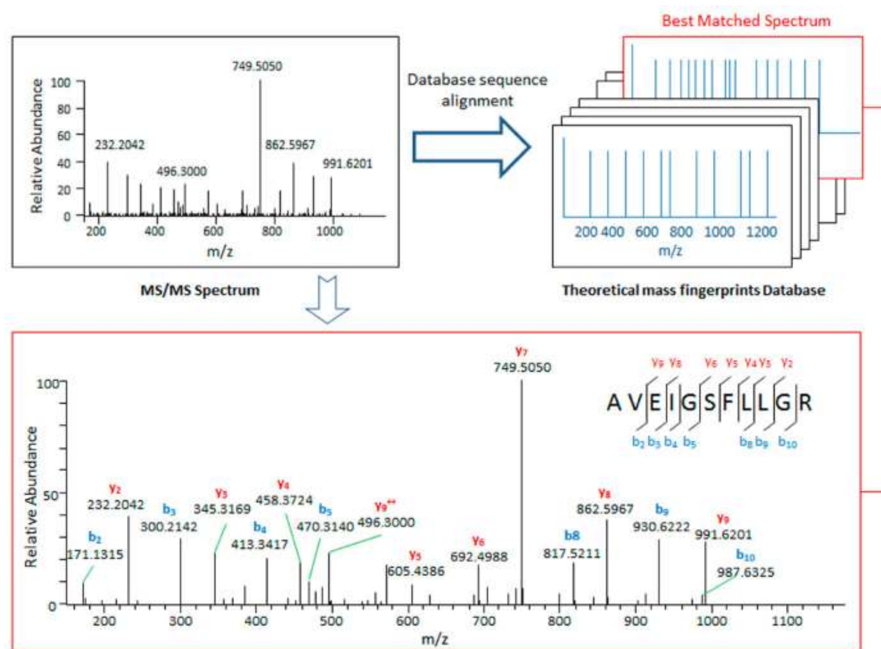


Figure 3.
Schematic overview of the peptide fingerprint alignment.

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript

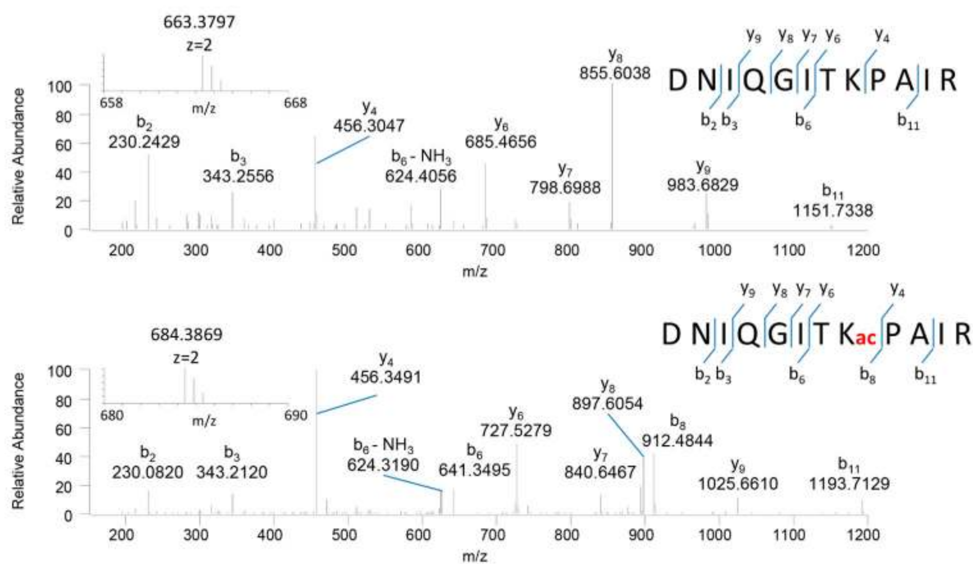


Figure 4.

An example of a mass shift caused by lysine acetylation. Insets show the precursor ion masses. In addition to the parent ions, the daughter ions containing the acetyllysine have also a mass shift of 42.0106 Da caused by acetylation.

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

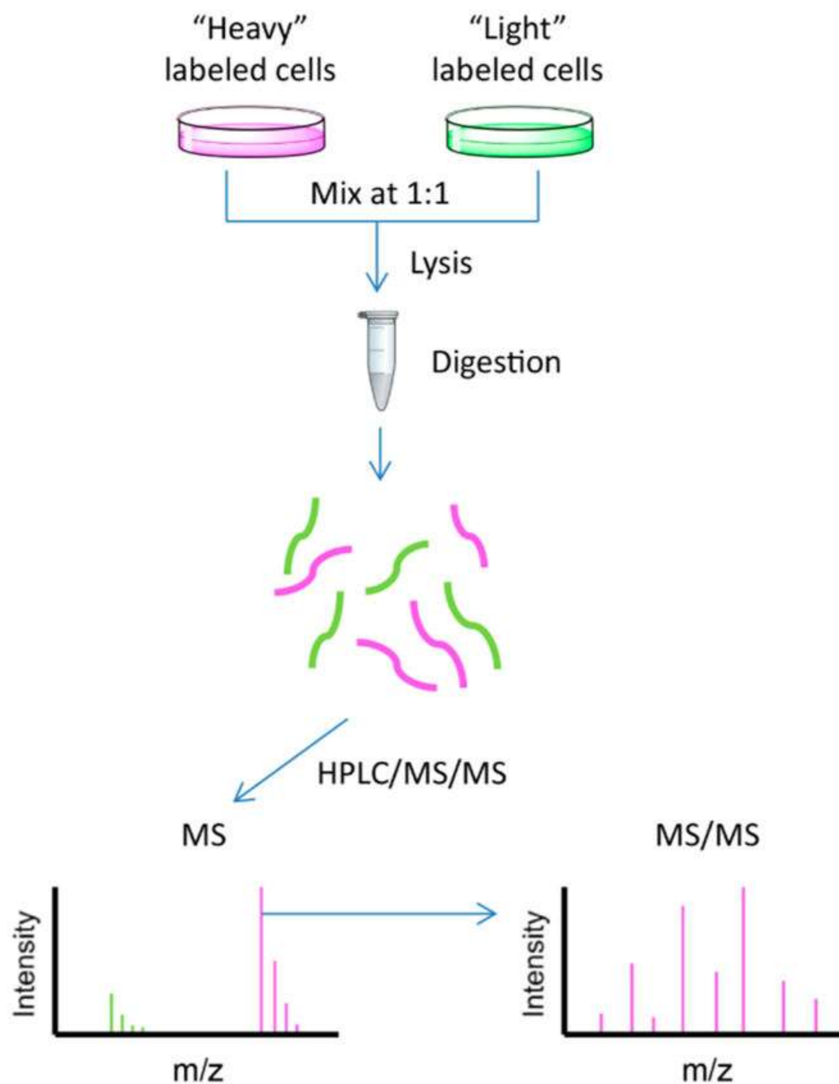


Figure 5.

Schematic showing SILAC followed by HPLC/MS/MS. Treatment and control cells are cultured in different media, containing "light" and "heavy" isotope-containing amino acids, respectively. Equal numbers of "light" and "heavy" cells are mixed for collecting proteins. The protein mixture is then digested by a protease of choice. The resulting peptide mixture can be subjected to HPLC/MS/MS analysis, with or without prior fractionation. The same peptide from the two cell populations can be identified and quantified by MS analysis.

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

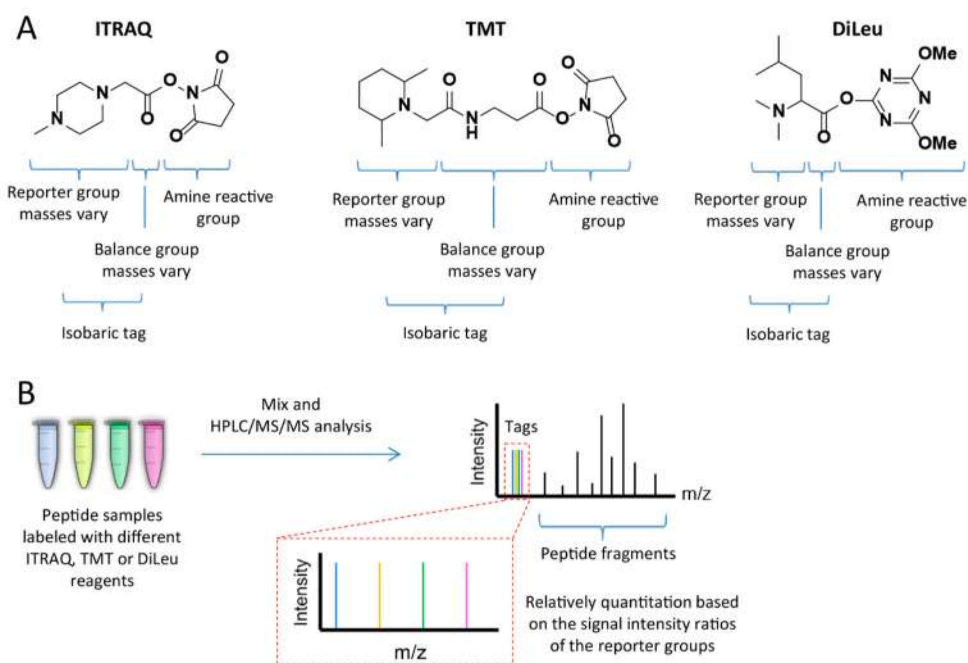


Figure 6.

(A) Chemical structures of ITRAQ, TMT, and DiLeu reagents, showing design principles for these reagents. In multiplexing reagents, the reporter group carries different numbers of ^{13}C and/or ^{15}N atoms, resulting one dalton different in mass among different tags. The balance group is also labeled by different numbers of stable isotopes. Thus, the combined report and balance groups have the same total molecular weights. The amine reactive group reacts with amine groups on peptides' N-termini and unmodified lysine residues, adding the isobaric tag onto the peptides. (B) Example of multiplex proteomic quantitation with ITRAQ, TMT, or DiLeu reagents.

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

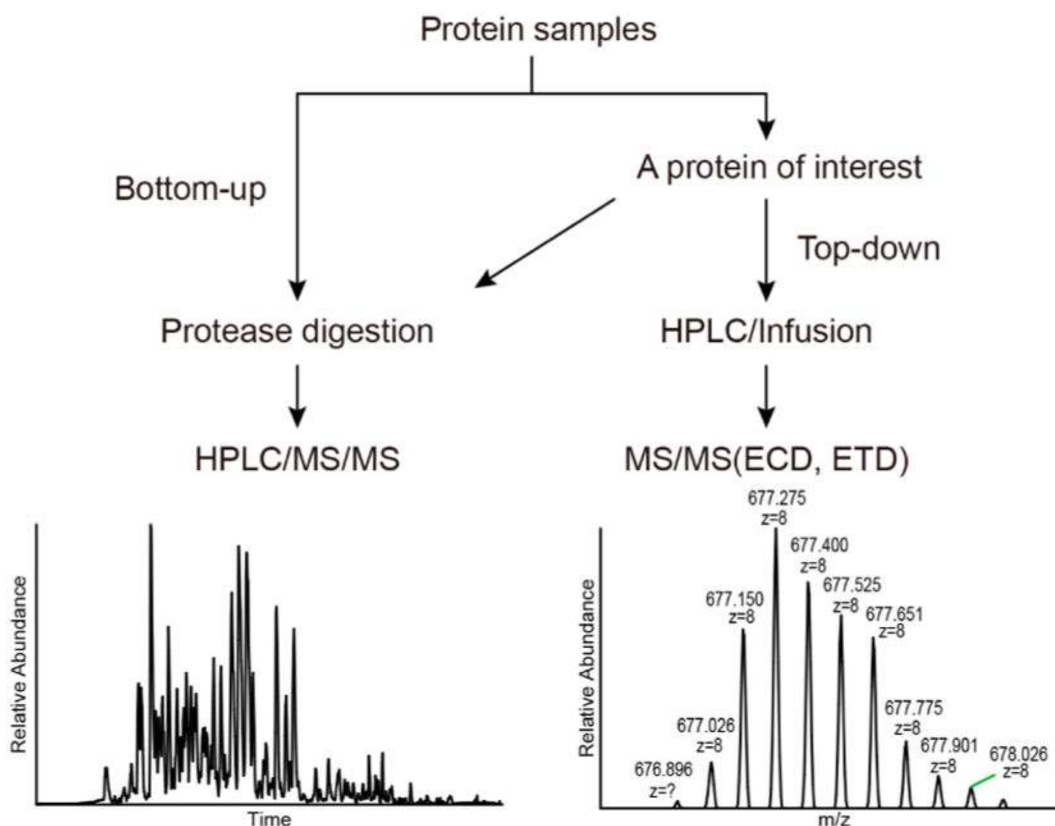


Figure 7.

Bottom-up and top-down mass spectrometry methods. In the bottom-up method, proteins are digested by a protease such as trypsin, Arg-C, Glu-C, or Asp-N protease. The peptides are then subjected to HPLC/MS/MS analysis. The chromatogram shows the base-peaks of a bottom-up MS run. In the top-down approach, a purified protein is analyzed in MS that is either directly infused or separated in HPLC before MS analysis. ETD is usually chosen as the MS/MS fragmentation technique. The spectrum shows isotope distribution of a protein ion population of charge +8.

Author Manuscript
Author Manuscript
Author Manuscript
Author Manuscript

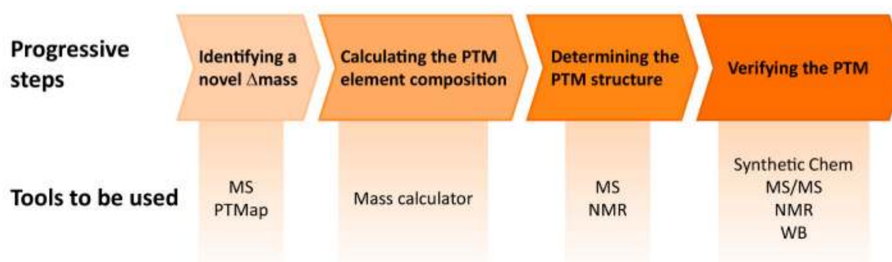


Figure 8.
Progressive steps for identification, verification, and systematic analysis of novel PTMs in histones.

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

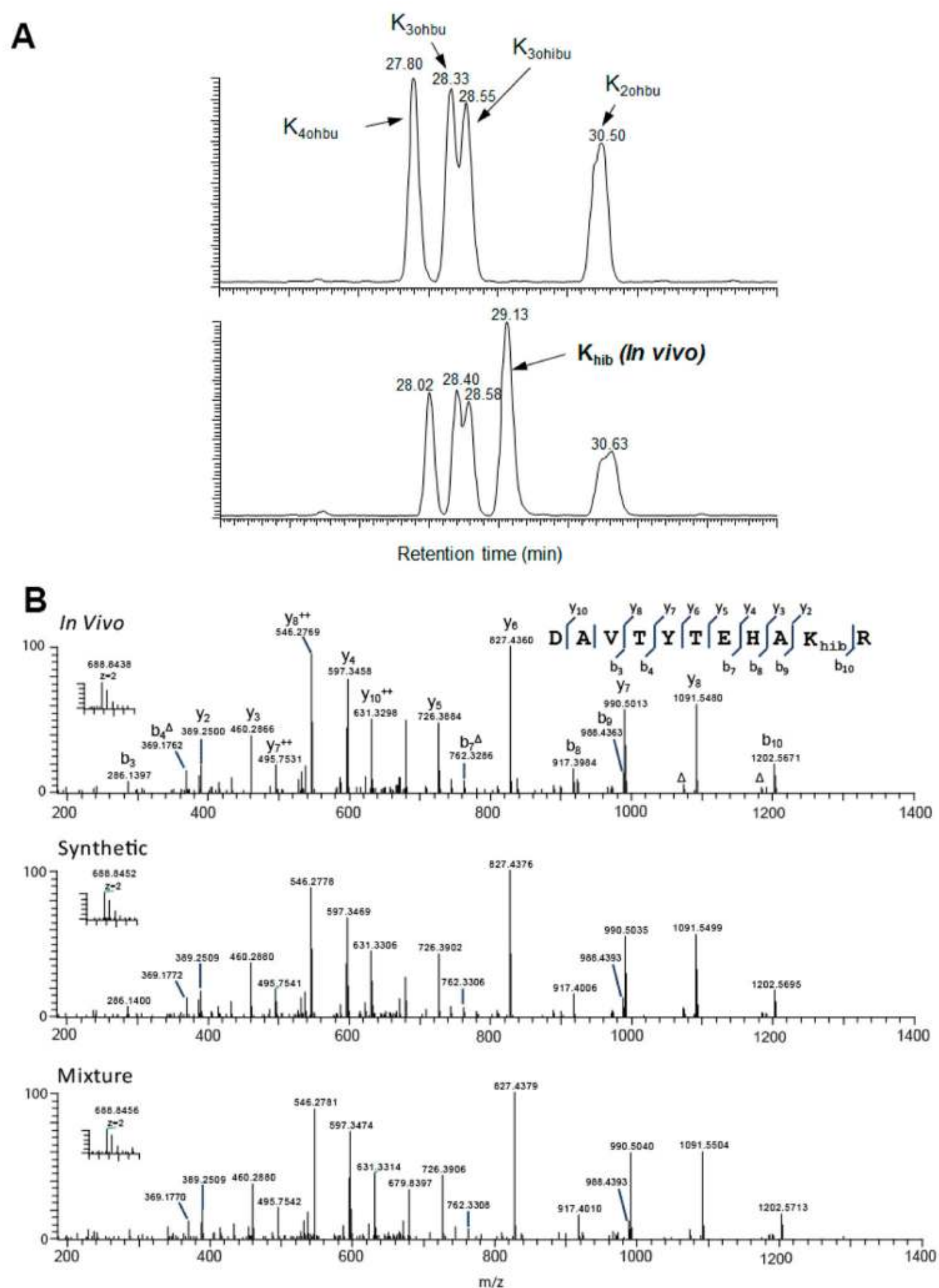


Figure 9.

Verification of lysine 2-hydroxyisobutyrylation. (A) Extracted ion chromatograms (XICs) from HPLC/MS analysis of a mixture of four synthetic peptides (DAVTYTEHAK$_{(\pm)\text{-2ohbu}}$R, DAVTYTEHAK$_{(r)\text{-3ohbu}}$R, DAVTYTEHAK$_{(s)\text{-3ohibu}}$R, and DAVTYTEHAK$_{\text{4ohbu}}$R), without (top) or with (bottom) the in vivo peptide. The *x* axis represents retention time of HPLC/MS analysis, while the *y* axis represents the MS signal. (B) The MS/MS spectra of an in vivo peptide bearing a PTM (DAVTYTEHAK$_{+86.0354}$R) (top), a synthetic lysine 2-hydroxyisobutyrylated peptide corresponding to the sequence of the in vivo peptide (middle), and a mixture of the two peptides (bottom). The label Δ designates b or y ions with

water and/or ammonia loss. Insets show the precursor ion masses. The data are from the published literature.[6a]

Author Manuscript
Author Manuscript
Author Manuscript
Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript

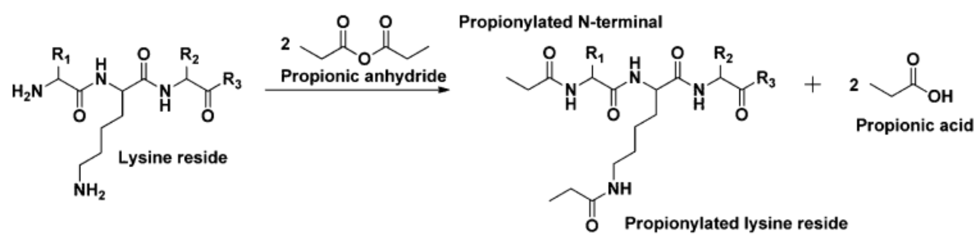


Figure 10.

Chemical derivatization of lysine-containing peptides by propionic anhydride. R_1 and R_2 represent amino acid side chains, and R_3 represents other residues. After the derivatization, both the N-terminal amine group and the amine group on the unmodified lysine residue are modified with propionyl group. Propionic acids are the side products of this reaction.

Author Manuscript
Author Manuscript
Author Manuscript
Author Manuscript

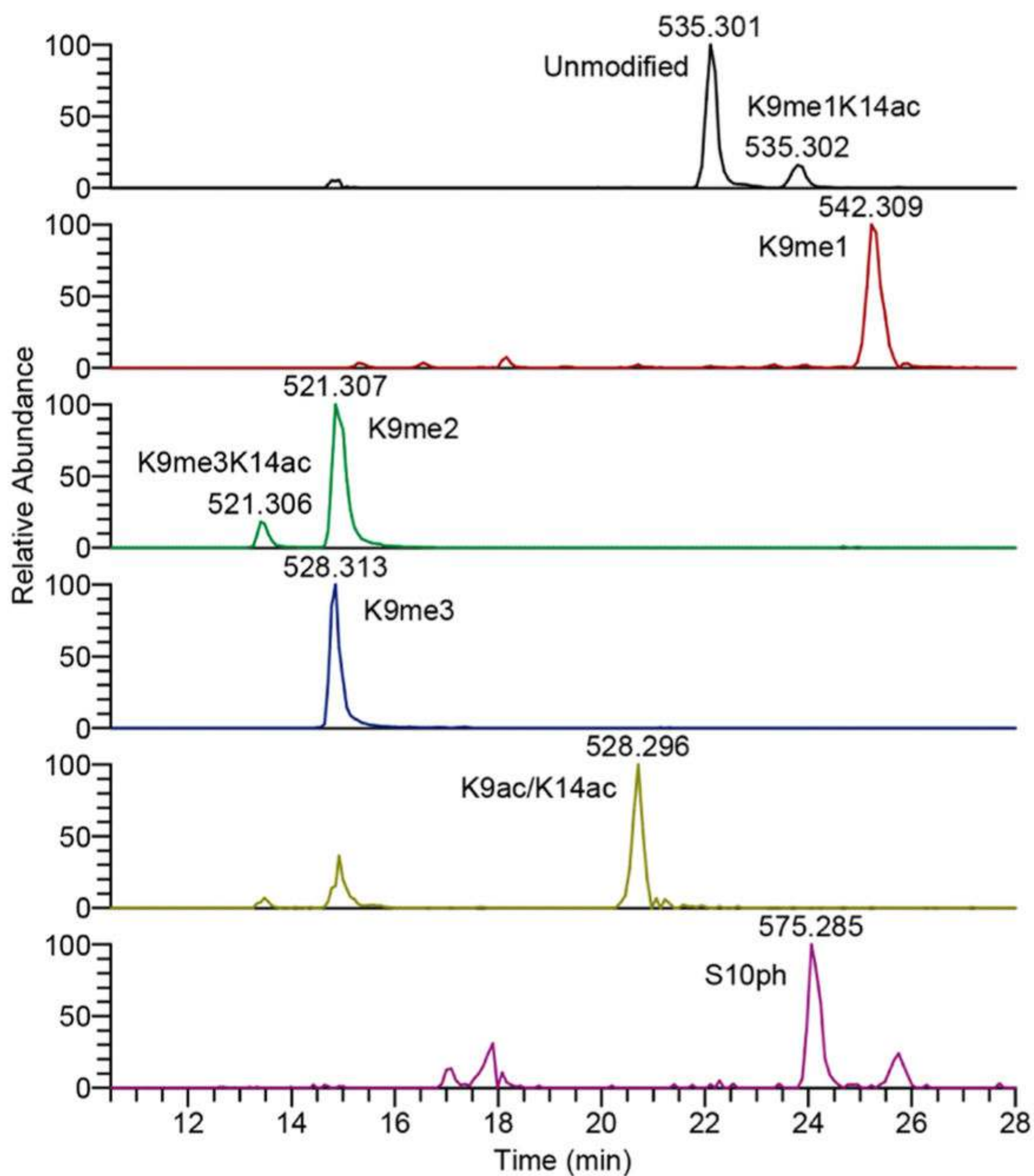


Figure 11.

HPLC/MS/MS analysis of the $[M+2H]^{2+}$ ions of histone H3 9–17 peptide KSTGGKAPR. Eight forms with various PTMs were detected and shown. The peptides are propionylated and digested by trypsin. The PTMs and *m*/*z* values of the peptides are indicated.

Author Manuscript

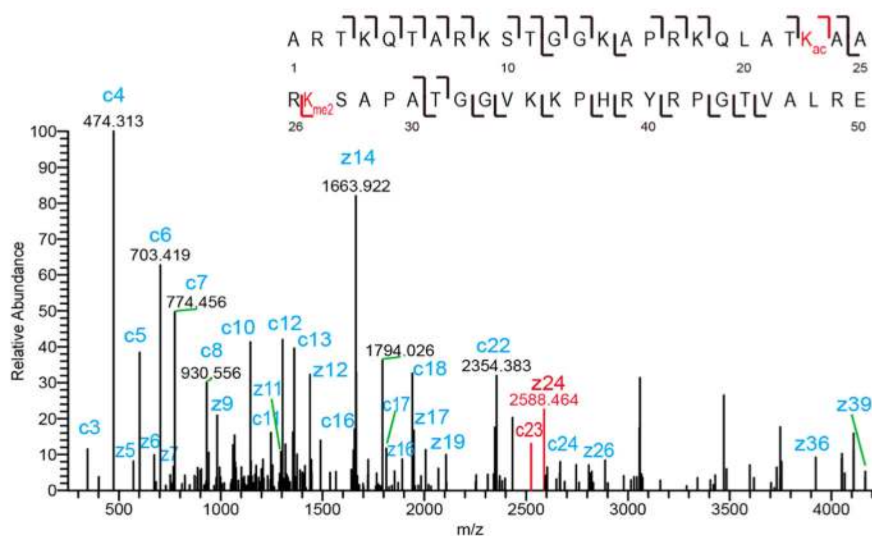


Figure 12.

Example of MS/MS achieved by ETD fragmentation for the H3 1–50 peptide with K23 acetylated (ac) and K27 dimethylated (me2). The raw spectrum was deconvoluted and deisotoped by the Xtract program as described.[145c,146] The relative abundance of the c4 ion (*m/z* = 474.313) was set to be 100%. Identified c and z ion peaks are annotated both on the peptide sequence and in the spectrum. The modified lysine residues and critical fragment ions, c23 and z24, are highlighted in red.

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

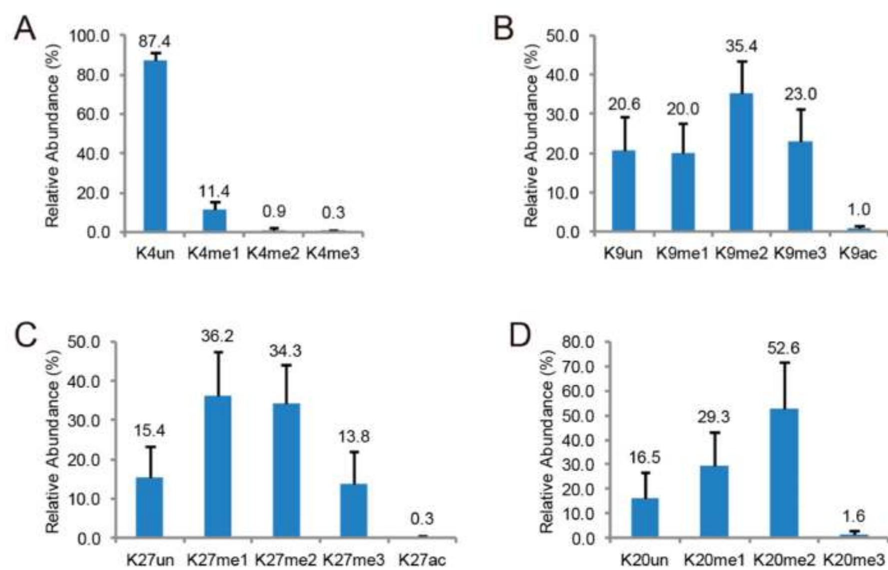


Figure 13.

Averaged PTM data from 24 human cell lines; data achieved from ref 142). (A) Histone H3K4 PTMs; (B) H3K9 PTMs; (C) H3K27 PTMs; (D) H4K20 PTMs. The numbers above each bar indicate the percentiles of the particular PTM over the total H3 signal (100%). Error bars represent standard deviation among 24 cell lines.

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

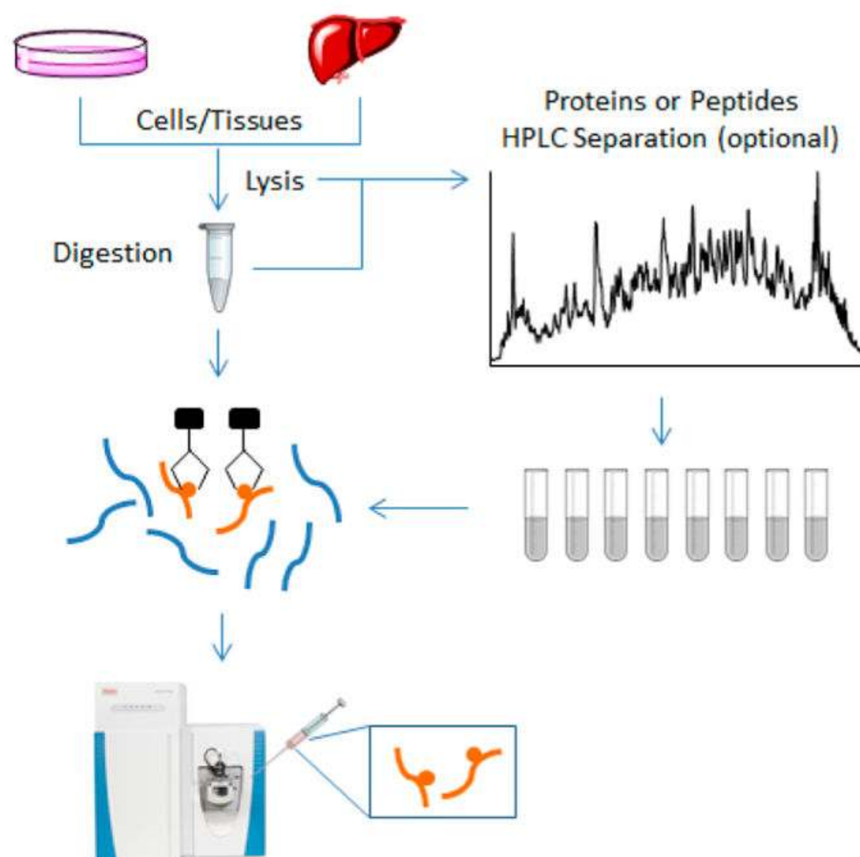


Figure 14.
Typical workflow of the proteomic analysis of PTMs.

Author Manuscript
Author Manuscript
Author Manuscript
Author Manuscript

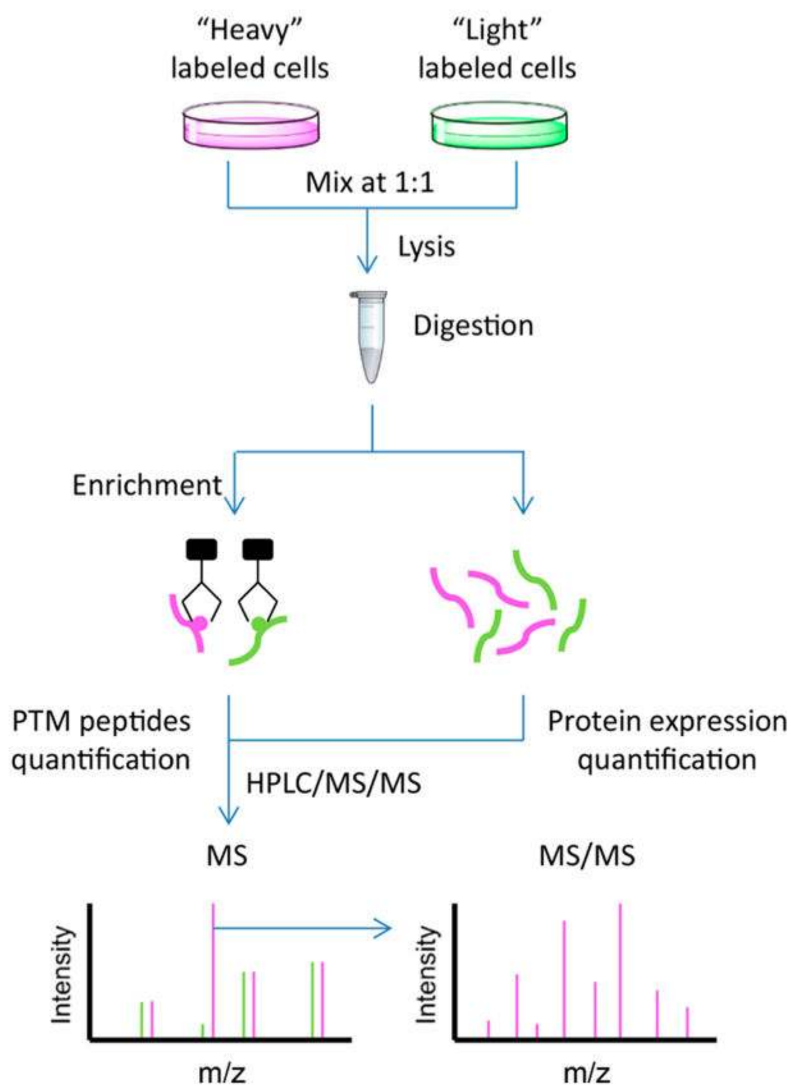


Figure 15.
Schematic representation of experimental workflow for quantitative proteomics of PTMs.

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

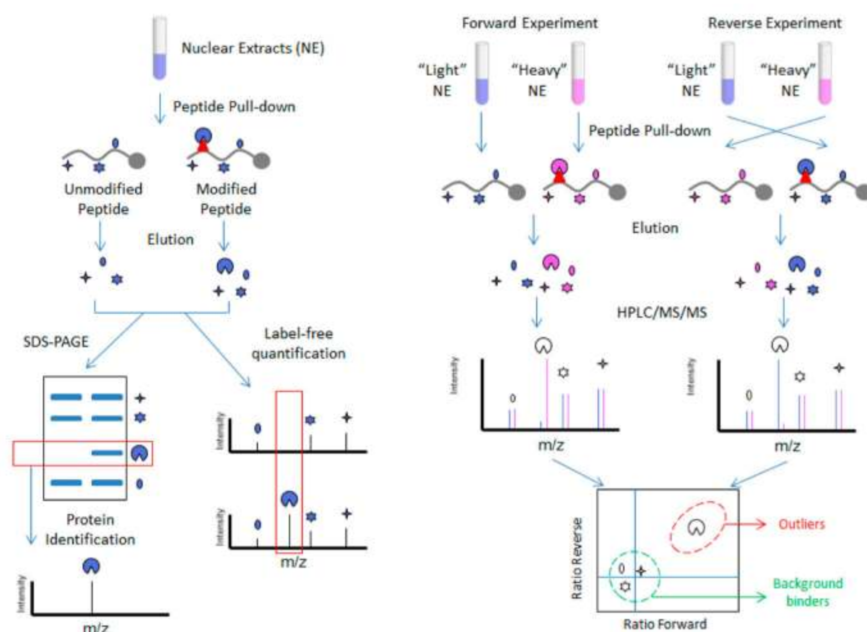


Figure 16.

Schematic overview of the workflow for identifying histone-mark "binders". A pair of biotinylated peptides are synthesized and used as baits to incubate with a protein lysate for pulling down experiment. Gray curves indicate immobilized histone tail peptide. Red triangle indicates histone mark. (Left) The proteins isolated with modified and unmodified peptides were resolved in SDS-PAGE; proteins specific to histone marks were visualized and then identified by MS. Alternatively, the enriched proteins from the two affinity enrichment experiments are digested and then analyzed by HPLC/MS/MS for identifying and quantifying proteins. The histone mark-specific binding proteins will be identified. (Right) SILAC-based quantitative proteomic approach for identifying and quantifying proteins that bind to a histone mark of interest.

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript

Table 1

Bottom-Up MS for Detection of Histone H3.1/2 and H4 Modifications

protein	peptide position	peptide sequence	modifications
H3.1/2	3–8	TKQTAR	K4me1, me2, and me3
	9–17	KSTGGKAPR	K9ac; me1, me2, and me3; S10ph; K14ac
	18–26	KQLATKAAR	K18ac, me1; K23ac, me1
	27–40	KSAPATGGVKKPHR	K27ac, me1, me2, and me3; K36me1, me2, and me3
	54–63	YQKSTELLIR	K56ac
	73–83	EIAQDFKTDLR	K79me1, me2, and me3
H4	4–17	GKGGKGLGKGGAKR	K5ac; K8ac; K12ac; K16ac

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

Table 2

Histone Variants and Their Functions[a]

protein	variant	function	distribution[b]
H2A	H2A.1/2	canonical, genome packaging	universal
	H2A.Z	transcription regulation	universal
	H2A.X	DNA double-strand break repair	universal
	macroH2A	X chromosome inactivation and others	animals
H3	H3.1/2	canonical, genome packaging	widespread
	H3.3	replacement and transcriptional activation	universal
	CENP-A	centromere identity	universal

[a]Modified from ref 183.

[b]Universal means all eukaryotes have the corresponding histone variants, which typically are shown to have similar biological functions. Widespread means most higher eukaryotes have the corresponding histone variants, H3.1/2. The exceptions are some unicellular organisms, including *S. cerevisiae* and some algae species. They only have one version of noncentromeric H3. This single version is more similar to H3.3 than H3.1/2.

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

Table 3

Histone Phosphorylation Marks

site	ref	site	ref	site	ref
H3T3	190	H4Y72	189	H2BS75	141
H3T6	191	H4H75	192	H2BS87	193
H3S10	194	H4Y88	193	H2BT88	193
H3T11	195	H2AS1	196	H2BS91	193
H3S28	197	H2AY50	157	H2BT115	157
H3Y41	198	H2AT59	157	H2BT119	157
H3T45	199	H2AT101	157	H1S1	182a
H3S57	200	H2AT120	157, 201	H1T3	182a
H3T80	150f	H2AS122	157	H1S35	182a
H3S86	157	H2BS6	141	H1S40	157
H3T107	141	H2BS14	202	H1S54	157
H3T118	203	H2BS32	204	H1Y70	157
H4S1	205	H2BS36	206	H1T145	182a
H4H18	192b	H2BT52	157	H1T164	6d
H4S47	157, 203, 207	H2BS56	157	H1S172	157
H4Y51	157	H2BS64	157	H1T179	182a

Author Manuscript
Author Manuscript
Author Manuscript
Author Manuscript

Table 4

Histone Acetylation Marks

site	ref	site	ref	site	ref
H3K4	208	H4K77	141	H2BK108	141, 209
H3K9	210	H4K79	141, 209	H2BK116	141
H3S10	211	H4K91	141, 209	H2BK120	141, 209, 212
H3K14	213	H2AK5	214	H2BK125	141
H3K18	213a, 215	H2AK9	141	H1K16	141
H3T22	211	H2AK36	216	H1K33	217
H3K23	152, 218	H2AK74	141	H1S35	141
H3K27	141, 208a, 219	H2AK95	141	H1K45	141
H3S28	211	H2AK118	129a	H1S50	141
H3K36	208a, 220	H2AK127	141	H1K51	141
H3K37	221	H2AK129	141	H1K63	222
H3K56	208a, 223	H2BK5	152, 224	H1K74	123c
H3K64	225	H2BK11	141, 152	H1K89	6b
H3K79	6b	H2BK12	141, 152, 226	H1K96	6d
H3T80	141	H2BK15	141, 152	H1K105	123c
H3K115	141, 209	H2BK16	141, 152, 227	H1S112	141
H3K122	133d, 141, 209	H2BT19	141	H1K167	141
H4K5	228	H2BK20	141, 152	H1K168	6d
H4K8	152, 213a	H2BK23	216	H1K190	6d
H4K12	229	H2BK24	216	H4 N-term	230
H4K16	229a, 231	H2BK46	129a	H2A N-term	230a, c
H4K20	141, 209	H2BK57	141		
H4K31	123c, 208a	H2BK85	209		

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

Table 5

Histone Lysine and Glutamine Methylation Marks[a]

site	ref	site	ref	site	ref
H3K4me1	232	H3K122me1	6b, 203	H2BK23me1	6b
H3K4me2	233	H3K122me2	193	H2BK23me2	209
H3K4me3	233, 234	H4K5me3	141, 235	H2BK34me1	141
H3K9me1	236	H4K12me1	209	H2BK43me1	209
H3K9me2	237	H4K16me1	141	H2BK46me1	152
H3K9me3	238	H4K16me3	141	H2BK47me1	152
H3K14me1	193	H4K20me1	141, 208a	H2BK57me1	6b
H3K14me2	193	H4K20me2	141	H2BK57me2	152
H3K14me3	141	H4K20me3	239	H2BK85me1	6b
H3K18me1	6b, 208a	H4K31me1	240	H2BK108me1	152
H3K23me1	6b, 208a	H4K31me2	152	H2BK116me1	6b
H3K23me2	193	H4K59me1	6b, 203	H1K16me2	193
H3K27me1	241	H4K59me2	6b	H1K21me2	193
H3K27me2	208a, 241	H4K77me1	6b, 152	H1K25me2	242
H3K27me3	233, 241	H4K79me1	138c	H1K33me1	6b
H3K36me1	141, 208a	H2AK9me1	6b	H1K51me1	6b
H3K36me2	141, 208a	H2AK9me2	141	H1K62me1	6b
H3K36me3	141, 208a	H2AK74me1	243	H1K63me1	6b
H3K37me1	244	H2AK95me1	141	H1K89me1	6b
H3K56me1	208a, 245	H2AK95me2	141	H1K96me1	6b
H3K56me3	208a, 246	H2AK99me1	141	H1K105me1	6b
H3K64me1	247	H2AK99me2	141	H1K128me1	6b
H3K64me3	248	H2AK118me1	6b	H1K147me1	6b
H3K79me1	208a, 249	H2AK125me1	6b	H1K167me1	6b
H3K79me2	208a, 249	H2BK5me1	13	H1K186me1	250
H3K79me3	208a, 249, 251	H2BK12me3	141	H2AQ104me1	252
H3K83me1	141	H2BK15me1	6b		
H3K83me2	141	H2BK20me1	141		

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

[a] me1, monomethylation; me2, dimethylation; me3, trimethylation.

Author Manuscript
Author Manuscript
Author Manuscript
Author Manuscript

Table 6

Histone Arginine Methylation Marks[a]

site	ref	site	ref	site	ref
H3R2me1	253	H4R3me	141	H4R92me1	203
H3R2me2a	254	H4R3me2a	255	H2AR11me2	256
H3R2me2s	257	H4R3me2s	258	H2AR20me1	259
H3R8me1	258	H4R17me1	141	H2AR29me2a	256
H3R8me2a	260	H4R17me2	141	H2AR42me1	6b
H3R8me2s	258	H4R19me1	141	H2AR71me1	141
H3R17me2a	261	H4R19me2	141	H2AR88me1	6b, 262
H3R26me2a	261b	H4R23me1	141, 262	H2BR79me1	6b
H3R42me2a	263	H4R23me2	141	H2BR86me1	133c, 262
H3R63me1	6b	H4R35me1	6b	H2BR92me1	133c
H3R83me1	264	H4R55me1	6b, 152	H2BR99me1	6b, 209
H3R128me1	6b	H4R67me1	6b		

[a] me1, monomethylation; me2a, asymmetric dimethylation; me2s, symmetric dimethylation; me2, dimethylation, uncertain symmetry.

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript

Table 7

Histone Lysine Propionylation and Butyrylation Marks

Propionylation					
site	**ref**	**site**	**ref**	**site**	**ref**
H3K23	265	H4K12	122	H4K77	122
H3K56	266	H4K16	122	H4K79	122
H4K5	122	H4K31	122	H4K91	122
H4K8	122	H4K44	122	H2AK125	141

Butyrylation					
site	**ref**	**site**	**ref**	**site**	**ref**
H3K14	266	H4K12	122	H4K77	122
H3K27	266	H4K16	122	H4K79	122
H4K5	122	H4K31	122	H4K91	122
H4K8	122	H4K44	122		

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

Table 8

Histone Lysine Crotonylation and 2-Hydroxyisobutyrylation Marks

Crotonylation					
site	**ref**	**site**	**ref**	**site**	**ref**
H3K4	6b	H4K77	267	H2BK20	6b
H3K9	6b	H4K91	141	H2BK23	6b
H3K18	6b	H2AK36	6b	H2BK34	6b
H3K23	6b, 267	H2AK95	141	H2BK108	141, 267
H3K27	6b	H2AK118	6b	H2BK116	267
H3K56	6b	H2AK119	6b	H1K33	6b
H3K122	141, 267	H2AK125	6b, 141	H1K63	6b
H4K5	6b	H2BK5	6b	H1K84	6b
H4K8	6b	H2BK11	6b	H1K89	6b
H4K12	6b	H2BK12	6b	H1K96	6b
H4K16	6b	H2BK15	6b	H1K158	6b
H4K59	267	H2BK16	6b	H1K167	6b

2-Hydroxyisobutyrylation					
site	**ref**	**site**	**ref**	**site**	**ref**
H3K4	6a	H2AK5	6a	H1K25	6a
H3K9	6a	H2AK9	6a	H1K26	6a
H3K14	6a	H2AK36	6a	H1K33	6a
H3K18	6a	H2AK74	6a	H1K45	6a
H3K23	6a	H2AK75	6a	H1K51	6a
H3K27	6a	H2AK95	6a	H1K62	6a
H3K36	6a	H2AK118	6a	H1K63	6a
H3K56	6a	H2BK5	6a	H1K74	6a
H3K64	6a	H2BK12	6a	H1K80	6a
H3K79	6a	H2BK20	6a	H1K84	6a
H3K122	6a	H2BK23	6a	H1K89	6a
H4K5	6a	H2BK24	6a	H1K96	6a
H4K8	6a	H2BK34	6a	H1K109	6a

2-Hydroxyisobutyrylation					
site	**ref**	**site**	**ref**	**site**	**ref**
H4K12	6a	H2BK43	6a	H1K116	6a
H4K16	6a	H2BK46	6a	H1K120	6a
H4K31	6a	H2BK57	6a	H1K128	6a
H4K44	6a	H2BK85	6a	H1K135	6a
H4K59	6a	H2BK108	6a	H1K147	6a
H4K77	6a	H2BK116	6a	H1K158	6a
H4K79	6a	H2BK120	6a	H1K167	6a
H4K91	6a	H1K22	6a	H1K212	6a

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

Table 9

Histone Lysine Malonylation, Succinylation, and Glutarylation Marks

Malonylation					
site	ref	site	ref	site	ref
H3K14	150a	H4K79	150a	H1K45	150a
H3K18	150a	H2AK95	150a	H1K62	150a
H3K23	150a	H2BK5	150a	H1K63	150a
H3K56	123b, 150a	H2BK34	150a	H1K74	150a
H3K79	150a	H2BK46	150a	H1K84	150a
H3K122	150a	H2BK108	150a	H1K89	150a
H4K8	150a	H2BK116	123b	H1K96	150a
H4K31	150a	H2BK120	150a	H1K105	150a
H4K77	150a	H1K33	150a	H1K159	150a

Succinylation					
site	ref	site	ref	site	ref
H3K14	123b	H2AK9	123b	H1K33	123c
H3K23	123c	H2AK36	123b	H1K45	123c
H3K27	123c	H2AK95	123b	H1K62	123c, 268
H3K56	123b, c, 268	H2BK5	268	H1K63	123c
H3K79	123b, c, 268	H2BK34	123b	H1K74	123c
H3K122	123b, c, 268	H2BK43	268	H1K84	268
H4K12	123b, c	H2BK46	268	H1K89	123c, 268
H4K31	123b, c, 268	H2BK85	268	H1K96	123c, 268
H4K77	123b, c, 268	H2BK108	123c	H1K105	123c, 268
H4K79	123b	H2BK116	123b	H1K120	123c
H4K91	123b, c, 268	H2BK120	123b, 268		

Glutarylation					
site	ref	site	ref	site	ref
H2BK5	128b	H2BK116	128b	H2BK120	128b

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

Table 10

Histone Formylation and Ubiquitination Marks

Formylation					
site	ref	site	ref	site	ref
H3K18	269	H2AK36	269	H1K33	269
H3K23	269	H2AK95	269	H1K45	269
H3K56	6b	H2AK118	269	H1K62	269
H3K64	269	H2BK5	269	H1K63	269
H3K79	269	H2BK34	269	H1K74	269
H4K12	269	H2BK43	269	H1K80	6b
H4K31	269	H2BK46	269	H1K84	269
H4K59	269	H2BK108	269	H1K89	269
H4K77	269	H2BK116	6b	H1K96	269
H4K79	269	H2BK120	6b	H1K139	269
H4K91	269	H1K16	269	H1K159	269

Ubiquitination					
site	ref	site	ref	site	ref
H3K14	270	H2AK36	270	H1K45	270
H3K18	270	H2AK95	270	H1K51	270
H3K23	270, 271	H2AK118	270	H1K63	270
H3K27	270	H2AK119	272	H1K74	270
H3K36	270	H2AK125	270	H1K89	270
H3K56	270	H2BK20	273	H1K96	270
H3K79	270	H2BK34	274	H1K105	270
H3K122	270, 275	H2BK46	270	H1K109	270
H4K31	270	H2BK57	270	H1K116	270
H4K59	270	H2BK108	270	H1K126	270
H4K77	270	H2BK116	270	H1K139	270
H4K91	270	H2BK120	276	H1K159	270
H2AK15	277	H1K33	270	H1K167	270

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

Table 11

Other Histone Marks[a]

Citrullination					
site	**ref**	**site**	**ref**	**site**	**ref**
H3R2	278	H3R26	278b, 279	H4R19	280
H3R8	278	H4R3	278	H2AR3	281
H3R17	278	H4R17	280	H1R53	282

Hydroxylation					
site	**ref**	**site**	**ref**	**site**	**ref**
H4Y51	6b	H2AY39	6b	H2BY83	6b
H4Y88	6b	H2BY37	6b	H1Y70	6b

O-GlcNAcylation					
site	**ref**	**site**	**ref**	**site**	**ref**
H3S10	119c	H4S47	119b	H2BS36	119b
H3T32	283	H2AT101	119b	H2BS112	284

ADP Ribosylation					
site	**ref**	**site**	**ref**	**site**	**ref**
H3K27	285	H2AK13	285	H1E2	286
H3K37	285	H2BE2	286		
H4K16	285	H2BK30	285		

Proline Isomerization					
site	**ref**	**site**	**ref**	**site**	**ref**
H3P16	287	H3P30	287a, 288	H3P38	287a, 288

SUMOylation					
site	**ref**	**site**	**ref**	**site**	**ref**
H3K14	289	H4K12*	290	H2BK16*	290
H3K18	289	H4K16*	290	H2BK17*	290
H3K23	289	H4K20*	290	H2BK20	289

		SUMOylation			
site	ref	site	ref	site	ref
H3K56	289	H2AK118	289	H2BK108	289
H3K79	289	H2AK126*	290	H2BK116	289
H3K122	289	H2BK5	289	H2BK120	289
H4K5*	290	H2BK6*	290	H1K63	289
H4K8*	290	H2BK7*	290		

[a]Asterisk indicates that the histone sequences from *S. cerevisiae* are used as templates for numbering of modified residues.

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

Table 12

Methods Developed for Proteomic Analysis of PTMs by HPLC/MS/MS

PTM	enrichment methods	ref
phosphorylation	IMAC	325a
	TiO2	326a
	immunoaffinity purification	319a
	SCX	332a
	HILIC	330a
	ERLIC	331
	phosphate precipitation	335
	HAP	336
acetylation	immunoaffinity purification	129b
	off-line fractionation using basic HPLC, and immunoaffinity purification	324
methylation	immunoaffinity purification	169a
	MBT domain affinity enrichment	321c
	derivatization and immunoaffinity purification	321a
ubiquitination	immunoaffinity purification	323
	off-line fractionation using basic HPLC, and immunoaffinity purification	322a
O-GlcNAcylation	BEMAD	346
	conjugation of glycoproteins to a solid support using hydrazide chemistry	345b
	tagging-via-substrate, protein/peptide enrichment with biotin/click chemistry	341a, 344a
	lectin weak affinity chromatography (LWAC)	339a
	immunoaffinity purification	356
	boronate affinity	357
ADP ribosylation	macro domains affinity enrichment	340a
	boronate affinity	313

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

Table 13

A List of Presentative Proteomic Analysis of the Popular PTMs

PTM	cell line/tissue/species	ref
phosphorylation	mouse liver; drosophila embryos	328
	mouse brain	319b
	Jurkat T cells	358
	Saccharomyces cerevisiae	158
	mouse liver	359
	human leukemia cells	149
	Saccharomyces cerevisiae	360
	human tissues	361
	luminal and basal type breast cancer cells	362
	mouse synaptosome	363
acetylation	$SIRT1^{+/+}$ and $SIRT1^{-/-}$ MEF cells	129a
	drosophila	364
	mouse liver	318
	mouse liver	365
	human leukemia cells	149
	Saccharomyces cerevisiae	366
	E. coli	320b
	human acute myeloid leukemia cells	129c
	human liver	129d
	E. coli	131
	mouse liver	367
	rat tissues	368
methylation	HCT116 cells, MEF cells, mouse brain	259
	HeLaS3 cells	264
	HeLaS3 cells	321b
malonylation	MCD cells; Sirt5 KO mouse liver	150a
succinylation	$SIRT5^{+/+}$ and $SIRT5^{-/-}$ MEF cells; mouse liver	123c
	E. coli; S. cerevisiae; HeLa cells; mouse liver	268
formylation	MCF7, HeLa, and A549 cells; mouse liver	269
ubiquitination	HCT116 and 293T cells	270
	Jurkat E6-1 cells	273
	human leukemia cells	149
	HCT116 cells, Hela cells	322b
O-GlcNAcylation	rat forebrain	344a
	HEK293 cells	369
	mouse cerebrocortical brain tissue	370
	NIH 3T3 cells	371
	HeLa cells	372
	mouse synaptosome	363