Review

https://dx.doi.org/10.14889/jpdm.2021.0001

JPDM
Journal of Proteome Data and Methods

Comprehensive mass spectrometric investigation strategies of the human methylproteome

Byron Baron[1*]

1. Centre for Molecular Medicine and Biobanking, University of Malta, Msida, Malta.

ORCID: Byron Baron: https://orcid.org/0000-0001-5722-6295

* Corresponding authors.
E-mail addresses: byron.baron@um.edu.mt

© 2021 The Authors. This is an open access article under the CC BY license (http://creativecommons.org/licenses/by/4.0/).

Keywords

Methylproteomics, mass spectrometry, immunoaffinity, methyl-SILAC

Abstract

Protein methylation has been studies for over two decades, mainly in histones but more recently also in non-histone proteins. Despite the great advances in technologies, particularly the application of Orbitrap technology, proteomic analysis of methylation still remains a challenge. This is mainly due to the small size of methylations, their position on fragmented peptides and issue of incorrect identification. A number of different methods have been developed or modified to varying degrees of success in order to enrich for protein methylation prior to mass spectrometric analysis. Despite the availability of different techniques, it is important to understand the underlying strategy employed, the type of sample being analysed and the methylation-related question being investigated. Without such knowledge it is very easy to become lost analysing large volumes of insignificant data or misinterpret the data output.

Contents

1. Introduction
    1.1 Protein Methylation
2. Mass spectrometry
    2.1 Effect of methylation on enzyme digestion
    2.2 Fragmentation of methylated peptides and position identification
    2.3 Fractionation by FAIMS
3. Methods for methylation enrichment
    3.1. Antibody-based enrichment
        3.1.1. Immunoaffinity purification
        3.1.2. Chemical Derivatisation
    3.2. Antibody-free enrichment - chemical affinity methods
        3.2.1. Low-pH SCX chromatography
        3.2.2. High pH SCX chromatography
        3.2.3. Immobilised Metal Affinity Chromatography (IMAC)
        3.2.4. HILIC
        3.2.5. IEF
        3.2.6. RP-LC
4. Orthogonal Validation
    4.1. Isotope Labeling
        4.1.1. Stable istope labeling with amino acids in culture (SILAC)
        4.1.2. Heavy methyl-SILAC (hmSILAC)
        4.1.3. Isomethionine methyl-SILAC
        4.1.4. Methylation by Isotope Labeled SAM (MILS)

Received 26 April 2021; Revised 21 June 2021; Accepted 22 June 2021; Available online 29 June 2021

4.2. SAM-analogue labeling
5. Other considerations
5.1. Methylation cross-talk with other modifications
5.2. Application of proteogenomics
5.3. Bioinformatics Tools
6. Conclusion

1. Introduction

As the field of proteomics grows and evolves, it is becoming ever more evident that the proteome of individual cells, irrespective of their background, is spatially, temporally, and chemically dynamic. Moreover, a cell can assign different roles to a single protein through chemical modifications, in order to fulfill different functions. Functional proteomics is thus becoming less about the changes in total protein and more about the changes in the different forms of a protein, contributing towards a better understanding of cellular biology under both healthy and disease conditions[1].

Such chemical modifications on proteins are called post-translational modifications (PTMs) and these can be described collectively as epiproteomics. They are considered as such because they are added to proteins in response to an environmental context such as the cell type in which the protein is expressed, as a result of signalling, following a stressful condition, due to interaction with the micro-environment and so on. Regardless of the origin, PTMs lead to some type of alteration in one of the protein properties, and by consequent, its function. There are over 200 known PTMs but some of the most commonly found within the proteome are phosphorylation, ubiquitination, acetylation and methylation[1].

1.1. Protein Methylation

Methylation involves the addition of one or more methyl groups to an amino acid residue within a protein. In these reactions a methyl group is transferred from a donor, which is usually S-adenosyl-L-methionine (SAM or AdoMet) to the acceptor amino acid by a class of enzymes called protein methyltransferases[2–4].

Methyl groups are most frequently found on arginine and lysine residues, but have also been reported on the side chains of histidine, glutamic acid, glutamine, aspartic acid, asparagine, cysteine and methionine[5], as well as on protein N- and C-termini[6,7]. Lysine can undergo mono- (Kme1), di- (Kme2) or tri- (Kme3) methylation on its epsilon (ε) amine, while arginine can be mono- (Rme1), asymmetric di- (aRme2) or symmetric di- (sRme2) methylated on the guanidinium moiety[1,8,9]. Protein methylation can also be sub-divided by acceptor element (i.e. the atom to which the methyl group attaches) into N-, O- and S-methylation, such that N-methylation occurs on lysine, arginine, histidine, alanine, proline, glutamine, phenylalanine, asparagine, and methionine residues, O-methylation occurs on aspartic acid, glutamic acid and cysteine residues, while S-methylation occurs on methionine and cysteine residues[2,10].

Very simplistically, protein methyltransferases are generally classified into protein lysine methyltransferases (PKMTs) and protein arginine methyltransferases (PRMTs), although there are numerous protein methyltransferases that fall outside such a grouping. PKMTs thus catalyse mono-, di- and tri- methylation on lysine residues while PRMTs catalyse mono-, asymmetric or symmetric di-methylation on arginine residues[8,11,12]. Protein methyltransferases can also be classified based on their structural domains, with all PRMTs being seven-β-strand methyltransferases (Class I), while most PKMTs are Su(var)3-9, Enhancer of Zeste, Trithorax (SET) domain methyltransferases (Class V)[13,14]. Other PKMTs fall within Class I as in the case of Disruptor Of Telomeric silencing 1-Like (DOT1L) and Methyltransferase-like (METTL) family members[2,15].

Methylation of a residue brings about an increased residue bulkiness (depending on how many methyl groups are added) and hydrophobicity (due to removal of the amino hydrogens) but minimal change in the overall charge (about 0.5 pKa units) or pI of the protein. This results in small but significant changes in protein structure and decreased H-bonding ability[16,17], consequently affecting function, by altering protein stability or half-life, sub-cellular localisation, protein interactions (with other proteins, DNA promoters or target molecules) with consequences on protein activity as well as altering PTM cross-talk[18,19]. This aspect of cross-talk is particularly complex when it comes to lysine because it is not only related to the possibility of lysine methylation interacting with other PTMs on adjacent residues but also to competition through its ability to undergo multiple PTMs mainly methylation, acetylation, ubiquitination, sumoylation, and glycation, further adding another level of regulation[2]. This is exemplified in the tumour suppressor p53, where different positions or degrees of lysine methylation alter a number of p53 protein properties such as suppression of gene transcription and increased affinity for 53BP1[20].

Thus through the addition of such methylations numerous important biological processes such as gene activation or repression, transcriptional regulation, RNA processing, DNA repair and signal transduction, are regulated[11,21]. These changes are significant to both healthy human cell biology and disease, but their study through comprehensive proteomics, particularly for clinical applications is still a research area in its infancy[22].

2. Mass spectrometry

Mass spectrometry is deemed one of the focal tools in comprehensive proteomics, thanks to its sensitivity and versatility, with a capacity to identify target proteins even within highly complex sample matrices. The approach most commonly adopted (and most relevant to this discussion) is bottom-up proteomics, also referred to as "Shotgun Proteomics" where the proteins in the sample are digested enzymatically (e.g. by trpsin, chymotrypsin, Lys-C, Lys-N) or chemically[23], the peptides generated (0.8-3kDa) are separated by liquid chromatorgraphy (LC) and then identified by tandem mass spectrometry (MS/MS) to determine the peptide masses through their fragmentation ions, hence the workflow being defined as LC-MS/MS. Software is then used to identify the parent proteins in the sample by comparing the parent (precursor) and daughter (derived fragmentation) ion masses generated experimentally against a database of theoretical peptide ion masses. Such software parameters can be altered to also include a search for a number of PTMs including the various methylations. The inclusion of a PTM on an amino acid residue results in a spe-

cific mass shift, which can be detected in an MS/MS, meaning that the different degrees of methylation should be unambiguously detected, though this is not always the case, due to multiple reasons (explored in Section 7). Interestingly, although mass spectral analysis in MS2 can distinguish between aRme2 and sRme2 through specific fragment ions, due to their identical mass, for convenience, most proteomic analyses consider them as a single form[24,25].

The development and refinement of the Orbitrap, which is a type of ion trap mass analyser consisting of three electrodes (two outer cup-shaped electrodes forming a barrel and a central spindle-like electrode) that oscillate ions in a harmonic spiral motion around the central spindle[26], has greatly improved the achievable resolving power and mass accuracy. These properties are critical considering the low abundance of peptides harbouring PTMs in comparison to the abundance of the total unmodified parent ions.

Despite the popularity of bottom-up proteomics identifying peptide mixtures, intact proteins have also been analysed by mass spectrometry in order to identify methylation sites on individual proteins of interest in what is described as a top-down approach. One of the major advantages of top-down is the complete coverage of the protein sequence, which effectively allows the detection of all the mass shifts associated with the PTMs of interest[27]. However, major drawbacks include the inadequacy for proteins greater than 50 kDa and reduced sensitivity compared to the bottom-up approach[28].

It is also possible to reach a compromise between these two approaches by performing enzymatic (e.g. by Glu-C, Asp-N, OmpT, Sap9) or chemical digests which produce few cleavages within target proteins[23]. This is referred to as a 'middle-down' approach, in which long peptides (2.5-10kDa) undergo analysis by mass spectrometry. In so doing the ability to detect combinations of PTMs in the top-down approach is coupled to the higher sensitivity, through greater ionisation efficiency of peptides over whole proteins, of the bottom-up approach[28].

A number of high-throughput methylation studies have been performed over the past decade by applying modifications to the basic LC-MS/MS workflow, resulting in the identification of hundreds of lysine and arginine methylations throughout the human proteome, in both healthy and disease samples[25,29–38]. Unfortunately though, only a small proportion of the total methylated residues have ever been isolated. For example, it has been estimated that 0.7-1.0% of all arginine residues in proteins are methylated[39]. Moreover, the methyltransferases responsible for a significant proportion of these methylations have not been identified yet[15].

In order to compensate for the low (sub-stoichiometric) abundance of methylated peptides in a sample, most methylation analyses by LC-MS/MS are often preceded by a range of enrichment procedures in order to improve the likelihood of selection for MS/MS analysis and detection sensitivity for methyl-peptides, which will be detailed in Section 3. The digestion step may be included either before or after the enrichment phase, depending on the type of technique used to enrich for methylation i.e. whether targetting the structure or domains of proteins (based on interactions) as opposed to targetting peptide sequences.

2.1. Effect of methylation on enzyme digestion

The most commonly used enzymee for protein digestion in bottom-up proteomic workflows is trypsin, which cleaves peptide bonds at the C-terminus of arginine or lysine residues (except when arginine and lysine are followed by proline)[28,40]. Thus upon complete digestion of a protein, each tryptic peptide would have one positively charged arginine or lysine residue at the C-terminus, as well as the positively charged N-terminus, such that, with the exception of peptides with histidine residues, all tryptic peptides carry two positive charges under low pH conditions[40].

However, when methylations are present on arginine or lysine residues the ability of trypsin to cleave at said sites is somewhat inhibited such that it was experimentally shown that following 16h incubation with trypsin, less than 30% of the mono-methylated peptides were cleaved, while no di- or tri-methylated peptides were cleaved[41]. This would result in the incomplete digestion of most methylated tryptic peptides and the inclusion of a miscleavage site and an extra positive charge. Nevertheless, methylation events on tryptic peptides with C-terminal arginine or lysine residues are frequently identified in mass spectrometric analyses.

In order to reduce the proportion of miscleavages in a sample, multiple proteases with the same cleavage specificity such as trypsin, Lys-C (cleaves at the C-terminus of lysine) and Arg-C (cleaves at the C-terminus of arginine) may be employed to enhance the efficiency of protein digestion and decrease the number of unmethylated miscleaved sites. In fact, the combination of trypsin, Lys-C and Arg-C was shown to reduce the proportion of miscleaved peptides from 14.3% to 4.8%[42].

The use of proteases with alternate cleavage site specificity may also be useful. For example, since chymotrypsin cleaves at the C-terminus of phenylalanine, tryptophan or tyrosine, it generatdes unique peptides particularly relevant for arginine methylation with respect to RGG-repeat regions, which give an identification advantage over trypsin, which often produces peptides that are too short to be assigned to specific RGG-repeat proteins. Moreover, since trypsin is inefficient at cleaving methylated arginine residues, chymotrypsin can be used as an alternative to improve the relative quantification of methylated peptides with respect to their corresponding unmethylated counterparts[36].

When proteins were digested with trypsin, chymotrypsin and Glu-C (cleaves at the C-terminus of either aspartic acid or glutamic acid) prior to mass spectrometric identification, it was observed that 2147, 497 and 274 methylation sites were identified respectively, with 355 methylation sites only identified through the latter two enzymes and not isolated using trypsin[31].

2.2. Fragmentation of methylated peptides and position identification

When analysing complex mixtures of digested peptides, the identification of the methylation type and position on a peptide is challenging. The MS/MS spectra are used to identify the peptide backbone fragments which carry mass shifts characteristic of methylation. Shotgun proteomics can make use of four different peptide fragmentation strategies in order to generate ions and collect peptide information. These are: Collission-Induced Dissociation (CID), Higher-energy Collision-Induced Dissociation (HCD), Electron Capture Dissociation (ECD) and Electron Transfer Dissociation (ETD). However, the ones of most relevance to methylation analysis are CID and ETD.

In CID, peptide ions are activated via multiple collisions with the inert gas, such as helium. As a result of these collisions, the weakest bonds within a peptide, irrispective whether in the peptide backbone or in the side chains (including PTMs), are cleaved preferentially[43]. CID typically generates y- and b-type fragment ions[44], with a preference for cleavages near aspartic acid, glutamic acid and proline residues, with abundant internal fragmentation and frequent losses of PTMs[43]. Unfortunately, the prominence of ions resulting from neutral losses often dominate MS/MS spectra, such that the resulting informative peptide backbone fragments detected tend to be few[45–48].

This means that methyl groups on the side chains of peptides are easily lost under CID conditions because they are highly labile[47]. Nevertheless, CID has been successfully used to localise methylation on lysine, arginine, aspartic acid and glutamic acid residues[48–51]. CID spectra of methylated peptides generate losses which are mostly water[52], although other abundant low-mass losses include methylamine, methylguanidine, or methylcarbodiimide[45,48,53].

In ECD, a thermal electron is captured and the energy released in this exothermic process achieves a fast (non-ergodic) fragmentation without prior internal redistribution of energy, while in ETD (which is a variant of ECD) a radical anion is used as the electron-transfer species, since electrons cannot easily be trapped with peptide ions in radio frequency (RF) traps[43,54].

The gas-phase radicals generated in ETD efficiently fragment peptide N-Cα bonds, mainly at the sites where the radical was initially trapped[55], independently of the sequence of peptides (unlike CID or HCD), yielding c- and z-type ion fragments while leaving side-chains with labile PTMs largely intact[43,54,56,57]. However, these electron-based fragmentation techniques require the presence of multiply charged peptide ions[58]. In ETD, methylations are well-preserved on c and z ions, enabling sequence identification and exact localisation of the methylations on the peptides[47,59].

Compared to CID, ETD generates a more complete series of ions, being particularly effective for highly charged peptides[46]. The spectra generated by ETD present extensive backbone fragmentation, providing more extensive sequence information, while leaving the labile PTMs intact, allowing more confident assignment of the methylation sites[43,46,60].

The spectra produced by both CID and ETD can be useful for identifying and validating methylations on a peptide through the presence of methylation-specific ion signals in the low mass range resulting from neutral losses, which are unique among the peptides generated from the proteome[61] . Together with other peptide backbone fragments these characteristic ions can be used to identify, validate and localise the methylation sites (Table 1).

Table 1. Characteristic ions and neutral losses associated with methylated peptides.

Amino Acid	Methylation Type	Mass Shift (Da)	Neutral Loss (Da)	Immonium ions (m/z)
Arginine	Rme1	14.0156	31.042 (mono-methylamine); 73.064 (mono-methylguanidine); 74.0718 (monomethylguanidinium); 31.0422 (monomethylamine); 73.0640 (monomethylguanidine)	143 (not unique for arginine)
	aRme2	28.0312	45.0578 (di-methylamine); 87.087 (di-methylguanidine): 46.0657 (dimethylammonium); 71.0609 (dimethylcarbodiimidium); 45.0578 (dimethylamine)	71.06
	sRme2	28.0312	31.0422 (mono-methylamine); 87.087; 71.0609 (dimethylcarbodiimidium); 31.0422 (monomethylamine); 70.0531 (dimethylcarbodiimide)	71.06
Lysine	Kme1	14.0156		98.096 84.081 (C-terminal lysine) 101 (N-terminal lysine); 98.0970 (immonium ion)
	Kme2	28.0312		98.096 112.1 84.081; 98.0970 (immonium ion); 112.1126 (immonium ion)
	Kme3	42.0470	59.0735; 59.0735 (trimethylamine)	

For arginine methylation, CID fragmentation can be useful in providing information on peptide sequence and to characterise the type of methyl arginine (i.e., Rme1, sRme2, or aRme2)[47], despite the loss of labile side-chain fragments[46]. On the other hand, ETD is more efficient than CID in the identification of arginine methylated peptides because it produces better fragmentation whilst retaining side-chains, especially since such peptides usually carry more positive charges[47]. Some side-chain fragmentation still occurs in ETD leading to neutral losses which allows the determination of the type of methylation on the basis of the ETD spectra alone[46]. Nevertheless, determination of the arginine methylation type by ETD may still be difficult[46].

A number of characteristic ions can be used to distinguish between the various types of arginine methylation. Protonated Rme1 peptides have neutral losses of monomethylguanidine (73 Da) and monomethylcarbodiimide (56 Da)[24,47,50,53,62], while despite aRme2 and sRme2 present no mass difference, characteristic fragment ions can be used to distinguish between the two forms using either CID or ETD fragmentation[24].

In CID, protonated sRme2 peptides undergo characteristic neutral losses of monomethylamine (31 Da) and dimethylcarbodiimide (70 Da) as opposed to aRme2 peptides which present neutral losses of dimethylamine (45 Da) and dimethylguanidine (87 Da)[45]. The neutral loss of dimethylamine (45 Da) is produced from both forms of Rme2 but predominantly indicates the presence of aRme2 rather than sRme2[46].

In ETD fragmentation, neutral losses from charge-reduced precursor ions result in highly abundant low-mass product ions which allow for the reliable discrimination of aRme2 and sRme2[45,46,53] (Table 1), with a neutral loss of monomethylamine (31 Da) being characteristic of Rme1 and sRme2, dimethylamine (45 Da) characteristic of aRme2, and dimethylcarbodiimidium (70 Da) characteristic of both aRme2 and sRme2[45,46,48].

Precursor ion scanning is also used to identify the type of arginine methylation since the dimethylammonium ion (46 Da) and dimethylcarbodiimidium ion (71 Da) side-chain fragments of Rme2 can be used as characteristic ions for arginine methylation[24,46,48,53].

As an alternative, a negative ion mode approach was proposed to provide a more reliable identification of the type of arginine methylation compared to the standard positive ion mode, especially for arginine methylated peptides harbouring additional basic residues, which cause further suppression of the protonated methylation-specific fragmentation due to sequestration of the external proton[62].

For lysine methylation, a number of characteristic ions can be used to distinguish between Kme1 and Kme2 or between the different forms of lysine methylation and acetylation. The characteristic immonium ion is produced by both Kme1 (m/z 98) and Kme2 (m/z 112)[63] making it possible to distinguish between these two degrees of lysine methylation using precursor ion scanning[4] as well as allowing the use of immonium ion scanning with different collision energies to improve the discovery rate of lysine methylations[64].

Through the characteristic fragment ions at m/z 98 and 115 displayed by Kme1 and the characteristic fragment ion at m/z 129 displayed by Kme2 it is possible to clearly distinguish them from acetylation, which produces characteristic fragment ions at m/z 143 and 126. Similarly, the presence of characteristic ions and neutral losses in MS/MS spectra of Kme3 (+42.05 Da) allow discrimination from acetylation (+42.01 Da), despite the mass difference between these two modifications being very small (0.0364 Da) such that it can only be discriminated within <30 ppm mass accuracy[63,65] (Table 1).

Thus in comprehensive methylation analyses, the use of CID alone does not provide sufficient spectral information for identification, especially for arginine methylated peptides which undergo side-chain fragmentation, whilst being much less relevant for lysine methylated peptides. The use of ETD fragmentation mostly avoids side-chain fragmentation and the neutral losses generated for arginine methylated peptides, providing sufficient amount of backbone fragmentation to enable confident peptide sequence determination, localisation of the methyl groups and assignment of the degree of methylation on the basis of the ETD spectra alone[46].

As the majority of methylated peptides under reversed-phase conditions are expected to carry a charge ≥ 3, ETD enables the assignment of methylation localisations for a wide variety of peptides, not possible using just CID spectra, through increased backbone cleavage and better MS/MS spectral data compared to CID fragmentation[46]. Despite this, the combined use of the ETD and CID fragmentation methods produces two complementary, independent and orthogonal datasets within the timeframe of a single LC run, obtaining information from differentially charged precursor ions, which significantly increases the confidence of sequencing and localising methylations[43].

Further to the use of neutral losses and immonium ions for identification, other methylation-specific, MS/MS-based strategies including precursor ion scanning[24,64] and multistage MS/MS[66] have been successfully applied[65]. However, the application of the methylation-specific ions to validate the identification of methylated peptides in high throughput experiments requires further investigation.

2.3. Fractionation by FAIMS

Field Asymmetric Ion Mobility Spectrometry (FAIMS) is a type of ion mobility spectrometry (IMS) that separates gaseous ions based on their mobility in an asymmetric waveform which generates an oscillating electric field between a high and low electric field[67,68]. The FAIMS device is inserted after the electrospray emitter such that the ions are carried by an inert gas such as nitrogen and/or helium through a gap/channel and oscillate between two electrodes to which the high voltage asymmetric radio frequency field is applied[69]. The maximum peak amplitude of the asymmetric waveform is called the dispersion voltage (DV). Ions move within the electric field, towards the electrodes, at a velocity based on their mobility resulting in a drift in a particular direction. By applying a DC voltage, called the compensation voltage (CV), the trajectory of the ions can be altered resulting in a net drift towards one of the electrodes. In this way a set of ions can be selected, allowing them to pass through the FAIMS device instead of being neutralised on the electrode. Thus by altering the CV voltage, population of ions having different mobilities can be separated, focused and transmitted into the ion trap of the mass spectrometer[70,71]. This allows the consecutive analysis of numerous unique fractions to be analysed over the course of an single experimental run, with a comparable depth of analysis to conventional two-dimensional LC strategies[67].

Using the tryptic digests of the human cell line K562, it was shown that for a single shot 4 hour run, 7818 proteins could be quantified with FAIMS compared to 6809 proteins without FAIMS. Furthermore, from a single shot 6 hour run using FAIMS, 8007 proteins were identified, while from an LC fractionation experiment consisting of four fractions analysed for 1.5 hours each, 7776 proteins were identified[67]. It was reported that a single shot 4 hour run with or without the application of FAIMS generated 815196 and 444965 respectively, effectively almost doubling the number of precursor ions generated by using FAIMS. Concomitantly, for single shot analyses, peptide identifications was increased up to 2-fold and protein identification increased up to 55%, depending on the duration of the run[67]. FAIMS has also been combined with SILAC in HEK293 cells, which through 16 replicate LC–MS/MS analyses identified 10574 SILAC pairs as opposed to just 3519 SILAC pairs without FAIMS. We identified 3257 pairs with both techniques, whereas 7317 and 262 pairs were uniquely detected with and without FAIMS, respectively[71].

FAIMS can thus be used as a separation technique for peptide analysis, orthogonal to mass spectrometry, improving the depth of proteome analysis as well as allowing peptide isomers and PTM sites to be resolved[71–74]. By combining FAIMS with LC-MS/MS, the transmission of multiply charged ions can be enhanced and the detection of low abundance peptide ions improved by increasing the signal-to-noise ratio[75–77], as well as reducing contaminants and increasing the score assigned to correct sequences, which reduces the FDR[71]. The application of FAIMS could in effect act as a replacement to LC fractionation for many applications[67]. Moreover since separation by FAIMS does not only rely on m/z ratio, it can be useful in the separation of isobaric species as in the case of methylation analysis[67]. FAIMS has displayed the complete separation of isobaric peptides with different degrees of methylation on peptides up to around 3 kDa, where the methyl moiety makes up just 0.5% of the peptide mass[78,79].

The application of FAIMS to separate isobaric Rme peptides was shown through the use of six methylation combinations on a model peptide con-

taining three RGG motifs, generating three Rme1 and three Rme2 peptides, making up two sets of isobaric Rme peptides. Under the experimental conditions used, each of the six peptide ions in the the +3 and +4 states were produced in relatively high abundances and generated FAIMS-MS spectra which could be used to separate each of the variants[79]. However, it was noted that the sites and degrees of methylation have an effect on the required He concentration in the carrier gas and the electric field used for ion transmission in order to successfully separate isobaric methylated peptides in a mixture[79].

FAIMS was also used to successfully discern the methylation positions on two adjacent lysine residues in isobaric peptides of human histone H3 tails. The first peptide pair was monomethylated at K4 or K9, while the second peptide pair was dimethylated at K27 or K36. A range of He (0–40%) or H_2 (50–70%) concentrations in the carrier gas were tested and it was found that resolution increased at higher He or H_2 concentrations and was comparable between the two gases at similar concentrations[78].

3. Methods for methylation enrichment

Although the resolution of LC-MS/MS has improved significantly thanks to numerous technological advances, identifying differentially expressed protein methylations from biological samples is still a challenging process. When performing high-throughput experiments, it is highly likely that since the expression of methylated proteins may be low, the methylated peptides derived from them would be in low stoichiometric amounts, and thus missed in the process of fragmentation[80]. Provided such methylated peptides are fragmented for MS/MS, the low abundance of their parent ions could result in fragmentation spectra which are of poor quality, and as a result the ability to determine the positions of such methylated residues is reduced. To overcome this issue, most protein methylation studies apply an enrichment technique in order to submit a larger amount of methylated peptides for analysis[28].

3.1. Antibody-based enrichment

3.1.1. Immunoaffinity purification

The use of antibodies against Kme1, Kme2, Kme3, Rme1, aRme2 or sRme2 motifs to enrich for methylations, whether on intact proteins and complexes or digested peptides, defined as immunoaffinity purification (IAP), prior to a mass spectrometry workflow (IAP LC-MS/MS) has been widely used in order to identify and quantify methylation sites on arginine and lysine, in a variety of human samples[32,81].

A number of antibodies targetting Rme1, aRme2 or sRme2 motifs have been designed by various labs[32,48,81,82], while other groups have preferred using commercially available options such as PTMScan motif antibodies[33,83]. The IAP LC-MS/MS workflow has allowed the identification of 1000–1700 arginine methylation sites from the cervical cancer cell line HeLa[36] or the colorectal cancer cell line HCT116, specifically 1000 Rme1 sites and 300–400 aRme2 sites[84]. The analysis of Jurkat T-cells and primary T-cells between them yielded 2502 unique arginine methylation sites derived from 1257 unique proteins[31].

While relatively successful for arginine methylation, IAP LC-MS/MS unfortunately on average does not identify more than 100-500 lysine methylation sites. For example, Guo et al. [32] identified a total 165 lysine methylation sites from the HCT116 cell line, specifically 132 Kme1, 35 Kme2, and 31 Kme3 sites. Slightly better, Cao et al.[30] identified a total of 552 lysine methylation sites from 493 peptides in 413 distinct proteins from the HeLa cell line, specifically 323 Kme1, 127 Kme2 and 102 Kme3 sites.

A number of reasons have been proposed for this from the non-selectivity and low binding efficiency of anti-pan Kme antibodies[29,48,85] to the possibility that lysine methylation as a PTM is limited to a small subset of proteins[32], supported by similarly low numbers isolated through chemical enrichment methods[86].

As an alternative to using antibodies, it is possible to enrich using domains which recognise methylated sites. As an alternative to using antibodies, it is possible to enrich using naturally-occurring domains which recognise methylated sites presenting a combination of broad sequence specificity and high methyl selectivity[87]. One such affinity reagent is composed of three repeats of the malignant brain tumor domain (3xMBT) of Lethal (3) Malignant Brain Tumor-Like Protein 1 (L3MBTL1), which can recognise and thus isolate peptides containing Kme1 and Kme2, independent of the surrounding protein sequence[88]. However, while it has been shown that 3xMBT does not bind to non-methylated peptides, it does bind weakly to some trimethylated peptides[34]. Moreover, to eliminate false positives due to non-specific binding, the D355N mutant of the 3xMBT was used as a negative control[87]. The 3xMBT has been used to isolate potential G9a-methylated substrates using protein arrays[85] and HEK293T cells[34], where 544 methylated proteins were identified with 313 enriched 2-fold. This approach was somewhat expanded upon by[89], by looking at the binding-affinity of a variety of chromatin-binding modules (CBMs) using permutation peptide arrays (where every residue in the Kme-containing peptide is substituted) followed by quantification of the binding affinity to the peptides in the permutation array to generate a scoring matrix. Upon determining the specificity of a CBM, the sites of lysine methylation that can be isolated using such a CBM from a pool of proteins can be predicted[89]. However, the application of such domains is restricted by the recognition specificity of the domain for the residues flanking the methylation site.

3.1.2. Chemical Derivatisation

Chemical derivatisation of the monomethylated ε-amine group of the lysine residues in the protein lysates by treatment with propionic anhydride, which produced propionyl-methyl ε-amines. This was followed by affinity enrichment using a pan anti-propionyl mono-methyl lysine antibody to isolate the propionylated Kme1 peptides, which resulted in the identification of 446 Kme1 sites in 398 proteins from a combination of five human cancer cell lines (HeLa, K562, SW620, A549 and SMM7721) and human liver cancer tissue samples[37]. Propionylation increases the structural size of mono-methylated lysine, increasing the affinity between antigen and antibody as well as increasing the mass shift observed, since propionyl mono-methylation has a mass shift equivalent to that of lysine butyrylation. However, despite this success, such a strategy is not amenable to

Kme2 and Kme3-containing peptides due to the difficulty of chemically derivatising di- and tri-methylated lysines[37].

It is also possible to greatly reduce the interference of non-methylated lysine and arginine in miscleaved peptides by derivatising their side chains to suppress the positive charges of side chains primary amine groups of unmethylated arginine and lysine residues and free N-terminal amines of peptides, while unaffecting their methylated counterparts, allowing methylated lysine and arginine side chains to retain their charges. This was achieved by sequentially treating the tryptic peptide mixture with malondialdehyde (MDA) and ortho-phthalaldehyde (OPA), which block most functional groups carrying positive charges within the peptide digest, whilst mainly retaining a positive charge on methylated peptides at neutral to basic pH. Any non-methylated peptides non-specifically purified by affinity methods would have difficulty ionising with no charges and thus reduce their likelihood of being identified in the subsequent mass spectrometric analysis. Followed by SCX fractionation, this resulted in 399 lysine, 240 arginine and 103 histidine methylated peptides being identified from HEK293 cells. As a note, this method does not block histidine, which is positively charged and as a result there is an enrichment of histidine-containing peptides. Furthermore, the esterification side reactions should be minimised to avoid false-positive identification of mono-, di- or tri-methylation[90].

3.2. Antibody-free enrichment - chemical affinity methods

Since, enrichment by both antibodies and domains relies on sequence affinity, to overcome the limitation of sequence selectivity, enrichment techniques based on chemical properties unique to methylated peptides may be preferred. Furthermore, the use of chemical affinity techniques to isolate methylated peptides is preferred over the use of chemical derivatisation methods since the latter tend to inevitably produce side reactions, which then require the inclusion of additional steps in the workflow in order to distinguish these products from the real methylations present *in vivo*[91].

However, methylation only bring about small changes in the physicochemical properties of peptides, limiting the strategies that can be employed to develop efficient enrichment approaches based on chemical properties. That being said, peptides obtained from trypsin digestion containing Kme and Rme are highly basic and hydrophilic. As a result it is possible to enrich for methylation using Strong Cation Exchange (SCX) chromatography, Hydrophilic Interaction Liquid Chromatography (HILIC), Isoelectric Focusing (IEF) or Reverse-Phase Liquid Chromatography (RP-LC)[25,29,31,33].

3.2.1. Low-pH SCX chromatography

SCX chromatography separates peptides by single charge differences, based on the pH of the sample solution, which affects the protonation and deprotonation of basic and acidic groups on the peptides. Most tryptic peptides (around 70%) in a sample have a net charge $\geq$+2 at pH 2.7 (two terminal positive charges and positive charges at miscleaved sites), and the higher the number of charges, the greater the strength of binding to and retention on the negatively charged SCX resin[92]. Methylation of lysine and arginine residues reduces the cleavage efficiency of trypsin, producing miscleavage sites[41] (described in Section 2.1), which increase the overall number of positive charges on such peptides to $\geq$ +3 (at pH 2.7), allowing for methylated peptides to be enriched by low-pH SCX chromatography in the $\geq$ +3 fractions[25]. Apart from the electrostatic interactions, hydrophilic and hydrophobic interactions between the tryptic peptides and the SCX resin also play a role, and as a result peptides having the same net charge can still elute differently. Low-pH SCX chromatography can therefore be used to reduce the sample complexity, and concomitantly improve the detection sensitivity and identification success rate for methylated peptides[41]. Although SCX fractionation can be used in isolation, it can also be performed prior to or even following enrichment by immunoaffinity[93,94]. Using low-pH fractionation 860 arginine and lysine methylations on 765 sites were identified from BEL cells, namely 276 Rme1, 350 Rme1, 78 Kme1, 82 Kme2 and 74 Kme3 i.e. 27.21% were lysine methylation forms. The number of methylated peptides in the SCX fractions increased with the increasing elution strength, as the number of positive charges increased due to miscleavages. Compared to direct analysis of the peptide digest, which yielded 85 methylated peptides, SCX fractionation resulted in the identification of 217 unique methylated peptides, meaning a 2.5 times higher yield[40].

3.2.2. High pH SCX chromatography

When using low-pH SCX enrichment, histidine-containing peptides interfere with the identification of methylated peptides, with between 60–70 % of peptides within the enriched methylation fractions actually being histidine-containing peptides, which also have a charge $\geq$ +3 (at pH 2.7) but do not include methylated residues[25]. This is because each histidine provides an additional positive charge at low pH, resulting in peptides with more than two positive charges overall, which consequently bind strongly to the SCX resin. To solve the issue of interference from histidine-containing peptides, high-pH SCX was proposed, the rationale being that at a solution pH higher than 9, the histidine residues would be mostly uncharged and thus could be depleted, allowing for improved detection and identification of methylated peptides[41].

The application of the high-pH SCX strategy resulted in the identification of 218 Rme1, 587 Rme2, 35 Kme1, 19 Kme2 and 28 Kme3 peptides[41]. However, lysine methylation accounted for just 9.24% of methylations. This low recovery was attributed to two reasons, firstly that *in vivo*, methylation on lysine is known to occur less frequently than on arginine, and secondly to the weak binding of lysine methylated peptides to the SCX resin at pH 9. This is primarily due to numerous acidic residues flanking methylated lysine residues, which induce strong electrostatic repulsion between these acidic residues and the sulfone groups in the SCX resin and barely allow these peptides to bind the resin. This issue of retention on the SCX resin is minimal at low pH due to suppressed ionisation[40].

3.2.3. Immobilised Metal Affinity Chromatography (IMAC)

Another option for depleting histidine-containing peptides in order to improve the identification of methylated peptides is copper-IMAC (Cu-IMAC) and to much a much lesser extent nickel-IMAC (Ni-IMAC). Whilst Ni-IMAC depleted histidine-containing peptides by 8.14% and resulted in the identification of 67 new methylated peptides (an increase in identification by 34.72%), the use of Cu-IMAC depleted histidine-containing peptides by 23.03% and resulted in the identification of 95 new methylated peptides (an increase in identification by 49.22%)[40]. Ni-IMAC is not as successful at depletion as Cu-IMAC because a large portion of histidine-containing peptides cannot bind the Ni-IMAC. With IMAC the reverse is also true and some methylated peptides are lost because they bind to the IMAC beads and end up being depleted[40].

3.2.4. HILIC

HILIC separates peptides by having the sample loaded at a high organic solvent concentration and the elution is achieved by increasing the polarity of the mobile phase, so as to function as the opposite strategy to RP-LC[95]. HILIC has been coupled with reverse-phase fractionation and proven to be quite successful to enrich highly hydrophilic peptides[96,97]. In fact, HILIC can be successfully applied for the enrichment of methylated peptides because a substantial proportion of tryptic peptides containing Rme are highly hydrophilic. In the identification of methylated peptides from T-cells, HILIC was found to be mmuch more effective than IEF or low pH SCX chromatography, resulting in the identification of 249 Rme sites in 131 proteins[25].

3.2.5. IEF

IEF separates peptides based on isolelectric points and in so doing, methylated peptides can be enriched due to the high isoelectric points they possess as a result of the extra charges brought about by miscleavages. Most Rme tryptic peptides are highly basic and IEF was used to enrich ≥ 5% of peptides containing Rme at a pH 9-11 (or greater) of the IEF gradient. However, the major limitation of this enrichment technique are IPG strips with a suitable pH range[25], since highly basic peptides would require strips which cover a pH range beyond pH 11 for effective enrichment.

3.2.6. RP-LC

RP-LC has been successfully coupled with SCX fractionation in order to reduce sample complexity and improve PTM identification, combining the power of separation of peptides by charge using SCX followed by separation by hydrophobicity using RP-LC[98]. It is also possible to replace SCX fractionation by low-pH RP-LC fractionation and then follow by high-pH RP-LC to produce a 2D system relying solely on RP-LC, which eliminates the need for sample desalting, reducing sample losses and processing time and thus improving peptide and protein identification[99]. As applied to methylation, the fractionation of tryptic peptides from HEK293 cells by RP-LC followed by affinity enrichment identified 8030 Rme sites from 3300 proteins[33].

4. Orthogonal Validation

As things currently stand, the major challenge in high-throughput analysis of protein methylation apart from specific enrichment of methylated peptides, is high confidence identification of such methylated peptides.

The identification of peptides in mass spectrometry is generally inferred from the degree of matching obtained between observed and predicted sequence-specific patterns. The target-decoy database search strategy has become the preferred method to do this in large-scale proteomics because of its high throughput and accuracy, accessiblity to any laboratory in terms of minimal computational resources required, and its use with any mass spectrometry system and any database-searching algorithm. The first step in building a target-decoy database is to generate a 'target' protein sequence database using a protein mixture appropriate for the sample to be analysed. Following this, a 'decoy' database is created to have the general composition of the target database, but containing the minimum number of peptide sequences in common with the target database. This is generally done by reversing the target protein sequences. From this, a False Discovery Rate (FDR) can be determined by doubling the percentage of peptide identifications obtained from the decoy database[100].

However, when it comes to correctly identifying methylated peptides, it was found that the target-decoy approach, whether applied globally or to peptide sub-groups, does not provide an accurate estimate of the FDR for methylation peptide spectrum matches (PSMs). The two most likely reasons proposed for this are: 1) the difference in mass between multiple pairs of amino acids being identical to that observed for the addition of a degree of methylation, that results in unmodified peptides being incorrectly identified as methylated peptides and 2) esterification reactions in sample preparation protocols using methanol or ethanol, which produce methylation or ethylation as artifacts, that in turn are misidentified as mono- or di-methylation respectively, on proximal arginine or lysine residues. Furthermore, it was found that wrongly assigned methylation of residues other than arginine or lysine to the latter only contribute minimally, whilst it is more common for methylations to be wrongly assigned between adjacent lysine or arginine residues[101].

FDRs for methyl-PSM were shown to greatly exceed the <1% FDRs estimated by the global target-decoy approach, typically exceeding 80%, irrespective of the sample preparation or mass spectrometer used. These findings suggest that to obtain reliable and sensitive methyl-PSMs in large-scale methylation experiments, orthogonal methylpeptide validation is required, with heavy-methyl SILAC (hmSILAC), or any of its variants being ideal for such validation[101]. It is possible to rescue low-scoring peptides by evaluating each identification manually, but is very inconvenient for large-scale studies[100] and software e.g. MethylQuant[102] has been designed to automate this process.

4.1. Istope Labeling

4.1.1. Stable istope labeling with amino acids in culture (SILAC)

SILAC involves culturing two cell populations, one population in culture medium containing all normal (light) amino acids and the second population in culture medium containing one or more stable isotope labeled (heavy) amino acids. Such heavy amino acid can contain ^{2}H instead of ^{1}H, ^{13}C instead of ^{12}C, or ^{15}N instead of ^{14}N. The incorporation of amino acids containing stable heavy isotopes into all newly synthesised proteins within a cell culture leads to a known mass shift within the peptides generated as compared to peptide containing the light version of the amino acid, without additional chemical changes or affecting cell morphology and growth. For incorporation of the heavy amino acid into all proteins of a cell population only essential amino acids are used for SILAC labeling and the cells have to be grown with dialysed serum. Complete incorporation takes a number

of cell doublings. Upon analysing the combined proteins from both cell population by mass spectrometry, pairs of identical peptide sequences can be identified presenting a definite mass difference due to the stable isotope inclusion, where the ratio of peak intensities in the mass spectrum for such peptide pairs reflects their abundance ratio (the difference in abundance) between the two cell populations. SILAC has been used for expression proteomics, for quantification of the proteome within the two SILAC populations, as well as for functional proteomics, where one cell population acts as a control, while the other population is treated in a specific way and the effect on the proteome is investigated[103].

4.1.2. Heavy methyl-SILAC (hmSILAC)

In the specific case of hmSILAC, the amino acid selected for heavy labeling is L-methionine because it is the precursor to SAM, which will act as the methyl donor for protein methylation. Culturing cells in media containing either unlabeled (light) or $^{13}CD_3$-labeled (heavy) L-methionine over multiple cell divisions will result in the complete incorporation of either light or heavy labeled methyl groups into all enzyme-generated methylation sites. Heavy labeling adds 4 Daltons (Da) for every methylation, resulting in a distinctive mass shift of 18.037 Da per methyl group[48].

By generating an experiment in which protein samples from unlabeled (light) cells and protein samples from $^{13}CD_3$-labeled (heavy) cells are analysed in parallel by high resolution LC-MS/MS, biologically real methylated peptides (i.e. peptides methylated enzymatically) can be identified from the resulting heavy/light peptide pairs (called methyl-SILAC pairs or spectral pairs). The characteristic mass shifts associated with either light or heavy labeling of methylated sites allows the identification of biologically-relevant methylations from potential false positives in an unambiguous and unbiased manner, as well as relative quantification of the methyl-SILAC pairs. Relative quantification can be achieved by comparing the measured relative signal intensities of methylated peptide pairs between the light and heavy samples[48]. The criteria set for strict validation are: 1) that light and heavy methylated peptides should be identified simultaneously, having closely matching elution profiles, and 2) that the extracted ion chromatogram (XIC) should present peaks of light and heavy peptides of very similar intensity ratios (ideally 1:1)[61].

The mass shift of 18 Da produced by heavy isotope methylations avoids false positive identification of such methylated sites by ignoring all those mass differences of 14 Da observed between numerous amino acids, which is the main source of false positive methylation identification. This in turn greatly boosts the confidence of identifying enzymatically generated methylations, making hM-SILAC more reliable than the traditional target-decoy approach that cannot effectively filter out the aforementioned false positive methylation sites[40,48]. Thanks to its ability to improve the confidence levels in the mass spectrometric identification of methylated peptides, hmSILAC has become the most commonly applied strategy for performing orthogonal validation for methylation studies[25,31,37,48].

However, hmSILAC is not without its drawbacks, which most notably include the high cost of the reagents and the labour intensive nature of the procedure. Furthermore, it should be noted that this method can only be applied to cell culture samples, but not to patient tissue samples. For this reason, the generation of patient-derived organoid cultures combined with hmSILAC might be the best option[104]. This would allow the complete labeling of a cellular system which more closely recapitulates the interactions of the proteome as expected of the *in vivo* situation.

The major technical limitation of hmSILAC is that all proteins synthesised in hmSILAC cultures also have heavy methionine incorporated, such that after enzymatic digestion methionine-containing peptides also present spectra pairs in the in precursor scans of methylation analyses. In order to heavy label methylation sites whilst avoiding methionine labeling, and allowing better discrimination between peptides from the two arms of a labeling experiment, isomethionine methyl-SILAC (iMethyl-SILAC)[31] and Methylation by Isotope-Labeled SAM (MILS)[105] have been developed.

4.1.3. Isomethionine methyl-SILAC

The iMethyl-SILAC technique was developed as a solution to eliminate the ambiguity resulting from methionine-containing peptides. This was achieved by replacing the light L-methionine with L-methionine-$^{13}C_4$, and combining this with the heavy $^{13}CD_3$-methionine workflow. In this way, methyl groups added to a protein during a methylation event would still be considered either light ($^{13}CH_3$) or heavy ($^{13}CD_3$), but either of the two types of stable isotope-labeled methionine used in protein synthesis are nearly isobaric (have nearly identical molecular weight) but differ in the distribution of the additional mass. The methyl groups between the heavy and light arms of the experiment would have a mass difference of 4 Da, such that methylated peptides would generate spectra pair with a 4Da difference for every methylation present on a peptide. As a result, methyl-SILAC spectral pairs are only generated for methylated peptides, contributing to reduce the FDR[31].

The iMethyl-SILAC strategy was combined with three different proteases (trypsin, chymotrypsin and GluC) and immunoaffinity purification using the anti-mono-methylated arginine antibodies developed by[32] and identified 2,502 arginine methylation sites in 1,257 proteins when using Jurkat T cells and primary T cells[31]. However, it should be noted that the misidentification of unmodified methionine-containing peptides as methylated peptides in a hmSILAC experiment is not a common occurrence and thus the utilisation of the iMethyl-SILAC strategy only reduces the FDR slightly[61].

4.1.4. Methylation by Isotope Labeled SAM (MILS)

The MILS technique involves the direct inclusion of isotope labeled SAM in the culture medium. In this set-up, the *E. coli* AdoMet synthetase MetK was used to generate SAM and $^{13}CD_3$-SAM from methionine and $^{13}CD_3$-methionine respectively. These methylation donors were then used in combination with an auxotrophic *S. cerevisiae* strain for SAM through the deletion of the SAM synthases genes *SAM1* and *SAM2*. Since SAM biosynthesis was completely abolished, proteins were exclusively labeled with CH3 or $^{13}CD_3$ depending on the SAM variant added to the culture. Upon mixing the proteomes of these two cell populations and analysing by mass spectrometry, MS/MS spectral pairs for methylated peptides were produced[105].

This technique has never having been used in a human cell study, probably due to the laborious requirement of having to delete the SAM synthase enzymes in order to exclusively label proteins with the exogenous isotope labeled SAM generated separately. Nevertheless it should be possible to reproduce the experimental procedure in a human cell line despite the difficulty of having to delete multiple gene copies when dealing with polyploid cancer cell lines.

4.2. SAM-analogue labeling

Another solution for orthogonal validation of methylation is by using SAM analogues which have a chemical moiety not found in nature as a PTM replacing the methyl group, in order to generate a unique signal or else to act as an entry point for further reactions. This method has been described as Bioorthogonal Profiling of Protein Methylation (BPPM)[106].

Through the use of click chemistry, a number of SAM analogues have been generated over the past decade in which the transferable chemical moiety includes a ketone[107] or alkene derivatives[106,108–111], among others. Following the addition of the SAM analogue by the methyltransferase enzyme, it is possible to go directly to mass spectrometry however an enrichment step for methylated peptides involving for example a copper-catalysed azide-alkyne cyclo-addition (CuAAC) would need to be performed to attach motifs such as biotin[109] or an azide-FLAG probe[110], using streptavidin-coated beads or immunoaffinity respectively.

The major disadvantage of a number of the developed approaches is that the methyltransferases used need to be modified in order to be able to use the modified SAM analogues and so are limited to *in vitro* systems. Interestingly, it has been shown that the point mutation M233G in PRMT3 makes this enzyme a promiscuous alkyltransferase that can use sp2-β-sulfonium-containing SAM analogues. Thus by mutating this conserved methionine alone or in combination with specific residues nearby, it could be possible to modulate the cofactor promiscuity of multiple type I PRMTs[84].

Moreover, although some SAM analogues can be efficiently used by methyltransferases found in nature, there is still a restriction to certain methyltransferases, as was reported for propargyl SAM which could be used by SETDB1 but not SET7/9, SMYD2, PRMT4, or PRMT1[110]. One option might be to use Se-alkyl Se-adenosyl-l-selenomethionine (SeAM) analogues, which upon comparison with their equivalent SAM analogues indicated that the sulfonium-to-selenonium substitution can enhance the compatibility these SeAMs with certain protein methyltransferases, allowing for the use of otherwise less reactive SAM analogues[112].

In vivo protein allylation has been only performed in *S. cerevisiae* to capture protein methylation candidates by combining this derivatisation with chemical tagging and affinity enrichment[113]. Thus, while *in vivo* enrichment of methylation using strategies involving SAM analogues is possible, there is still a long way forward before these become a staple in methylation tagging and enrichment.

5. Other considerations

Despite the development of multiple strategies to improve the reliability of methylation analysis, the availability of reliable methylation-directed chemistry in not enough. There are numerous other facets of large-scale protein methylation studies that need to be considered. Although these are beyond the scope of the current review it is important to appreciate the weighting these have on the outcome of such a study.

5.1. Methylation cross-talk with other modifications

Methylations can interact with other PTMs in two ways, either by direct competition for the same residue (as in methylation, acetylation or ubiquitination on the same lysine) or by positive or negative interaction on neighbouring residues[18]. When considering proximal interaction, this could be it either in relation to the sequence or else in 3D space once the protein is folded. Furthermore, the interactions encompassed in this category would be between methylation and any other PTM within a certain distance, which could either be impacted sterically or chemically.

This means that in a proteome-wide enrichment study for any other PTM, there is a potential for indirect enrichment of methylations due to sequence proximity cross-talk. In fact, methylation sites can be regularly identified following isolation of peptides enriched for PTMs such as acetylation or phosphorylation. The presence of methylation together with another PTM could indicate a synergistic interaction while the isolation of peptides with mutually exclusive PTMs could indicate an antagonistic interaction.

The cross-talk between methylations on different residues and with other PTMs including acetylation, phosphorylation and ubiquitination has been described for a number of proteins. For example, methylations interacting with each other have been described in p53, where mono-methylation at K372 (by SET9) inhibits mono-methylation at K370 (by SMYD2), thereby regulating transcription activity of p53[114]. An example of methylation and acetylation cross-talk was described in Estrogen receptor alpha (ERα), where mono-methylation at K266 (by SMYD2) represses chromatin recruitment of ERα and target gene activation, while acetylation at the same site promotes ERα transactivation activity[115]. The cross-talk between methylation and ubiquitylation was shown in RNA binding protein 15 (RBM15), where asymmetric di-methylation at residue R578 (by PRMT1) triggers protein degradation by the E3 ligase CCR4-NOT Transcription Complex Subunit 4 (CNOT4)[116].

Similarly, the cross-talk between methylation and phosphorylation has been shown in a number of proteins. Antagonistic interaction is observed in Flap Endonuclease 1 (FEN1), where asymmetrical di-methylation at R192 (by PRMT5) suppresses FEN1 phosphorylation at S187 (by Cdk2–cyclin E), such that once methylated (but not phosphorylated) FEN1 interacts with proliferating cell nuclear antigen (PCNA), ensuring correct localisation and DNA replication or repair[117]. Similarly, in Sex determining region Y-box 2 (SOX2), mono-methylation at K119 (by SET7) inhibits the transcription activity of SOX2 bringing about its ubiquitination by the E3 ligase WW domain-containing protein 2 (WWP2) and degradation, while phosphorylation at T118 (by AKT1) antagonises methylation and stabilises SOX2[118]. Likewise in Lymphoid-specific Helicase (LSH), asymmetrical di-methylation at R309 (by PRMT5) is downregulated by phosphorylation at S503 (by MAPK1), which in turn promotes stemness properties[119]. A reciprocal interaction between methylation and phosphorylation was reported in the cardiac voltage-gated sodium channel NaV1.5, where asymmetrical di-methylation at R513 (by PRMT3), which has been associated with cardiac conduction disease significantly decreases phosphorylation at S516 (by PKA), while the latter reciprocally blocks R513 methylation[120]. Even more extreme is the interaction in DNA methyltransferase-1 (DNMT1), where mono-methylation at K142 (by SET7) promotes degradation, while phosphorylation at S143 (by AKT1) stabilises the enzyme in such a way as to have a mutually exclusive PTM pair regulating DNMT1 activity through protein stability[121]. Conversely, methylation and phosphorylation work together in Apoptosis signal–regulating kinase 1 (ASK1), where asymmetrical di-methylation at R89 (by PRMT5) promotes phosphorylation of ASK1 at S83 (by

AKT1), negatively regulating ASK1 activity[122].

5.2. Application of proteogenomics

There is also a genetic component to be considered in the analysis of protein methylation data, since mutations the add or remove arginine and lysine from proteins affect the possibility of methylation. In the development of certain diseases, particularly cancers, one or more somatic mutations may be observed, which have an effect on both the characteristics and the protein function within that disease. In cancer, genomic instability and mutations together with promoter methylation, copy number aberrations and chromosomal translocations drive disease progression[123].

However, since in proteomics in order to identify peptides and proteins in a sample, the spectra generated are matched against a reference protein database which typically contains a generic set of protein sequences, meaning that besides possibly being incomplete or inaccurate, any sequence variants unique to a given sample, including disease-relevant mutated sequences would be absent in such a reference database, impeding their detection[124].

This issue can be addressed by integrating information of genomic and trancriptomic origin with the proteomic data generated. This research field is called proteogenomics and in essence uses genomic and transcriptomic data to build a protein database against which the MS/MS spectra generated from a proteomics experiment can be compared to. This improves the identification accuracy of peptides[124]. The major drawback of such an approach is that it generates a large protein search space, increasing the possibility of false positive identifications, although the use of transcriptomic data significantly reduces such search spaces[124].

Such disease-specific peptide sequences, which may originate from novel or unannotated proteins, as well as which may contain mutations or be fusions resulting from the disease, would have high biological relevance and contribute to the understanding of human biology in general[124]. Pertaining to the study of cancer, this involves the detection of tumour-specific changes throughout the proteome and thus is commonly called onco-proteogenomics[125].

Onco-proteogenomics performed using tumour-specific databases allows the characterisation of cancer phenotypes at a molecular level and improves the overall understanding of the role played by specific mutations on protein abundances, subcellular localisation, protein interactions and the effects of PTMs[125]. This data is then used to construct protein interaction networks and signalling pathways, which would be help in the understanding of interaction alterations and signalling cascades, including through aberrant PTMs, including methylations[126].

Protein variants identified through this strategy may be useful as biomarkers for a particular cancer type or assist in identifying a therapeutic target, thanks to its functional role in that specific cancer. Moreover, numerous cancers are known to dysregulate the function of their protein methyltransferases, through the accumulation of somatic mutations or else by altering their level of activity[2,127].

The workflow of onco-proteogenomics has been applied in a number of studies. In one study, 95 colon and rectal tumours were analysed at the proteomic level and 796 amino acid variants were identified across 86 tumours for which matched RNA-seq data was available, resulting in the categorisation of five major proteomic colorectal cancer subtypes[128]. In another study using 110 lung adenocarcinoma tumours and 101 matched normal adjacent tissues, genomics, epigenomics, deep-scale proteomics, phosphoproteomics and acetylproteomics were combined in order to obtain a better understanding of therapeutic options[129].

5.3. Bioinformatics Tools

As with all large-scale proteomic analyses, bioinformatics plays a critical role in methylation studies, in that specific algorithms are applied in order to identify and validate methylated peptides, in order to generate a biologically relevant output. The biggest computational issue is that sequence database searches for comprehensive LC-MS/MS methylation experiments are known to produce higher FDRs for methyl-PSMs than predicted using the target-decoy approach[101], with high specificity and sensitivity being difficult to attain. This is because the prediction of methylated peptides is dependent on the quality of existing methylation data within databases.

Until now, most researchers in the field have manually sifted through their methylation data in order to confirm the validity of each peptide, by looking for characteristic ions and other properties of methylated peptides (such as missed cleavages). There have been however a few groups that have designed in-house prediction tools using experimentally verified methylation sites as positive datasets. The major limitation of such an approach is that it treats all other datasets as negative datasets, which reduces the accuracy of prediction, since not all sites from such datasets are truly unmodified[61].

Prediction tools for methylated peptides have come a long way, and have evolved from very simplistic sequence-based tools to ones which take into account multiple peptide properties. Early prediction tools for protein methylation (such as AutoMotif), relied upon a positive dataset of methylated sites (flanked by multiple unmethylated residues) and a negative dataset containing unmethylated peptides[130]. The prediction accuracy of such an algorithm was later improved by taking into account disordered structure sequence features preferred by lysine and arginine methylations[131]. A web tool for predicting methylations was later developed (called MeMo) based on experimentally verified methylated sites in the Swiss-Prot database[132]. Further improvements were made by taking into account structural features (such as solvent-accessible surface area and secondary structure), sequence features (such as hydrophobicity and charge number), normalised van der Waals volume and position weight amino acid composition[133–135]. More recently, the development of a tool (called GPS-MSP) which uses the classification of methylation sites based on the degree of methylation, which was found to be linked to preferential involvement in different biological processes and pathways and unique sequence preference associated with each degree of methylation, considerably improved the prediction accuracy[136].

Despite the improvement in the identification of methylation sites, their validity is still questionable. In order to improve the identification of true methylated peptides hmSILAC has been applied. However, when it comes to the computational aspect of detecting methyl-SILAC pairs, software packages used for traditional SILAC or chemical labeling are not particularly useful in this case because the methylated peptides under investiga-

tion cannot be assumed to exist as both labeled and unlabeled pairs, and such softwares do not allow the identification of the particular mass shifts associated with methyl-SILAC pairs[102].

In order to validate the true or false positive status and quantify hmSILAC methylated peptides, a few groups have developed specific tools for matching hmSILAC pairs[31,102]. For example, the software MethylQuant detects putative methyl-SILAC peptide pairs within a list of methylated peptides identified from a sequence database search of a large-scale LC-MS/MS experiment. It then validates such methylated peptide pairs by generating various scores and provides relative quantification for each[102].

The ultimate computational tool for methylation analysis would incorporate machine learning, which could improve its prediction accuracy gradually, as more real methylations are identified and validated, and be able to inspect thousands of MS/MS spectra using validation parameters and site localisation algorithms.

6. Conclusion

Thanks to all these techniques and innovations, the isolation and identification of methylations at proteome level has improved by a great margin but a lot still remains to be done in order to bring methylation analyses up to the level of other PTMs, with the standard being set by phosphoproteomics. While large-scale methylation analyses have focused on arginine and lysine, the methylation data available for other amino acids is very scarce and still requires much investigation. Also, despite the availability of arginine and lysine methylation data, most of it is not properly validated with an adequate orthogonal method to discern real from artifactual identifications. This means that unvalidated methylation datasets would need to be reanalysed. Finally, once the target protein, residue position and degree of methylation are determined, development by wet-lab assays are required to confirm that these are truly functional methylations, determine the methyltransferase responsible for the addition and understand the role in both healthy cellular biology and disease.

As more importance is placed on methylation function, there will be a greater tendency for its inclusion in biomarker assays or diagnostic kits similar to the application of phosphorylations. Therefore, the further development of large-scale methylation enrichment, detection, identification and orthogonal validation strategies, together with ancillary bioinformatics tools are required in order to advance this research field. This being said, the current limitations are only expected to be short-term, when considering the fast pace of technological advances in mass spectrometry bolstering resolution, and the development of innovative biological techniques to enrich for methylation peptides.

Acknowledgements

This research as part of project Kme-CRC was financed by the Malta Council for Science & Technology, for and on behalf of the Foundation for Science and Technology, through the FUSION: R&I Technology Development Programme.

References

[1] Baron, B. Lysine Methylation of Non-Histone Proteins. *Biochemistry and Modern Applications* **2015** *1*, 1–2.
[2] Baron, B. The Lysine Multi-Switch: The Impact of Lysine Methylation on Transcription Factor Properties. *Biohelikon Cell Bio.* **2014**, *1*, 1–8.
[3] Schubert, H. L.; Blumenthal, R. M.; Cheng, X. 1 Protein Methyltransferases: Their Distribution Among the Five Structural Classes of AdoMet-Dependent Methyltransferases. In *The Enzymes*; Clarke, S. G., Tamanoi, F., Eds.; Protein Methyltransferases; Academic Press, **2006**, *(24)*, 3–28.
[4] Fontecave, M.; Atta, M.; Mulliez, E. S-Adenosylmethionine: Nothing Goes to Waste. *Trends Biochem. Sci.* **2004**, *29*, 243–249.
[5] Murn, J.; Shi, Y. The Winding Path of Protein Methylation Research: Milestones and New Frontiers. *Nat. Rev. Mol Cell Biol.* **2017**, *18*, 517–527.
[6] Ng, S. S.; Yue, W. W.; Oppermann, U.; Klose, R. J. Dynamic Protein Methylation in Chromatin Biology. *Cell Mol Life Sci.* **2009**, *66*, 407–422.
[7] Grillo, M. A.; Colombatto, S. S-Adenosylmethionine and Protein Methylation. *Amino Acids* **2005**, *28*, 357–362.
[8] Lake, A. N.; Bedford, M. T. Protein Methylation and DNA Repair. *Mutat. Res.* **2007**, *618*, 91–101.
[9] Smith, B. C.; Denu, J. M. Chemical Mechanisms of Histone Lysine and Arginine Modifications. *Biochim. Biophys. Acta* **2009**, *1789*, 45–57.
[10] Kim, J. H.; Yoo, B. C.; Yang, W. S.; Kim, E.; Hong, S.; Cho, J. Y. The Role of Protein Arginine Methyltransferases in Inflammatory Responses. *Mediators Inflamm.* **2016**, 4028353.
[11] Bedford, M. T.; Richard, S. Arginine Methylation an Emerging Regulator of Protein Function. *Mol. Cell* **2005**, *18*, 263–272.
[12] Kaniskan, H. Ü.; Jin, J. Chemical Probes of Histone Lysine Methyltransferases. *ACS Chem. Biol.* **2015**, *10*, 40–50.
[13] Schubert, H. L.; Blumenthal, R. M.; Cheng, X. Many Paths to Methyltransfer: A Chronicle of Convergence. *Trends Biochem. Sci.* **2003**, *28*, 329–335.
[14] Petrossian, T. C.; Clarke, S. G. Uncovering the Human Methyltransferasome. *Mol. Cell. Proteomics* **2011**, *10*, M110.000976.
[15] Falnes, P. Ø.; Jakobsson, M. E.; Davydova, E.; Ho, A.; Małecki, J. Protein Lysine Methylation by Seven-β-Strand Methyltransferases. *Biochem. J.* **2016**, *473*, 1995–2009.
[16] Polevoda, B.; Sherman, F. Methylation of Proteins Involved in Translation. *Mol. Microbiol.* **2007**, *65*, 590–606.
[17] Xin, F.; Radivojac, P. Post-Translational Modifications Induce Significant yet Not Extreme Changes to Protein Structure. *Bioinformatics* **2012**, *28*, 2905–2913.
[18] Hamamoto, R.; Saloura, V.; Nakamura, Y. Critical Roles of Non-Histone Protein Lysine Methylation in Human Tumorigenesis. *Nat. Rev. Cancer* **2015**, *15*, 110–124.
[19] Biggar, K. K.; Li, S. S.-C. Non-Histone Protein Methylation as a Regulator of Cellular Signalling and Function. *Nat. Rev. Mol. Cell Biol.* **2015**, *16*, 5–17.

[20] Huang, J.; Berger, S. L. The Emerging Field of Dynamic Lysine Methylation of Non-Histone Proteins. *Curr. Opin. Genet. Dev.* **2008**, *18*, 152–158.
[21] Egorova, K. S.; Olenkina, O. M.; Olenina, L. V. Lysine Methylation of Nonhistone Proteins Is a Way to Regulate Their Stability and Function. *Biochemistry (Mosc.)* **2010**, *75*, 535–548.
[22] Baron, B. Application of Proteomics to Cancer Therapy. *SMGEBooks.* **2016**.
[23] Pandeswari, P. B.; Sabareesh, V. Middle-down Approach: A Choice to Sequence and Characterize Proteins/Proteomes by Mass Spectrometry. *RSC Advances* **2019**, *9(1)*, 313–344.
[24] Rappsilber, J.; Friesen, W. J.; Paushkin, S.; Dreyfuss, G.; Mann, M. Detection of Arginine Dimethylated Peptides by Parallel Precursor Ion Scanning Mass Spectrometry in Positive Ion Mode. *Anal. Chem.* **2003**, *75*, 3107–3114.
[25] Uhlmann, T.; Geoghegan, V. L.; Thomas, B.; Ridlova, G.; Trudgian, D. C.; Acuto, O. A Method for Large-Scale Identification of Protein Arginine Methylation. *Mol. Cell. Proteomics* **2012**, *11*, 1489–1499.
[26] Hecht, E. S.; Scigelova, M.; Eliuk, S.; Makarov, A. Fundamentals and Advances of Orbitrap Mass Spectrometry. In *Encyclopedia of Analytical Chemistry*; American Cancer Society, **2019**, 1–40.
[27] Meng, F.; Forbes, A. J.; Miller, L. M.; Kelleher, N. L. Detection and Localization of Protein Modifications by High Resolution Tandem Mass Spectrometry. *Mass Spectrom. Rev.* **2005**, *24*, 126–134.
[28] Erce, M. A.; Pang, C. N. I.; Hart-Smith, G.; Wilkins, M. R. The Methylproteome and the Intracellular Methylation Network. *Proteomics* **2012**, *12*, 564–586.
[29] Bremang, M.; Cuomo, A.; Agresta, A. M.; Stugiewicz, M.; Spadotto, V.; Bonaldi, T. Mass Spectrometry-Based Identification and Characterisation of Lysine and Arginine Methylation in the Human Proteome. *Mol. Biosyst.* **2013**, *9*, 2231–2247.
[30] Cao, X.-J.; Arnaudo, A. M.; Garcia, B. A. Large-Scale Global Identification of Protein Lysine Methylation in Vivo. *Epigenetics* **2013**, *8*, 477–485.
[31] Geoghegan, V.; Guo, A.; Trudgian, D.; Thomas, B.; Acuto, O. Comprehensive Identification of Arginine Methylation in Primary T Cells Reveals Regulatory Roles in Cell Signalling. *Nat. Commun.* **2015**, *6*, 6758.
[32] Guo, A.; Gu, H.; Zhou, J.; Mulhern, D.; Wang, Y.; Lee, K. A.; Yang, V.; Aguiar, M.; Kornhauser, J.; Jia, X.; Ren, J.; Beausoleil, S. A.; Silva, J. C.; Vemulapalli, V.; Bedford, M. T.; Comb, M. J. Immunoaffinity Enrichment and Mass Spectrometry Analysis of Protein Methylation. *Mol. Cell. Proteomics* **2014**, *13*, 372–387.
[33] Larsen, S. C.; Sylvestersen, K. B.; Mund, A.; Lyon, D.; Mullari, M.; Madsen, M. V.; Daniel, J. A.; Jensen, L. J.; Nielsen, M. L. Proteome-Wide Analysis of Arginine Monomethylation Reveals Widespread Occurrence in Human Cells. *Sci. Signal.* **2016**, *9*, rs9.
[34] Moore, K. E.; Carlson, S. M.; Camp, N. D.; Cheung, P.; James, R. G.; Chua, K. F.; Wolf-Yadlin, A.; Gozani, O. A General Molecular Affinity Strategy for Global Detection and Proteomic Analysis of Lysine Methylation. *Mol. Cell* **2013**, *50*, 444–456.
[35] Shishkova, E.; Zeng, H.; Liu, F.; Kwiecien, N. W.; Hebert, A. S.; Coon, J. J.; Xu, W. Global Mapping of CARM1 Substrates Defines Enzyme Specificity and Substrate Recognition. *Nat. Commun.* **2017**, *8*, 15571.
[36] Sylvestersen, K. B.; Horn, H.; Jungmichel, S.; Jensen, L. J.; Nielsen, M. L. Proteomic Analysis of Arginine Methylation Sites in Human Cells Reveals Dynamic Regulation during Transcriptional Arrest. *Mol. Cell. Proteomics* **2014**, *13*, 2072–2088.
[37] Wu, Z.; Cheng, Z.; Sun, M.; Wan, X.; Liu, P.; He, T.; Tan, M.; Zhao, Y. A Chemical Proteomics Approach for Global Analysis of Lysine Monomethylome Profiling. *Mol. Cell. Proteomics* **2015**, *14*, 329–339.
[38] Liu, H.; Galka, M.; Liu, X.; Lin, Y.; Pittock, P.; Voss, C.; Dhami, G.; Li, X.; Miyaji, M.; Lajoie, G.; Chen, B.; Li, S. S.-C. A Method for Systematic Mapping of Protein Lysine Methylation Identifies New Functions for HP1β in DNA Damage Repair. *Mol. Cell* **2013**, *50*, 723–735.
[39] Bulau, P.; Zakrzewicz, D.; Kitowska, K.; Wardega, B.; Kreuder, J.; Eickelberg, O. Quantitative Assessment of Arginine Methylation in Free versus Protein-Incorporated Amino Acids in Vitro and in Vivo Using Protein Hydrolysis and High-Performance Liquid Chromatography. *Biotechniques* **2006**, *40*, 305–310.
[40] Wang, Q.; Liu, Z.; Wang, K.; Wang, Y.; Ye, M. A New Chromatographic Approach to Analyze Methylproteome with Enhanced Lysine Methylation Identification Performance. *Anal. Chim. Acta* **2019**, **1068**, 111–119.
[41] Wang, K.; Dong, M.; Mao, J.; Wang, Y.; Jin, Y.; Ye, M.; Zou, H. Antibody-Free Approach for the Global Analysis of Protein Methylation. *Anal. Chem.* **2016**, *88*, 11319–11327.
[42] Wang, W.; Zhang, H.; Wang, L.; Zhang, S.; Tang, M. MiR-613 Inhibits the Growth and Invasiveness of Human Hepatocellular Carcinoma via Targeting DCLK1. *Biochem. Biophys. Res. Commun.* **2016**, *473*, 987–992.
[43] Sobott, F.; Watt, S. J.; Smith, J.; Edelmann, M. J.; Kramer, H. B.; Kessler, B. M. Comparison of CID versus ETD Based MS/MS Fragmentation for the Analysis of Protein Ubiquitination. *J. Am. Soc. Mass Spectrom.* **2009**, *20*, 1652–1659.
[44] Steen, H.; Mann, M. The ABC's (and XYZ's) of Peptide Sequencing. *Nat. Rev. Mol. Cell Biol.* **2004**, *5*, 699–711.
[45] Brame, C. J.; Moran, M. F.; McBroom-Cerajewski, L. D. B. A Mass Spectrometry Based Method for Distinguishing between Symmetrically and Asymmetrically Dimethylated Arginine Residues. *Rapid Commun. Mass Spectrom.* **2004**, *18*, 877–881.
[46] Snijders, A. P. L.; Hung, M.-L.; Wilson, S. A.; Dickman, M. J. Analysis of Arginine and Lysine Methylation Utilizing Peptide Separations at Neutral PH and Electron Transfer Dissociation Mass Spectrometry. *J. Am. Soc. Mass Spectrom.* **2010**, *21*, 88–96.
[47] Wang, H.; Straubinger, R. M.; Aletta, J. M.; Cao, J.; Duan, X.; Yu, H.; Qu, J. Accurate Localization and Relative Quantification of Arginine Methylation Using Nanoflow Liquid Chromatography Coupled to Electron Transfer Dissociation and Orbitrap Mass Spectrometry. *J. Am. Soc. Mass Spectrom.* **2009**, *20*, 507–519.
[48] Ong, S.-E.; Mittler, G.; Mann, M. Identifying and Quantifying in Vivo Methylation Sites by Heavy Methyl SILAC. *Nat. Methods* **2004**, *1*, 119–126.
[49] Sprung, R.; Chen, Y.; Zhang, K.; Cheng, D.; Zhang, T.; Peng, J.; Zhao, Y. Identification and Validation of Eukaryotic Aspartate and Glutamate Methylation in Proteins. *J. Proteome Res.* **2008**, *7*, 1001–1006.
[50] Jung, S. Y.; Li, Y.; Wang, Y.; Chen, Y.; Zhao, Y.; Qin, J. Complications in the Assignment of 14 and 28 Da Mass Shift Detected by Mass Spectrometry as in Vivo Methylation from Endogenous Proteins. *Anal. Chem.* **2008**, *80*, 1721–1729.
[51] Zou, Y.; Wang, Y. Mass Spectrometric Analysis of High-Mobility Group Proteins and Their Post-Translational Modifications in Normal and Cancerous Human Breast Tissues. *J. Proteome Res.* **2007**, *6*, 2304–2314.
[52] Hart-Smith, G.; Low, J. K. K.; Erce, M. A.; Wilkins, M. R. Enhanced Methylarginine Characterization by Post-Translational Modification-Specific Targeted Data Acquisition and Electron-Transfer Dissociation Mass Spectrometry. *J. Am. Soc. Mass Spectrom.* **2012**, *23*, 1376–1389.
[53] Gehrig, P. M.; Hunziker, P. E.; Zahariev, S.; Pongor, S. Fragmentation Pathways of N(G)-Methylated and Unmodified Arginine Residues in Peptides Studied by ESI-MS/MS and MALDI-MS. *J. Am. Soc. Mass Spectrom.* **2004**, *15*, 142–149.

[54] Syka, J. E. P.; Coon, J. J.; Schroeder, M. J.; Shabanowitz, J.; Hunt, D. F. Peptide and Protein Sequence Analysis by Electron Transfer Dissociation Mass Spectrometry. *Proc. Natl. Acad. Sci. U. S. A.* **2004**, *101*, 9528–9533.

[55] Li, X.; Cournoyer, J. J.; Lin, C.; O'Connor, P. B. The Effect of Fixed Charge Modifications on Electron Capture Dissociation. *J. Am. Soc. Mass Spectrom.* **2008**, *19*, 1514–1526.

[56] Coon, J. J.; Shabanowitz, J.; Hunt, D. F.; Syka, J. E. P. Electron Transfer Dissociation of Peptide Anions. *J. Am. Soc. Mass Spectrom.* **2005**, *16*, 880–882.

[57] Mikesh, L. M.; Ueberheide, B.; Chi, A.; Coon, J. J.; Syka, J. E. P.; Shabanowitz, J.; Hunt, D. F. The Utility of ETD Mass Spectrometry in Proteomic Analysis. *Biochim. Biophys. Acta* **2006**, **1764**, 1811–1822.

[58] Swaney, D. L.; McAlister, G. C.; Coon, J. J. Decision Tree-Driven Tandem Mass Spectrometry for Shotgun Proteomics. *Nat. Methods* **2008**, *5*, 959–964.

[59] Riley, N. M.; Coon, J. J. The Role of Electron Transfer Dissociation in Modern Proteomics. *Anal. Chem.* **2018**, *90*, 40–64.

[60] Good, D. M.; Wirtala, M.; McAlister, G. C.; Coon, J. J. Performance Characteristics of Electron Transfer Dissociation Mass Spectrometry. *Mol. Cell. Proteomics* **2007**, *6*, 1942–1951.

[61] Wang, Q.; Wang, K.; Ye, M. Strategies for Large-Scale Analysis of Non-Histone Protein Methylation by LC-MS/MS. *Analyst.* **2017**, *142*, 3536–3548.

[62] Katsanovskaja, K.; Driver, T.; Pipkorn, R.; Edelson-Averbukh, M. Negative Ion Mode Collision-Induced Dissociation for Analysis of Protein Arginine Methylation. *J. Am. Soc. Mass Spectrom.* **2019**, *30*, 1229–1241.

[63] Zhang, K.; Yau, P. M.; Chandrasekhar, B.; New, R.; Kondrat, R.; Imai, B. S.; Bradbury, M. E. Differentiation between Peptides Containing Acetylated or Tri-Methylated Lysines by Mass Spectrometry: An Application for Determining Lysine 9 Acetylation and Methylation of Histone H3. *Proteomics* **2004**, *4*, 1–10.

[64] Couttas, T. A.; Raftery, M. J.; Bernardini, G.; Wilkins, M. R. Immonium Ion Scanning for the Discovery of Post-Translational Modifications and Its Application to Histones. *J. Proteome. Res.* **2008**, *7*, 2632–2641.

[65] Afjehi-Sadat, L.; Garcia, B. A. Comprehending Dynamic Protein Methylation with Mass Spectrometry. *Curr. Opin. Chem. Biol.* **2013**, *17*, 12–19.

[66] Zheng, Y.; Huang, X.; Kelleher, N. L. Epiproteomics: Quantitative Analysis of Histone Marks and Codes by Mass Spectrometry. *Curr. Opin. Chem. Biol.* **2016**, *33*, 142–150.

[67] Hebert, A. S.; Prasad, S.; Belford, M. W.; Bailey, D. J.; McAlister, G. C.; Abbatiello, S. E.; Huguet, R.; Wouters, E. R.; Dunyach, J.-J.; Brademan, D. R.; Westphall, M. S.; Coon, J. J. Comprehensive Single-Shot Proteomics with FAIMS on a Hybrid Orbitrap Mass Spectrometer. *Anal. Chem.* **2018**, *90*, 9529–9537.

[68] Xuan, Y.; Creese, A. J.; Horner, J. A.; Cooper, H. J. High-Field Asymmetric Waveform Ion Mobility Spectrometry (FAIMS) Coupled with High-Resolution Electron Transfer Dissociation Mass Spectrometry for the Analysis of Isobaric Phosphopeptides. *Rapid Commun. Mass Spectrom.* **2009**, *23*, 1963–1969.

[69] Shvartsburg, A. A.; Tang, K.; Smith, R. D. Understanding and Designing Field Asymmetric Waveform Ion Mobility Spectrometry Separations in Gas Mixtures. *Anal. Chem* **2004**, *76*, 7366–7374.

[70] Guevremont, R. High-Field Asymmetric Waveform Ion Mobility Spectrometry: A New Tool for Mass Spectrometry. *J. Chromatogr.* **2004**, **1058**, 3–19.

[71] Bonneil, E.; Pfammatter, S.; Thibault, P. Enhancement of Mass Spectrometry Performance for Proteomic Analyses Using High-Field Asymmetric Waveform Ion Mobility Spectrometry (FAIMS). *J. Mass Spectrom.* **2015**, *50*, 1181–1195.

[72] Shvartsburg, A. A.; Creese, A. J.; Smith, R. D.; Cooper, H. J. Separation of Peptide Isomers with Variant Modified Sites by High-Resolution Differential Ion Mobility Spectrometry. *Anal. Chem.* **2010**, *82*, 8327–8334.

[73] Shvartsburg, A. A.; Creese, A. J.; Smith, R. D.; Cooper, H. J. Separation of a Set of Peptide Sequence Isomers Using Differential Ion Mobility Spectrometry. *Anal. Chem.* **2011**, *83*, 6918–6923.

[74] Kanshin, E.; Michnick, S.; Thibault, P. Sample Preparation and Analytical Strategies for Large-Scale Phosphoproteomics Experiments. *Semin. Cell Dev. Biol.* **2012**, *23*, 843–853.

[75] Saba, J.; Bonneil, E.; Pomiès, C.; Eng, K.; Thibault, P. Enhanced Sensitivity in Proteomics Experiments Using FAIMS Coupled with a Hybrid Linear Ion Trap/Orbitrap Mass Spectrometer. *J. Proteome Res.* **2009**, *8*, 3355–3366.

[76] Xia, Y.-Q.; Wu, S. T.; Jemal, M. LC-FAIMS-MS/MS for Quantification of a Peptide in Plasma and Evaluation of FAIMS Global Selectivity from Plasma Components. *Anal. Chem.* **2008**, *80*, 7137–7143.

[77] Shvartsburg, A. A.; Seim, T. A.; Danielson, W. F.; Norheim, R.; Moore, R. J.; Anderson, G. A.; Smith, R. D. High-Definition Differential Ion Mobility Spectrometry with Resolving Power up to 500. *J. Am. Soc. Mass Spectrom.* **2013**, *24*, 109–114.

[78] Shvartsburg, A. A.; Zheng, Y.; Smith, R. D.; Kelleher, N. L. Separation of Variant Methylated Histone Tails by Differential Ion Mobility. *Anal. Chem.* **2012**, *84*, 6317–6320.

[79] Winter, D. L.; Mastellone, J.; Kabir, K. M. M.; Wilkins, M. R.; Donald, W. A. Separation of Isobaric Mono- and Dimethylated RGG-Repeat Peptides by Differential Ion Mobility-Mass Spectrometry. *Anal. Chem.* **2019**, *91*, 11827–11833.

[80] Michalski, A.; Cox, J.; Mann, M. More than 100,000 Detectable Peptide Species Elute in Single Shotgun Proteomics Runs but the Majority Is Inaccessible to Data-Dependent LC-MS/MS. *J. Proteome Res.* **2011**, *10*, 1785–1793.

[81] Boisvert, F.-M.; Côté, J.; Boulanger, M.-C.; Richard, S. A Proteomic Analysis of Arginine-Methylated Protein Complexes. *Mol. Cell. Proteomics* **2003**, *2*, 1319–1330.

[82] Wesche, J.; Kühn, S.; Kessler, B. M.; Salton, M.; Wolf, A. Protein Arginine Methylation: A Prominent Modification and Its Demethylation. *Cell. Mol. Life Sci.* **2017**, *74*, 3305–3315.

[83] Lehman, S. M.; Chen, H.; Burgos, E. S.; Maron, M.; Gayatri, S.; Nieves, E.; Bai, D. L.; Sidoli, S.; Gupta, V.; Marunde, M. R.; Bone, J. R.; Sun, Z.-W.; Bedford, M. T.; Shabanowitz, J.; Hunt, D. F.; Shechter, D. Transcriptomic and Proteomic Regulation through Abundant, Dynamic, and Independent Arginine Methylation by Type I and Type II PRMTs. *bioRxiv.* **2020**, 2020.06.23.167601.

[84] Guo, H.; Wang, R.; Zheng, W.; Chen, Y.; Blum, G.; Deng, H.; Luo, M. Profiling Substrates of Protein Arginine N-Methyltransferase 3 with S-Adenosyl-L-Methionine Analogues. *ACS Chem. Biol.* **2014**, *9*, 476–484.

[85] Levy, D.; Liu, C. L.; Yang, Z.; Newman, A. M.; Alizadeh, A. A.; Utz, P. J.; Gozani, O. A Proteomic Approach for the Identification of Novel Lysine Methyltransferase Substrates. *Epigenetics Chromatin.* **2011**, *4*, 19.

[86] Najbauer, J.; Aswad, D. W. Diversity of Methyl Acceptor Proteins in Rat Pheochromocytoma (PC12) Cells Revealed after Treatment with

Adenosine Dialdehyde. *J. Biol. Chem.* **1990**, *265*, 12717–12721.
[87] Carlson, S. M.; Moore, K. E.; Green, E. M.; Martín, G. M.; Gozani, O. Proteome-Wide Enrichment of Proteins Modified by Lysine Methylation. *Nat. Protoc.* **2014**, *9*, 37–50.
[88] Nady, N.; Krichevsky, L.; Zhong, N.; Duan, S.; Tempel, W.; Amaya, M. F.; Ravichandran, M.; Arrowsmith, C. H. Histone Recognition by Human Malignant Brain Tumor Domains. *J. Mol. Biol.* **2012**, *423*, 702–718.
[89] Liu, H.; Galka, M.; Iberg, A.; Wang, Z.; Li, L.; Voss, C.; Jiang, X.; Lajoie, G.; Huang, Z.; Bedford, M. T.; Li, S. S. C. Systematic Identification of Methyllysine-Driven Interactions for Histone and Nonhistone Targets. *J. Proteome Res.* **2010**, *9*, 5827–5836.
[90] Ning, Z.; Star, A. T.; Mierzwa, A.; Lanouette, S.; Mayne, J.; Couture, J.-F.; Figeys, D. A Charge-Suppressing Strategy for Probing Protein Methylation. *Chem. Commun. (Camb.)* **2016**, *52*, 5474–5477.
[91] Choudhary, C.; Mann, M. Decoding Signalling Networks by Mass Spectrometry-Based Proteomics. *Nat. Rev. Mol. Cell Biol.* **2010**, *11*, 427–439.
[92] Beausoleil, S. A.; Jedrychowski, M.; Schwartz, D.; Elias, J. E.; Villén, J.; Li, J.; Cohn, M. A.; Cantley, L. C.; Gygi, S. P. Large-Scale Characterization of HeLa Cell Nuclear Phosphoproteins. *Proc. Natl. Acad. Sci. U. S. A.* **2004**, *101*, 12130–12135.
[93] Cao, X.-J.; Garcia, B. A. Global Proteomics Analysis of Protein Lysine Methylation. *Curr. Protoc. Protein Sci.* **2016**, *86*, 24.8.1-24.8.19.
[94] Olsen, J. B.; Cao, X.-J.; Han, B.; Chen, L. H.; Horvath, A.; Richardson, T. I.; Campbell, R. M.; Garcia, B. A.; Nguyen, H. Quantitative Profiling of the Activity of Protein Lysine Methyltransferase SMYD2 Using SILAC-Based Proteomics. *Mol. Cell. Proteomics* **2016**, *15*, 892–905.
[95] Alpert, A. J. Hydrophilic-Interaction Chromatography for the Separation of Peptides, Nucleic Acids and Other Polar Compounds. *J. Chromatogr.* **1990**, *499*, 177–196.
[96] McNulty, D. E.; Annan, R. S. Hydrophilic Interaction Chromatography Reduces the Complexity of the Phosphoproteome and Improves Global Phosphopeptide Isolation and Detection. *Mol. Cell. Proteomics* **2008**, *7*, 971–980.
[97] Jian, W.; Edom, R. W.; Xu, Y.; Weng, N. Recent Advances in Application of Hydrophilic Interaction Chromatography for Quantitative Bioanalysis. *J. Sep. Sci.* **2010**, *33*, 681–697.
[98] Huang, J.; Wang, F.; Ye, M.; Zou, H. Enrichment and Separation Techniques for Large-Scale Proteomics Analysis of the Protein Post-Translational Modifications. *J. Chromatogr.* **2014**, **1372C**, 1–17.
[99] Wang, R.; Zheng, W.; Yu, H.; Deng, H.; Luo, M. Labeling Substrates of Protein Arginine Methyltransferase with Engineered Enzymes and Matched S-Adenosyl-L-Methionine Analogues. *J. Am. Chem. Soc.* **2011**, *133*, 7648–7651.
[100] Elias, J. E.; Gygi, S. P. Target-Decoy Search Strategy for Increased Confidence in Large-Scale Protein Identifications by Mass Spectrometry. *Nat. Methods* **2007**, *4*, 207–214.
[101] Hart-Smith, G.; Yagoub, D.; Tay, A. P.; Pickford, R.; Wilkins, M. R. Large Scale Mass Spectrometry-Based Identifications of Enzyme-Mediated Protein Methylation Are Subject to High False Discovery Rates. *Mol. Cell. Proteomics* **2016**, *15*, 989–1006.
[102] Tay, A. P.; Geoghegan, V.; Yagoub, D.; Wilkins, M. R.; Hart-Smith, G. MethylQuant: A Tool for Sensitive Validation of Enzyme-Mediated Protein Methylation Sites from Heavy-Methyl SILAC Data. *J. Proteome Res.* **2018**, *17*, 359–373.
[103] Mann, M. Functional and Quantitative Proteomics Using SILAC. *Nat. Rev. Mol. Cell Biol.* **2006**, *7*, 952–958.
[104] Baron, B. Biochemical Formulation of Chemically-Defined and Non-Xenogeneic Culture Media for Research and Clinical Applications of Human Stem Cells. *Basic Biochemistry* **2016**, 1–16.
[105] Wang, K.; Zhou, Y. J.; Liu, H.; Cheng, K.; Mao, J.; Wang, F.; Liu, W.; Ye, M.; Zhao, Z. K.; Zou, H. Proteomic Analysis of Protein Methylation in the Yeast Saccharomyces Cerevisiae. *J. Proteomics* **2015**, *114*, 226–233.
[106] Islam, K.; Bothwell, I.; Chen, Y.; Sengelaub, C.; Wang, R.; Deng, H.; Luo, M. Bioorthogonal Profiling of Protein Methylation Using Azido Derivative of S-Adenosyl-L-Methionine. *J. Am. Chem. Soc.* **2012**, *134*, 5909–5915.
[107] Lee, B. W. K.; Sun, H. G.; Zang, T.; Kim, B. J.; Alfaro, J. F.; Zhou, Z. S. Enzyme-Catalyzed Transfer of a Ketone Group from an S-Adenosylmethionine Analogue: A Tool for the Functional Analysis of Methyltransferases. *J. Am. Chem. Soc.* **2010**, *132*, 3642–3643.
[108] Wang, Y.; Yang, F.; Gritsenko, M. A.; Wang, Y.; Clauss, T.; Liu, T.; Shen, Y.; Monroe, M. E.; Lopez-Ferrer, D.; Reno, T.; Moore, R. J.; Klemke, R. L.; Camp, D. G.; Smith, R. D. Reversed-Phase Chromatography with Multiple Fraction Concatenation Strategy for Proteome Profiling of Human MCF10A Cells. *Proteomics* **2011**, *11*, 2019–2026.
[109] Peters, W.; Willnow, S.; Duisken, M.; Kleine, H.; Macherey, T.; Duncan, K. E.; Litchfield, D. W.; Lüscher, B.; Weinhold, E. Enzymatic Site-Specific Functionalization of Protein Methyltransferase Substrates with Alkynes for Click Labeling. *Angew. Chem. Int. Ed. Engl.* **2010**, *49*, 5170–5173.
[110] Binda, O.; Boyce, M.; Rush, J. S.; Palaniappan, K. K.; Bertozzi, C. R.; Gozani, O. A Chemical Method for Labeling Lysine Methyltransferase Substrates. *Chembiochem* **2011**, *12*, 330–334.
[111] Singh, S.; Zhang, J.; Huber, T. D.; Sunkara, M.; Hurley, K.; Goff, R. D.; Wang, G.; Zhang, W.; Liu, C.; Rohr, J.; Van Lanen, S. G.; Morris, A. J.; Thorson, J. S. Facile Chemoenzymatic Strategies for the Synthesis and Utilization of S-Adenosyl-(L)-Methionine Analogues. *Angew. Chem. Int. Ed. Engl.* **2014**, *53*, 3965–3969.
[112] Bothwell, I. R.; Luo, M. Large-Scale, Protection-Free Synthesis of Se-Adenosyl-L-Selenomethionine Analogues and Their Application as Cofactor Surrogates of Methyltransferases. *Org. Lett.* **2014**, *16*, 3056–3059.
[113] Zhang, Y.; Pan, Y.; Liu, W.; Zhou, Y. J.; Wang, K.; Wang, L.; Sohail, M.; Ye, M.; Zou, H.; Zhao, Z. K. In Vivo Protein Allylation to Capture Protein Methylation Candidates. *Chem. Commun.* **2016**, *52*, 6689–6692.
[114] Huang, J.; Perez-Burgos, L.; Placek, B. J.; Sengupta, R.; Richter, M.; Dorsey, J. A.; Kubicek, S.; Opravil, S.; Jenuwein, T.; Berger, S. L. Repression of P53 Activity by Smyd2-Mediated Methylation. *Nature* **2006**, *444*, 629–632.
[115] Zhang, X.; Tanaka, K.; Yan, J.; Li, J.; Peng, D.; Jiang, Y.; Yang, Z.; Barton, M. C.; Wen, H.; Shi, X. Regulation of Estrogen Receptor α by Histone Methyltransferase SMYD2-Mediated Protein Methylation. *Proc. Natl. Acad. Sci. U. S. A.* **2013**, *110*, 17284–17289.
[116] Zhang, L.; Tran, N.-T.; Su, H.; Wang, R.; Lu, Y.; Tang, H.; Aoyagi, S.; Guo, A.; Khodadadi-Jamayran, A.; Zhou, D.; Qian, K.; Hricik, T.; Côté, J.; Han, X.; Zhou, W.; Laha, S.; Abdel-Wahab, O.; Levine, R. L.; Raffel, G.; Liu, Y.; Chen, D.; Li, H.; Townes, T.; Wang, H.; Deng, H.; Zheng, Y. G.; Leslie, C.; Luo, M.; Zhao, X. Cross-Talk between PRMT1-Mediated Methylation and Ubiquitylation on RBM15 Controls RNA Splicing. *Elife* **2015**, ***4*.**
[117] Guo, Z.; Zheng, L.; Xu, H.; Dai, H.; Zhou, M.; Pascua, M. R.; Chen, Q. M.; Shen, B. Methylation of FEN1 Suppresses Nearby Phosphorylation and Facilitates PCNA Binding. *Nat. Chem. Biol.* **2010**, *6*, 766–773.
[118] Fang, L.; Zhang, L.; Wei, W.; Jin, X.; Wang, P.; Tong, Y.; Li, J.; Du, J. X.; Wong, J. A Methylation-Phosphorylation Switch Determines Sox2 Stability and Function in ESC Maintenance or Differentiation. *Mol. Cell.* **2014**, *55*, 537–551.
[119] Liu, N.; Yang, R.; Shi, Y.; Chen, L.; Liu, Y.; Wang, Z.; Liu, S.; Ouyang, L.; Wang, H.; Lai, W.; Mao, C.; Wang, M.; Cheng, Y.; Liu, S.; Wang, X.;

Zhou, H.; Cao, Y.; Xiao, D.; Tao, Y. The Cross-Talk between Methylation and Phosphorylation in Lymphoid-Specific Helicase Drives Cancer Stem-like Properties. *Signal Transduct. Target. Ther.* **2020**, *5*, 197.
[120] Beltran-Alvarez, P.; Feixas, F.; Osuna, S.; Díaz-Hernández, R.; Brugada, R.; Pagans, S. Interplay between R513 Methylation and S516 Phosphorylation of the Cardiac Voltage-Gated Sodium Channel. *Amino Acids* **2015**, *47*, 429–434.
[121] Estève, P.-O.; Chang, Y.; Samaranayake, M.; Upadhyay, A. K.; Horton, J. R.; Feehery, G. R.; Cheng, X.; Pradhan, S. A Methylation and Phosphorylation Switch between an Adjacent Lysine and Serine Determines Human DNMT1 Stability. *Nat. Struct. Mol. Biol* **2011**, *18*, 42–48.
[122] Chen, M.; Qu, X.; Zhang, Z.; Wu, H.; Qin, X.; Li, F.; Liu, Z.; Tian, L.; Miao, J.; Shu, W. Cross-Talk between Arg Methylation and Ser Phosphorylation Modulates Apoptosis Signal-Regulating Kinase 1 Activation in Endothelial Cells. *Mol. Biol. Cell* **2016**, *27*, 1358–1366.
[123] Hanahan, D.; Weinberg, R. A. Hallmarks of Cancer: The next Generation. *Cell* **2011**, *144*, 646–674.
[124] Sheynkman, G. M.; Shortreed, M. R.; Cesnik, A. J.; Smith, L. M. Proteogenomics: Integrating Next-Generation Sequencing and Mass Spectrometry to Characterize Human Proteomic Variation. *Annu. Rev. Anal. Chem. (Palo Alto Calif)* **2016**, *9*, 521–545.
[125] Alfaro, J. A.; Sinha, A.; Kislinger, T.; Boutros, P. C. Onco-Proteogenomics: Cancer Proteomics Joins Forces with Genomics. *Nat. Methods* **2014**, *11*, 1107–1113.
[126] Boja, E. S.; Rodriguez, H. Proteogenomic Convergence for Understanding Cancer Pathways and Networks. *Clin. Proteomics* **2014**, *11*, 22.
[127] Colón-Bolea, P.; Crespo, P. Lysine Methylation in Cancer: SMYD3-MAP3K2 Teaches Us New Lessons in the Ras-ERK Pathway. *Bioessays* **2014**, *36*, 1162–1169.
[128] Zhang, B.; Wang, J.; Wang, X.; Zhu, J.; Liu, Q.; Shi, Z.; Chambers, M. C.; Zimmerman, L. J.; Shaddox, K. F.; Kim, S.; Davies, S. R.; Wang, S.; Wang, P.; Kinsinger, C. R.; Rivers, R. C.; Rodriguez, H.; Townsend, R. R.; Ellis, M. J. C.; Carr, S. A.; Tabb, D. L.; Coffey, R. J.; Slebos, R. J. C.; Liebler, D. C. NCI CPTAC. Proteogenomic Characterization of Human Colon and Rectal Cancer. *Nature* **2014**, *513*, 382–387.
[129] Gillette, M. A.; Satpathy, S.; Cao, S.; Dhanasekaran, S. M.; Vasaikar, S. V.; Krug, K.; Petralia, F.; Li, Y.; Liang, W.-W.; Reva, B.; Krek, A.; Ji, J.; Song, X.; Liu, W.; Hong, R.; Yao, L.; Blumenberg, L.; Savage, S. R.; Wendl, M. C.; Wen, B.; Li, K.; Tang, L. C.; MacMullan, M. A.; Avanessian, S. C.; Kane, M. H.; Newton, C. J.; Cornwell, M.; Kothadia, R. B.; Ma, W.; Yoo, S.; Mannan, R.; Vats, P.; Kumar-Sinha, C.; Kawaler, E. A.; Omelchenko, T.; Colaprico, A.; Geffen, Y.; Maruvka, Y. E.; da Veiga Leprevost, F.; Wiznerowicz, M.; Gümüş, Z. H.; Veluswamy, R. R.; Hostetter, G.; Heiman, D. I.; Wyczalkowski, M. A.; Hiltke, T.; Mesri, M.; Kinsinger, C. R.; Boja, E. S.; Omenn, G. S.; Chinnaiyan, A. M.; Rodriguez, H.; Li, Q. K.; Jewell, S. D.; Thiagarajan, M.; Getz, G.; Zhang, B.; Fenyö, D.; Ruggles, K. V.; Cieslik, M. P.; Robles, A. I.; Clauser, K. R.; Govindan, R.; Wang, P.; Nesvizhskii, A. I.; Ding, L.; Mani, D. R.; Carr, S. A. Clinical Proteomic Tumor Analysis Consortium. Proteogenomic Characterization Reveals Therapeutic Vulnerabilities in Lung Adenocarcinoma. *Cell* **2020**, *182*, 200-225.e35.
[130] Plewczynski, D.; Tkacz, A.; Wyrwicz, L. S.; Rychlewski, L. AutoMotif Server: Prediction of Single Residue Post-Translational Modifications in Proteins. *Bioinformatics* **2005**, *21*, 2525–2527.
[131] Daily, K. M.; Radivojac, P.; Dunker, A. K. Intrinsic Disorder and Prote in Modifications: Building an SVM Predictor for Methylation. In *2005 IEEE Symposium on Computational Intelligence in Bioinformatics and Computational Biology*. **2005**, 1–7.
[132] Chen, H.; Xue, Y.; Huang, N.; Yao, X.; Sun, Z. MeMo: A Web Tool for Prediction of Protein Methylation Modifications. *Nucleic Acids Res.* **2006**, *34*, (Web Server issue), W249-253. 3.
[133] Shien, D.-M.; Lee, T.-Y.; Chang, W.-C.; Hsu, J. B.-K.; Horng, J.-T.; Hsu, P.-C.; Wang, T.-Y.; Huang, H.-D. Incorporating Structural Characteristics for Identification of Protein Methylation Sites. *J. Comput. Chem.* **2009**, *30*, 1532–1543.
[134] Shi, S.-P.; Qiu, J.-D.; Sun, X.-Y.; Suo, S.-B.; Huang, S.-Y.; Liang, R.-P. PMeS: Prediction of Methylation Sites Based on Enhanced Feature Encoding Scheme. *PLoS One* **2012**, *7*, e38772.
[135] Shi, S.-P.; Qiu, J.-D.; Sun, X.-Y.; Suo, S.-B.; Huang, S.-Y.; Liang, R.-P. PLMLA: Prediction of Lysine Methylation and Lysine Acetylation by Combining Multiple Features. *Mol. Biosyst.* **2012**, *8*, 1520–1527
[136] Deng, W.; Wang, Y.; Ma, L.; Zhang, Y.; Ullah, S.; Xue, Y. Computational Prediction of Methylation Types of Covalently Modified Lysine and Arginine Residues in Proteins. *Brief. Bioinform.* **2017**, *18*, 647–658.

Supporting Information

Supporting information is available online at http://dx.doi.org/10.14889/jpdm.2021.0001.