The histone code: what proteomics has taught us

Joshua Peterson Bsc

May 2014

Joshua Peterson

Cancer Stem cells and Developmental Biology

Graduate School of Life Sciences

Utrecht University

Supervisor: Dr. Michiel Vermeulen, Faculty of Medicine, University Medical Centre Utrecht, Dept. Molecular Cancer Research

Examiner: Prof Dr. Mark Timmers, Faculty of Medicine, University Medical Centre Utrecht, Dept. Molecular Cancer Research

Index

Abstract

Within eukaryotic cells chromatin plays an important role in the regulation of gene expression. Nucleosomes form the fundamental repeating unit within chromatin. Nucleosomes consist each of the four core histones around which ~146 basepairs of DNA is wrapped. Both DNA and histone proteins carry additional information which relates to gene expression. This information is present in the form of chemical modifications known as 'epigenetic marks'. Aberrant epigenetic regulation can result in disease. Histone post-translational modifications (PTMs) are associated with various functions within the cell and miss-regulation of histone PTM patterns have been linked to cancer. Historically many studies have analyzed histone PTMs using site-specific antibodies. However there are major drawbacks in using this approach. Mass spectrometry has been introduced as the tool of choice for studying novel epigenetic structures and understanding the role of histone PTMs. This review will cover the recent developments and applications of mass spectrometry in the field of chromatin biology.

Introduction

During embryonic development the inner cell mass of the embryo gives rise to all cells of the body through cellular proliferation and differentiation. Cellular differentiation can be viewed as a epigenetic process during which cellular phenotype is changed but the genomic DNA sequences remain the same.[1]

Gene expression is regulated through chromatin activity. Chromatin is a complex of DNA and proteins that forms chromosomes within the nucleus of eukaryotic cells. Chromatin is build up of nucleosomes. A nucleosome is composed of ~147 bp of DNA wrapped around a set of eight proteins called histones, which are known as a histone octamer. Each histone octamer is made out of two copies of the histone proteins H2A, H2B, H3, and H4.[2] The chain of nucleosomes are compacted by multiple proteins including the linker histone H1 to form the highly organized chromatin as well as the chromosomes. DNA and histone proteins can carry chemical modifications known as 'epigenetic marks'. DNA can be found to be methylated at cytosines in CpG dinucleotides and histones can be modified to a great number of post-translational modifications such as phosphorylation, ubiquitylation, methylation and acetylation. Together these epigenetic marks form the 'epigenome' which is known to influence the behavior of cellular chromatin.[3,4] Modifications of the chromatin state in a cell will result in changes in DNA-associated processes, such as DNA repair, replication and transcription. Therefore problems in epigenetic regulation can lead to genetic lesions and the loss of cell identity. Aberrant chromatin modification patterns have been reported in tissues collected from early stage and benign tumors.[5]

These observations suggest that abnormal chromatin modification patterns could precede and possible facilitate oncogenesis. Unlike genetic mutations, epigenetic marks can be added or removed through normal cellular processes. This has driven physicians and researchers to investigate the effects of small-molecule inhibitors that target epigenetic processes.[6] The development of these novel 'epigenetic therapies' will require detailed knowledge of the mechanisms involved in chromatin regulation. More research is required to better understand how epigenetic modifications influence cellular functions. Mass-spectrometric and proteomic methods have been applied to provide novel insights in epigenetics research.

In this review, we will focus on the recent developments and applications of mass spectrometry in the field of chromatin biology.

Mass spectrometry as a tool to study chromatin biology

The main function of the mass spectrometer in proteomics is to analyze proteins and peptides. Mass spectrometry (MS) is proving to be a valuable tool for the study of histone post translational modifications (PTMs). MS high sensitivity and accuracy is ideal for the identification of novel epigenetic structures and understanding the role of histone PTMs.

Traditionally, specific PTMs located on histones where detected using immunoblotting assays. This however raised the problem of antibody specificity where one antibody might interact with similar modifications on another histone protein. MS has proven to be a more unbiased technique for discovering and quantifying histone proteins.[7,8] Furthermore when MS is used in conjunction with standard biological techniques, it can compare protein expression differences between 'normal' versus for example cancerous cells.[9,10]

Histones intrinsic properties have made it difficult to study them using the "standard" proteomic approach (Fig.1). This is because histones are very basic, and have a high frequency and heterogeneity of PTMs. Digestion of proteins with trypsin, which cleaves the lysine and arginine residues, will result in many short overlapping peptides. In addition PTMs such as acetylation and methylation can block the cleavage site of trypsin leading to non-reproducible peptides. This makes label-free quantization of specific modification sites very difficult. This resulted in the development of novel techniques for the analysis of histone PTMs using MS.

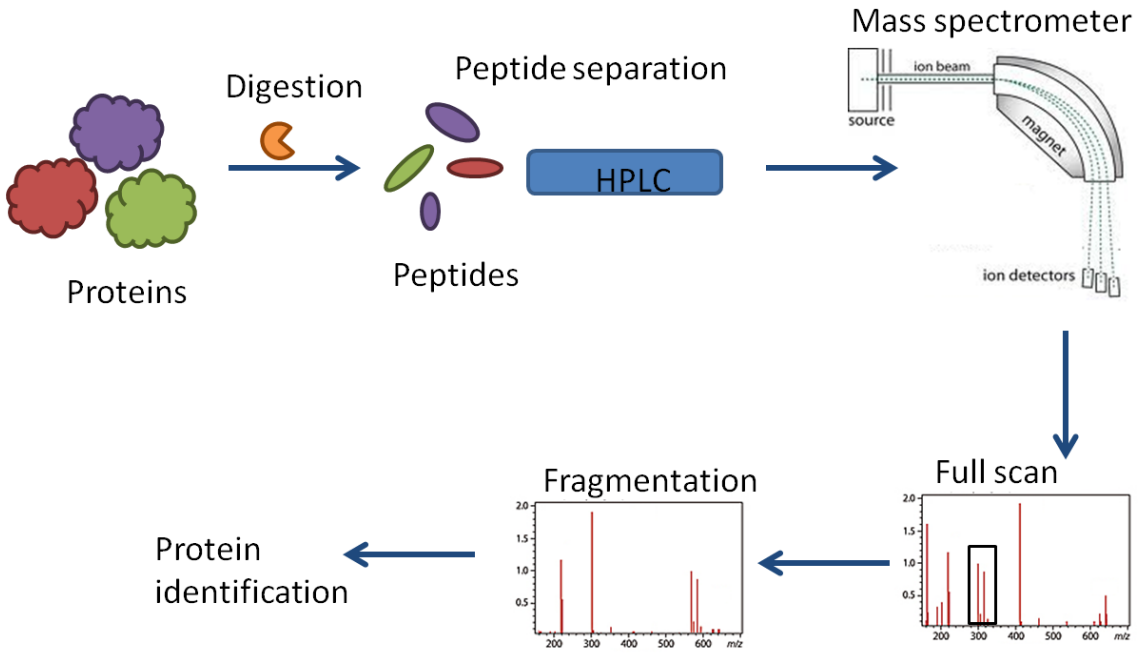

Figure 1: Standard proteomics workflow. This 'shotgun' approach is commonly used for analyzing cellular proteins. Proteins are first collected and digested using enzymes. The peptide mixture is then separated through nanoHPLC and electrosprayed by means of an applied voltage. The mass spectrometer can then determine the precursor masses which it records in the full scan. A selection from this full scan can be further fragmented and analyzed again (MS/MS). Eventually the proteins are identified by a database search

Bottom-up MS is often used for the identification of proteins and characterization of amino acid sequences and PTMs. What defines bottom-up MS is that proteins are digested prior to analysis.[11,12] This approach has the advantage of being very sensitive and has led to the finding of many new histone PTMs.[13]

There are however limitations for using bottom-up MS. As a result of digestion a bottom-up MS experiment generally analyses short peptides. This limits the detection of simultaneously occurring long-distance PTMs. This has led to the development of different MS approaches like middle- and top-down MS which allow the study of combinatorial modifications on large polypeptides or even intact proteins, respectively.[14]

Middle-down and top-down approaches have been advanced with the use of electron transfer dissociation (ETD) which allows a more regular fragmentation along the length of the peptide,

without regard to its modification state.[15]

Most middle-down and top-down approaches focus on examining histone H4. Histone H4 is an attractive histone to study because of its inherent properties such as possessing the best separation, ionization and fragmentation properties when compared to other histones.[16]

The histone code

Histone modifications can serve various purposes within the cell. Modifications such as and H3 lysine 27 methylation (H3K27me), H3 serine 10 phosphorylation (H3S10ph), and H3 lysine 9 acetylation (H3K9ac) are well known for their involvement in chromatin remodeling as well as their role in recruitment of gene activation and/or repression complexes.[17]

The great variety of histone modifications quickly led to the idea of a 'histone code'. This histone code hypothesis explains how specific histone modifications located on one or more N-terminal tails can act sequentially or in combination to form a 'histone code' (fig2.) which could be interpreted by specialized proteins to induce a specific cellular response.[18,19]

These specialized proteins can be grouped in three classes, epigenetic "writers", "erasers" and "readers".

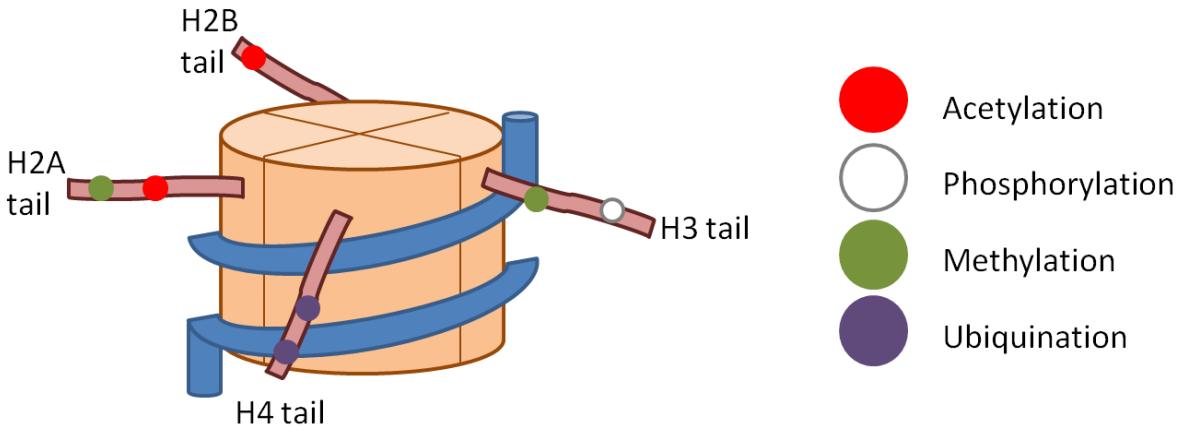

Figure 2: A schematic representation of a nucleosome containing multiple histone tails along with several histone PTMs. Orange represents the histone core, blue indicates DNA, pink represents the protruding histone tails. Histone modifications are labeled accordingly.

Epigenetic writers are specialized proteins which are able to add modifications either on histone tails or the DNA itself. These modifications are known as epigenetic marks, and contribute to gene expression and silencing. Examples of epigenetic "writers" are the histone

methyltransferases, which catalyze the transfer of methyl groups on the exposed histone tail, and histone acetyltransferases, which transfer acetyl groups onto histone lysine residues. Epigenetic “erasers”, as their name implies, are proteins which have the capability to remove epigenetic marks in order to alter gene expression. Multiple categories of epigenetic erasers exist which target histones, these include histone deacetylases, histone threonine/serine/tyrosine phosphatases, histone deubiquitinases and histone lysine/arginine demethylases. The last group known as epigenetic “readers”, specifically recognize certain modifications and exert their function at the site of recruitment. Examples of epigenetic readers include the binding of HP1 proteins to H3K9me3 and the numerous proteins which recognize the H3K4me3 mark such as, ING proteins[20], BPTF[21], PHF8[22], and SGF29[23].

Histone modification patterns have received a lot of interest because of their potential role in carcinogenesis. It has been observed that expression of histone PTMs is altered in cases of acute myeloid leukemia and chronic lymphocytic leukemia.[24]

Others have reported that the loss of the acetylated Lys16 (H4K16ac) and trimethylated Lys20 (H4K20me3) residues of histone H4 were associated with the hypomethylation of DNA sequences which developed early during tumorigenesis and accumulated over time.[25] Hypomethylation of DNA is one of the hallmarks of all common cancers. Recently researches have been looking at ways to modify histone PTMS and potential targets are the histone deacetylases. Drugs are currently designed to act as inhibitors of histone deacetylases and are currently in various phases of clinical trials.[26] These drugs can potentially treat several forms of cancer.

Identifying novel histone modifications using MS

MS is an excellent tool for the discovery of novel histone PTMs. For example, an early study of Strahl et al discovered by using a MS approach in combination with antibody labeling the arginine methylation of histone H4 (H4R3me1) to be linked to transcriptional repression.[27] This event was mediated by a protein PRMT, a methyltransferase.

Novel lysine methylation sites have also been reported using proteomic interrogation techniques. Weis et al, have characterized a new methylation site in the C-terminus of H1 which is the target of G9a/Glp1 both in vitro and in vivo.[28] The ‘top-down’ mass spectrometry approach can be used for identifying PTMs on full-length histones. This approach was used for the identification of a histone mark on lysine 37 of histone H2B which is dimethylated in yeast.[29]

MS is not restricted to identifying methylation marks, as acetylation sites have also been identified using MS. For example, H3K56ac and H3K36ac have been linked to the core transcriptional network in human embryonic stem cells.[30]

While most histone modifications reside on the N-tails few modifications have been identified within the histone core sequences. Using a MS approach it was found that threonine 45 of histone H3 (H3T45) plays a role in DNA replication in S. cerevisiae.[31] and as a site of active phosphorylation in apoptotic neutrophils in-vivo by protein kinase C-delta.[32]

The combination both in vivo biological and MS approaches has been used to identify *O*-GlcNAcylation as part of the histone code on H2A, H2B and H4. *O*-GlcNAcylation changes during mitosis and with heat shock.[33] Other atypical PTMs in addition to the *O*-GlcNA PTM, such as ADP-ribosylation, butyrylation, propionylation, and formylation have also been discovered using careful MS analysis.[33–36]

Quantification of histone PTMs

One of the major problems facing quantitative MS is to consistently obtain a reproducible digestion of histones peptides. As discussed earlier histones are rich in lysine and arginine residues making these proteins highly basic. On top of that, when histones are digested with simple proteases such as trypsin, which cleaves exclusively C-terminal to arginine and lysine residues, then a peptide mixture is obtained containing many short overlapping peptides containing the same resides.[37] This makes quantification of specific modification sites very challenging.

To circumvent the problems which arise using trypsin, researchers have looked into the use of other proteases which cleave at only one amino acid such as ArgC and GluC. The Arg-C protease preferentially cleaves at arginine in position P1 and GluC cleaves C-terminal to glutamate.[38,39]

This approach has successfully been demonstrated in 2004 by McKittrick et al. In this study they observed that histone H3.3 possesses marks associated with euchromatin.[40] Another study used a similar approach for examining the PTMs on H3.1 and H3.3 during chromatin assembly , they reported that lysine methylation is absent prior to histone incorporation into chromatin, except at H3K9.[41] The use of other proteases for digestion comes with certain downsides. For example, when comparing ArgC to trypsin there is a difference in efficiency and specificity. ArgC is less specific and efficient than trypsin and thus requires protocol optimization for obtaining

reproducible peptides.

The mass spectrometer can be used to obtain quantitative information relating to relative or absolute peptide abundance. Histone PTMs can be quantified based on the abundances relative to each other by a number of proteomic methods this could be a isotopically labeled approach or a label-free approach (Fig. 3).

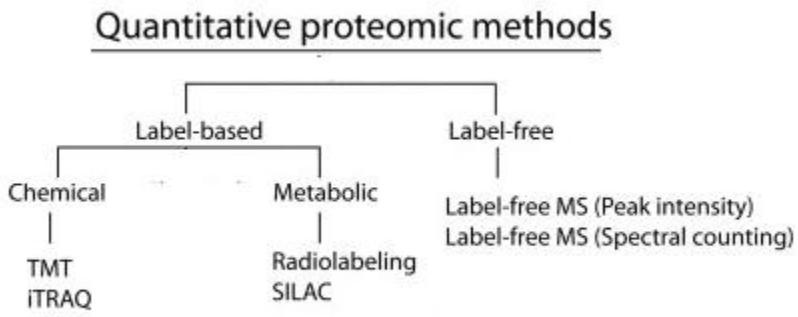

Figure 3: Overview of commonly used quantitative proteomic methods. Quantitative proteomic methods can be divided into label-based and label free methods. Label-based methods can be further divided into chemically or metabolically labeled.

Label-free quantification approaches come in various flavors the simplest being spectral counting and exponentially modified protein abundance index (emPAI).[42 ,43]

Both these approaches are based on the logic that abundant proteins in a sample is likely to result in more MS/MS fragmentation spectra than low-abundance proteins. The advantage of these techniques are that they can be used to any qualitative dataset. Label-free methods are relatively accurate because the ionization efficiencies of the same peptides remain the same. A major disadvantage of label-free techniques is that samples have to be individually processed and measured resulting in experimental variability. A partial solution to this problem is the use of intelligent label-free algorithms like MaxQuant which can correct for the variability between samples.[44]

In contrast to label-free methods there is the option to use stable isotope labels which can be used for protein quantification. Stable isotope labeling introduces a specific mass tag which can then be recognized by a mass spectrometer and provide quantitative information. Isotope labeling can be achieved using two methods, either chemical labeling or metabolic labeling. Chemical labeling is performed by introducing specific mass tags onto the reactive groups

present in peptides.[45] Commonly used methods of chemical labeling are iTRAQ and TMT.[46] What makes iTRAQ and TMT popular is that these methods allow for multiplexing and a relatively large number of specific experimental conditions can be compared in the same run.

Even though chemical labeling techniques can be applied to virtually all sample types, a general disadvantage is the introduction of chemical derivates into proteins. This can hinder the detection of rare PTMs.

Metabolic labeling techniques rely on the incorporation of heavy isotopes through metabolic pathways. This can be achieved thru addition of heavy isotopes into growth medium or in the food supply of experimental models. The advantage of this method is that the entire proteome is labeled, which allows samples to be mixed at an early time point during the experiment. This in turn minimizes experimental errors. One approach to introduce heavy isotopes is by replacing all nitrogen atoms in medium by the heavy isoform $_{15}$N.[47] Other approaches include the replacement of essential amino acids with a heavy isoform, this method is known as SILAC (stable isotope labeling by amino acids in cell culture).[48]

SILAC is a recent development which has been designed specifically for peptide quantization. It relies on the incorporation of non-radioactive, isotopically labeled amino acids into cultured cells.[45] In a SILAC setup cells are isotopically labeled either “light” (normal) or “heavy”. This approach makes peptides from two experimental states indistinguishable. Each peptide is therefore present as a light and heavy pair in the same mass spectrum. The ratio of their signals is the ratio of relative protein abundance under the two sets of experimental conditions.[49]

The quantification approach is significantly more accurate when compared to label-free methods.

SILAC experiments which study histone PTMs are becoming more prevalent. SILAC experiments have been used to analyze the core histone modifications which occur throughout the cell cycle. Bonenfant et al. showed that mitosis was a specific period of the cell cycle where many modifications exhibit dynamic changes.[50] During mitosis histone H3 and H4 phosphorylation increased and H3K27/K36 methylation decreased.[50]

SILAC has also been used to show methylation of histone H4 at lysine 20 (H4K20) during the cell cycle.[51] Others have used SILAC approaches to dissect the "histone modification signature" of human breast cancer cells.[52]

Another study used a SILAC approach in which 13CD3-methionine is introduced into methionine-depleted cell culture media. By taking this approach the cell is forced to use the heavy labeled methionine to synthesize *S*-adenosyl methionine. This will result in a heavy

labeled methyl donor group. This approach has been used to investigate the dynamics of histone methylation within synchronized cell populations.[53] Another study used SILAC labeling combined with this heavy methyl labeling to study the effects of H3K79me throughout the cell cycle and the turnover rates between the different H3K27K36 methylation states.[54, 55]

Gene locus-specific histone analysis

Current MS approaches generally isolate all histones from a sample before analysis. A next step would be to isolate a specific gene or genomic region to examine how the local chromatin environment contributes to gene expression. This would provide valuable information for studying how the chromatin environment relates to disease-associated genes and might reveal how epigenetics contribute to the underlying logic that governs the networks of gene expression within a cell.

In 2009 a study reported A SILAC-based DNA protein interaction screen that identifies candidate binding proteins to functional DNA elements.[56] Later that year another group published a report which demonstrated how to purify proteins associated with specific genomic loci.[57] This method makes use of specific nucleic acid probes to isolate genomic DNA based on affinity. Recently a study reported a technique which allows researchers to examine the differences between the transcriptionally active and silent states of a single genomic locus.[58] This method is called chromatin affinity purification with mass spectrometry (ChAP-MS) and they later enhanced the ChAP approach and called it TAL-ChAP-MS, which allows for the biochemical isolation of native chromatin sections without genomic engineering. This is achieved by performing purification of specialized TAL proteins prior to ChAP. TAL proteins recognize certain DNA sequences enabling the selective enrichment of a certain genomic loci.[59] Other methods like ChIP-MS have also been designed. These methods combine chromatin immuno precipitation (ChIP) approaches like ChIP-ChIP and ChIP-Seq to enrich for specific genomic populations prior to MS.[60] A recent study has introduced a refined approach of ChIP-MS termed ChroP (Chromatin Proteomics). The ChroP approach has been designed to combine ChIP and MS to dissect locus-specific proteomic landscapes of chromatin.[61] The combined use of ChIP, next-generation sequencing, and MS is a valuable toolset which can greatly contribute to our current understanding of chromatin biology. However there are some limitations still to overcome. One of these limitations is that chromatin immunoprecipitation relies on difficult to obtain good-quality antibodies. The great variability of histone modifications is an significant issue for developing

specific antibodies.[62] A related problem is that antibody-based techniques can only address a few modifications at a time. This limits the use of PTM-specific high throughput analysis. Future studies will have to develop novel approaches to handle these limitations.

Identifying epigenetic readers

MS has been employed for the identification of histone binding proteins. Histone modifications are mostly thought of as modifications found on the N-terminal histone tails which extend from the nucleosome as flexible structures. This idea let to the assumption that isolated peptides can give a good approximation of histone tails. Early studies preformed to identify histone PTM binding proteins made use of synthetic peptides which would carry trimethylated lysines to resemble the N-terminal tail of histones. Proteins which would bind to the histone modifications where then eluted and examined on an SDS-PAGE gel. Proteins where then excised and identified using MS. Experimental procedures like these resulted in the finding that the NuRD complex associates with unmodified histone H3 peptides and that binding is inhibited because of local trimethylation at lysine 4.[63] Lysine 4 methylation is commonly found on locations of active gene transcription. This suggests that lysine 4 methylation might disrupt the association of histones with a repressor complex.[63] Another study using a similar approach found that the NURF subunit BPTF was identified to interact with H3K4me3-modified histone peptides and show that this is required to maintain Hox gene expression patterns during development.[64] This peptide pull-down techniques was later combined with SILAC based quantitative proteomics. This approach was used to identify novel proteins which could interact with H3K4me3-modified histone tails like the transcription factor TFIID.[65] This finding has shown valuable for understanding the regulation of RNA polymerase II-mediated transcription. Later studies used this SILAC peptide pull-down approach to investigate numerous major lysine trimethilation marks in histone H3 (H3K4me3, H3K9me3, H3K27me3, and H3K36me) and histone H4 (H4K20me3).[23]

Aside from histone modifications located on the N-terminal of the histone tails, histone PTMs can also be embedded into the chromatin environment. The structural context of nucleosomes can contribute to the specificity of which epigenetic readers can recognize the modification.

In 2010 Bartke et al examined the “cross-talk” between DNA and histone modifications within

chromatin. In this study they made use of nucleosomes methylated on DNA and on histone H3 as baits in SILAC nucleosome affinity purifications.[66] Their analysis reveals that certain proteins which bind to nucleosomes are regulated by combinations of epigenetic marks on a nucleosomal template. These can be combinations of methylation of CpGs, H3K4, H3K9, and H3K27.

Other studies interested in the relation between the chromatin state and gene expression used an approach based on modified chromatin templates used in affinity purification experiments prior to MS analysis to identify the interactome of histone modification patterns. They compared epigenetic readers found with chromatin affinity purification to the classic histone N-terminal peptide affinity purification approach and found that only some factors associate with both chromatin and peptide matrices but that a large number of proteins differ in their association with these templates.[67] The data collected by peptide, nucleosome and chromatin purification shows overlap between some epigenetic readers but there are also differences to be found. Further studies are required to address these observations.

The use of immobilized DNA oligonucleotides has been reported as an approach for SILAC affinity purification. This approach has been used to determine methyl CpG-binding proteins and transcription factors by incorporating either methylated CpG dinucleotides or single base pair changes into the oligonucleotide sequence to act like an affinity bait.[66, 68, 69]

To study the role of RNA interaction proteins a unbiased screening technique has been developed which modified the oligonucleotide arrays to incorporate RNAs as bait for SILAC affinity purification.[70] This technique holds allot of value for researchers which are now able to study how RNA interaction, and non-coding RNAs contribute to epigenetic regulation.

Conclusion

The recent technological advancements in proteomic methodologies have greatly contributed to our current understanding of chromatin biology and epigenetics.

Today it is possible to quantify massive amounts of peptides and proteins in complex biological mixtures. We are able to scan through complete histone tails to study the histone code and are able to identify epigenetic readers using state of the art purification techniques. The development of ChIP-MS allows us to globally study all of the proteomic interactions within chromatin. Technological developments and overall decreasing costs is proving beneficial for researchers trying to understand the molecular interactions between histone PTMs and genes.

The role of epigenetic mechanisms and specifically histone PTMs is slowly revealing to play a major role in the development of several human disorders and diseases.

Modifications in histone PTMs has been shown to drive altered gene-expression which in turn can be directly linked to disease progression. For this reason, quantification of histone PTMs located on disease-associated genes will become increasingly interesting. One obstacle, however, is the identification of proteins at a single genomic loci. Recent advancements like TAL-ChAP-MS and ChroP-MS both have allot of potential and future studies will have to determine whether these approaches will be successful in analyzing proteins at single genomic loci. The recent development of the CRISPR-Cas9 technology opens a new window of opportunities for chromatin biology. The CRISPR (clustered regularly interspaced short palindromic repeats) system is a novel nuclease based approach to gene editing which makes use of a bacterially derived protein Cas9 to introduce site specific DNA breaks into the genome. The two components of CRISPR are the Cas9 protein, which cuts the DNA, and a small guide RNA molecule (gRNA) which shares a region of about 20-nt homology to direct Cas9 to its target DNA thru hybridization.[71] This system can be used to excise certain genes from the DNA which can be enriched for closer examination of the role of histone PTMs on a specific gene or gene-locus. However off target effects have been observed using the native Cas9 protein. To accommodate for off-target effects a modified Cas9 protein has been introduced which can be used as a nickase to induce single stranded breaks. By fusing two gRNAs targeted to opposite ends of the DNA to a Cas9 nickase, a double stranded break will be introduced with reduced off-target effects. Other applications of CRISPR can be to study changed chromatin states by fusing a histone modifying enzyme to Cas9 to alter histone PTMs.

Another promising development is the use of small-molecule inhibitors as bait to perform affinity purification of chromatin modulating proteins and epigenetic readers. This chemical proteomic approach can be challenging but has already been proven to be a valuable tool for drug discovery.[72]

The implementation of chemical proteomic strategies has only recently been initiated and future developments are bound to contribute greatly to our understanding of inhibitor specificities and aid in the study to identify the functions of inhibitor target proteins.

With the development of new technologies combined with novel MS approaches, MS-based proteomics will continue to reveal novel insights into the fundamentals concerning epigenetic processes and improve our abilities to understand the mechanics of chromatin regulation.

Acknowledgements

I would like to express my gratitude to Dr. Michiel Vermeulen for the supervision during the writhing of this thesis, and thank Prof. Dr. Mark Timmers for being the examiner.

References

1. Berger SL, Kouzarides T, Shiekhattar R, et al. An operational definition of epigenetics. Genes Dev 2009;23:781–3.

2. Luger K, Mader AW, Richmond RK, et al. Crystal structure of the nucleosome core particle at 2.8 A resolution. Nature 1997;389:251–60.

3. Jones PA. Functions of DNA methylation: islands, start sites, gene bodies and beyond. Nat Rev Genet 2012;13:484–92.

4. Kouzarides T. Chromatin modifications and their function. Cell 2007;128:693–705.

5. Feinberg AP, Ohlsson R, Henikoff S. The epigenetic progenitor origin of human cancer. Nat Rev Genet 2006;7: 21–33.

6. Rodriguez-Paredes M, Esteller M. Cancer epigenetics reaches mainstream oncology. NatMed 2011;17:330–9.

7. Brumbaugh J, Phanstiel D, Coon JJ. Unraveling the histone's potential: a proteomics perspective. Epigenetics. 2008; 3(5):254–257.

8. Trelle MB, Jensen ON. Functional proteomics in histone research and epigenetics. Expert Rev. Proteomics. 2007; 4(4):491–503.

9. Evertts AG, Zee BM, Garcia BA. Modern approaches for investigating epigenetic signaling pathways. J. Appl. Physiol. 2010; 109(3):927–933.

10. Eberl HC, Mann M, Vermeulen M. Quantitative proteomics for epigenetics. ChemBioChem. 2010; 12(2):224–234.

11 Aebersold R, Mann M. "Mass spectrometry-based proteomics". Nature 422, 2003, (6928): 198–207.

12. Chait BT. "Chemistry. Mass spectrometry: bottom-up or top-down?". Science 314, 2006, (5796): 65–6.

13. Yates JR, Ruse CI, Nakorchevsky A. "Proteomics by Mass Spectrometry: Approaches, Advances, and Applications". Annu. Rev. Biomed. Eng. 11, 2009, 49–79.

14. Young NL, Plazas-Mayorca MD, Garcia BA. Systems-wide proteomic characterization of combinatorial post-translational modification patterns. Expert Rev. Proteomics. 2010; 7(1):79–92.

15. Syka JE, Coon JJ, Schroeder MJ, Shabanowitz J, Hunt DF. Peptide and protein sequence analysis by electron transfer dissociation mass spectrometry. Proc. Natl Acad. Sci. USA. 2004; 101(26): 9528–9533.

16. Garcia BA. What does the future hold for top down mass spectrometry? J. Am. Soc. Mass

Spectrom. 2010; 21(2):193–202.
17. Campos EI, Reinberg D. Histones: annotating chromatin. Annu. Rev. Genet. 2009; 43:559–599.
18 Jenuwein T, Allis CD. Translating the histone code. Science. 2001; 293(5532):1074–1080.
19. Strahl BD, Allis CD. The language of covalent histone modifications. Nature. 2000; 403(6765):41– 45.
20. P.V. Peña, F. Davrazou, X. Shi, K.L. Walter, V.V. Verkhusha, O. Gozani, R. Zhao, T.G. Kutateladze Molecular mechanism of histone H3K4me3 recognition by plant homeodomain of ING2 Nature, 442 (2006), pp. 100–103
21. H. Li, S. Ilin, W. Wang, E.M. Duncan, J. Wysocka, C.D. Allis, D.J. Patel Molecular basis for site-specific read-out of histone H3K4me3 by the BPTF PHD finger of NURF Nature, 442 (2006), pp. 91–95
22. W. Feng, M. Yonezawa, J. Ye, T. Jenuwein, I. Grummt PHF8 activates transcription of rRNA genes through H3K4me3 binding and H3K9me1/2 demethylation Nat. Struct. Mol. Biol., 17 (2010), pp. 445–450
23. M. Vermeulen, H.C. Eberl, F. Matarese, H. Marks, S. Denissov, F. Butter, K.K. Lee, J.V. Olsen, A.A. Hyman, H.G. Stunnenberg, M. Mann Quantitative interaction proteomics and genome-wide profiling of epigenetic histone marks and their readers Cell, 142 (2010), pp. 967–980
24. Zhang L, Freitas MA, Wickham J, et al. Differential expression of histone post-translational modifications in acute myeloid and chronic lymphocytic leukemia determined by high-pressure liquid chromatography and mass spectrometry. J. Am. Soc. Mass Spectrom. 2004; 15(1):77–86.
25. Fraga MF, Ballestar E, Villar-Garea A, et al. Loss of acetylation at Lys16 and trimethylation at Lys20 of histone H4 is a common hallmark of human cancer. Nat. Genet. 2005; 37(4):391–400.
26. Minucci S, Pelicci PG. Histone deacetylase inhibitors and the promise of epigenetic (and more) treatments for cancer. Nat. Rev. Cancer. 2006; 6(1):38–51.
27. Strahl BD, Briggs SD, Brame CJ, et al. Methylation of histone H4 at arginine 3 occurs *in vivo* and is mediated by the nuclear receptor coactivator PRMT1. Curr. Biol. 2001; 11(12):996–1000.
28. Weiss T, Hergeth S, Zeissler U, et al. Histone H1 variant-specific lysine methylation by G9a/KMT1C and Glp1/KMT1D. Epigenetics Chromatin. 2010; 3(1):7.
29. Gardner KE, Zhou L, Parra MA, Chen X, Strahl BD. Identification of lysine 37 of histone H2B as a novel site of methylation. PLoS One. 2011; 6(1).

30. Xie W, Song C, Young NL, et al. Histone h3 lysine 56 acetylation is linked to the core transcriptional network in human embryonic stem cells. Mol. Cell. 2009; 33(4):417–427.
31. Baker SP, Phillips J, Anderson S, et al. Histone H3 Thr 45 phosphorylation is a replication associated post-translational modification in S. cerevisiae. Nat. Cell. Biol. 2010; 12(3):294–298.
32. Hurd PJ, Bannister AJ, Halls K, Dawson MA, Vermeulen M, Olsen JV, Ismail H,Somers J, Mann M, Owen-Hughes T, Gout I, Kouzarides T. Phosphorylation of histone H3 Thr-45 is linked to apoptosis. J Biol Chem. 2009 Jun 12;284(24):16575-83.
33. Sakabe K, Wang Z, Hart GW. β-*N*-acetylglucosamine (*O*-GlcNAc) is part of the histone code. Proc. Natl Acad. Sci. USA. 2010; 107(46):19915–19920.
34. Wisniewski JR, Zougman A, Mann M. Nepsilon-formylation of lysine is a widespread posttranslational modification of nuclear proteins occurring at residues involved in regulation of chromatin function. Nucleic Acids Res. 2008; 36(2):570–577.
35. Messner S, Altmeyer M, Zhao H, et al. PARP1 ADP-ribosylates lysine residues of the core histone tails. Nucleic Acids Res. 2010; 38(19):6350–6362.
36. Chen Y, Sprung R, Tang Y, et al. Lysine propionylation and butyrylation are novel posttranslational modifications in histones. Mol. Cell. Proteomics. 2007; 6(5):812–819.
37. Olsen JV, Ong SE, Mann M. Trypsin cleaves exclusively C-terminal to arginine and lysine residues. Mol Cell Proteomics. 2004 Jun;3(6):608-14.
38. Keil, B. Specificity of proteolysis. Springer-Verlag Berlin-Heidelberg-NewYork, pp.335. (1992
39.. Drapeau, G.R. Methods Enzymol., 45, 469-475, (1976).
40. McKittrick E, Gaften PR, Ahmad K, et al. Histone H3.3 is enriched in covalent modifications associated with active chromatin. Proceedings of the National Academy of Sciences of the United States of America 2004;1016:1525–1530.
41. Loyola A, Bonaldi T, Roche D, et al. PTMs on H3 variants before chromatin assembly potentiate their final epigenetic state. Mol Cell 2006;24:309–16.
42. H. Liu, R. G. Sadygov, J. R. Yates III, Anal. Chem. 2004, 76, 4193.
43. Y. Ishihama, Y. Oda, T. Tabata, T. Sato, T. Nagasu, J. Rappsilber, M. Mann, Mol. Cell. Proteomics 2005, 4, 1265.
44. C. A. Luber, J. Cox, H. Lauterbach, B. Fancke, M. Selbach, J. Tschopp, S. Akira, M. Wiegand, H. Hochrein, M. O'Keeffe, M. Mann, Immunity 2010, 32, 279.
45.Bantscheff M, Schirle M, Sweetman G, Rick J, Kuster B. Quantitative mass spectrometry in proteomics: a critical review. Anal Bioanal Chem. 2007 Oct;389(4):1017-31. Review.
46 .A. Thompson, J. Schafer, K. Kuhn, S. Kienle, J. Schwarz, G. Schmidt, T.Neumann, R.

Johnstone, A. K. Mohammed, C. Hamon, Anal. Chem. 2003, 75, 1895.
47.Y. Oda, K. Huang, F. R. Cross, D. Cowburn, B. T. Chait, Proc. Natl. Acad.
Sci. USA 1999, 96, 6591.
48. S. E. Ong, B. Blagoev, I. Kratchmarova, D. B. Kristensen, H. Steen, A. Pandey, M. Mann, Mol. Cell. Proteomics 2002, 1, 376.
49. Eberl HC, Mann M, Vermeulen M. Quantitative proteomics for epigenetics. Chembiochem. 2011 Jan 24;12(2):224-34.
50. Bonenfant D, Towbin H, Coulot M, Schindler P, Mueller DR, van Oostrum J. Analysis of dynamic changes in post-translational modifications of human histones during cell cycle by mass spectrometry. Mol. Cell. Proteomics. 2007; 6(11):1917–1932.
51. Pesavento JJ, Yang H, Kelleher NL, Mizzen CA. Certain and progressive methylation of histone H4 at lysine 20 during the cell cycle. Mol. Cell. Biol. 2008; 28(1):468–486.
52. Cuomo A, Moretti S, Minucci S, Bonaldi T. SILAC-based proteomic ana lysis to dissect the 'histone modification signature' of human breast cancer cells. Amino Acids. 2011; 41(2):387–399.
53. Zee BM, Levin RS, Xu B, et al. In vivo residue-specific histone methylation dynamics. J Biol Chem 2010;285:3341–50.
54. Sweet SM, Li M, Thomas PM, et al. Kinetics of re-establishing H3K79 methylation marks in global human chromatin. J Biol Chem 2010;285:32778–86.
55. Zheng Y, Sweet SM, Popovic R, etal. Total kinetic analysis reveals how combinatorial methylation patterns are established on lysines 27 and 36 of histone H3. ProcNatlAcad Sci USA 2012;109:13549–54.
56. Tackett AJ, Dilworth DJ, Davey MJ, et al. Proteomic and genomic characterization of chromatin complexes at a boundary. J. Cell Biol. 2005; 169(1):35–47.
57. Mittler G, Butter F, Mann M. A SILAC-based DNA protein interaction screen that identifies candidate binding proteins to functional DNA elements. Genome Res. 2009; 19(2):284–293.
58. Byrum SD, Raman A, Taverna SD, Tackett AJ. ChAP-MS: a method for identification of proteins and histone posttranslational modifications at a single genomic locus. Cell Rep. 2012 Jul 26;2(1):198-205.
59. Byrum SD, Taverna SD, Tackett AJ. Purification of a specific native genomic locus for proteomic analysis. Nucleic Acids Res. 2013 Nov 1;41(20)
60. Fuchs SM, Krajewski K, Baker RW, Miller VL, Strahl BD. Influence of combinatorial histone modifications on antibody and effector protein recognition. Curr. Biol. 2011; 21(1):53–58.

61. Soldi M, Bonaldi T. The ChroP Approach Combines ChIP and Mass Spectrometry to Dissect Locus-specific Proteomic Landscapes of Chromatin. J Vis Exp. 2014 Apr 11;(86)
62. T.A. Egelhofer, A. Minoda, S. Klugman, K. Lee, P. Kolasinska-Zwierz, A.A. Alekseyenkoet al. An assessment of histone-modification antibody quality. Nat Struct Mol Biol, 18 (2011), pp. 91–93
63. Zegerman P, Canas B, Pappin D, et al. Histone H3 lysine 4 methylation disrupts binding of nucleosome remodeling and deacetylase (NuRD) repressor complex. J Biol Chem 2002;277:11621–4.
64. Wysocka J, Swigut T, Xiao H, et al. A PHD finger of NURF couples histone H3 lysine 4 trimethylation with chromatin remodelling. Nature 2006;442:86–90.
65. Vermeulen M, Mulder KW, Denissov S, et al. Selective anchoring of TFIID to nucleosomes by trimethylation of histone H3 lysine 4. Cell 2007;131:58–69.
66. Bartke T, Vermeulen M, Xhemalce B, et al. Nucleosomeinteracting proteins regulated by DNA and histone methylation. Cell 2010;143:470–84.
67. Nikolov M, Stutzer A, Mosch K, et al. Chromatin affinity purification and quantitative mass spectrometry defining the interactome of histone modification patterns. Mol Cell Proteomics 2011;10.
68. Bartels SJ, Spruijt CG, Brinkman AB, et al. A SILAC-based screen for Methyl-CpG binding proteins identifies RBP-J as a DNA methylation and sequence-specific binding protein. PLoS One 2011.
69. Markljung E, Jiang L, Jaffe JD, et al. ZBED6, a novel transcription factor derived from a domesticated DNA transposon regulates IGF2 expression and muscle growth. PLoS Biol 2009;7.
70. Butter F, Scheibe M, Morl M, et al. Unbiased RNA protein interaction screen by quantitative proteomics. Proc Natl Acad Sci USA 2009;106:10626–31.
71. Yu Z, Ren M, Wang Z, Zhang B, Rong YS, Jiao R, Gao G. Highly efficient genome modifications mediated by CRISPR/Cas9 in Drosophila. Genetics. 2013 Sep;195(1):289-91. PubMed
72. Bantscheff M, Eberhard D, Abraham Y, Bastuck S, Boesche M, Hobson S, Mathieson T, Perrin J, Raida M, Rau C, Reader V, Sweetman G, Bauer A, Bouwmeester T, Hopf C, Kruse U, Neubauer G, Ramsden N, Rick J, Kuster B, Drewes G. Quantitative chemical proteomics reveals mechanisms of action of clinical ABL kinase inhibitors. Nat Biotechnol. 2007 Sep;25(9):1035-44.