The challenge of complexity: High-content peptide libraries to display sequence diversity and posttranslational modifications

Reimer U.[1]*, Knaute T.[1], Eckey M.[1], Schutkowski M.[2]

[1]JPT Peptide Technologies, Berlin, Germany

[2] Institute of Biochemistry & Biotechnology, Martin-Luther-University, Halle, Germany

Introduction:

One gene, one protein: this was textbook knowledge from a time when we thought that proteins were identical in every individual. This simple axiom gave way to more complex models that took into account isoforms, splice variants, polymorphisms, mutations, and post-translational modifications. Now, we recognize variability in more than half of the proteins of the human proteome, and 41 % of the UniProt entries contain post-translational modifications. For certain proteins, up to 88% of the primary structure can vary between individuals and 16% of the amino acids can be modified post-translationally as well documented in Histones.

Histones package DNA and thereby are crucial for a variety of processes starting from mitosis to transcription and repair. For all but one Histone multiple isoforms are observed. Moreover, all Histones are characterised by an astonishing level of diversity in form of post-translational modifications (PTM). The theoretical number of possible modifications far exceeds the experimentally observed ones. However, investigation of the role of such modifications is dependent on the availability of suitable tools, methods and appropriate model proteins with defined PTMs. Such model proteins are not easily accessible. Peptide libraries are a good approximation and useful for a number of applications.

There is no single Histone H2B:

- Isoforms are observed for most Histones (Fig. 1)
- Different Genes code for some Histones (e.g. H2B1C, H4)
- PTMs lead to further complexity
- **Example H2B: 1189 protein variants** would be needed for complete coverage of Isoforms and selected PTMs on Lys, Arg and phosphoacceptor sites.
- **H2B:** Complete coverage is achieved using a peptide library of **570 peptides**.

Fig. 1 Alignment of H2B isoforms with potential PTM-sites indicated (only for Phosphorylation, Arg-methylation and Lys-Acylation). The number of missense mutations with consequence for the protein sequence and the maximum allele frequency for a missense mutation is indicated to the right of the sequences. (from [1]).

Peptide Libraries:

Histone Code Peptide Microarrays: consists of

a) Knowledge Based Library: 1300 Peptides

- Histone H2A[1,19,85,108,110], H2B[1,22,42,106], H3[1,23,43,63,105,115], H4[1,14,44,68,81]
- PTMs at all reported sites: KBut, KAc, KProp, KMe1, KMe2, KMe3, RMe1, RMe2a, RMe2s, Cit, pT, pS, pY
- Known PTM combinations (up to 6 PTMs)
- Representation of natural sequence variants
- H1 wildtype scan + published modifications [2]

b) Systematic Library: 2400 Peptides

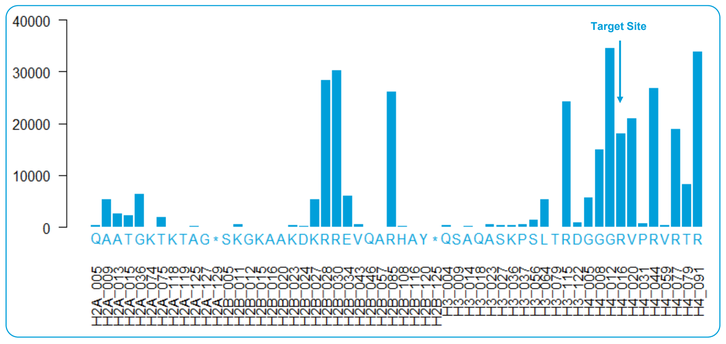
- Focussed scan through histones H2A, H2B, H3, H4 with single modifications (PTMs at all potential sites, + KMal, KSucc)
- Peptides were synthesized using SPOT synthesis and chemo-selectively immobilized onto glass slides.

Acetylome Microarrays: consists of 5599 peptides derived from human acetylation sites on two arrays with Lys in central position (acetylated and free)

SpikeTides™ Set Histone: consists of 37 stable isotope-labeled and absolutely quantified proteotypic peptides for Histone H3 for direct usage in mass-spectrometry based proteomics (MRM). The set comprises wild-type peptides and their Lys(Ac), Lys(Me), Lys(Me2) and Lys(Me3) modifications.

Application: Antibody Profiling

- PTM-specific antibody are important research tools for chromatin research.
- Specificity of such antibodies is not always well characterized
- Such antibodies are frequently raised using peptides.
- Therefore peptide microarrays are perfect tools for characterization of binding specificty. (Fig. 2 and 3)



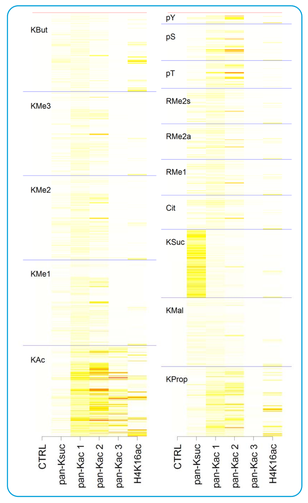


Figure 2 (above). Barplot of Signal Intensities. Interaction of H4K16ac with Kac modifications of main Histone variants (H2A 1-B/E, H2B 1-C/E/F/G/I, H3.1, H4). Shown is the mean signal intensity for all peptides carrying the single modification. The letters below the bars indicate the amino acid following the respective acetylated Lysine. When a star appears the respective lysine is at the C-terminus of the protein.

Figure 3 (left). Heatmap of signal intensities for different Histone- and pan-specific antibodies. Peptides with indicated single modifications are shown on the y-axis. Signal intensity ranges from white (noise) to red (max signal). The map is divided into two parts and each lane represents one antibody

Application: MS-based Proteomics

- Histone Peptides with different PTMs are available for MRM Assay setup and relative/absolute quantification of Histones in biological samples [3].
- High throughput synthesis enables efficient library generation of modified peptides for MRM assay set-up (e.g. acylations, phosphorylations, methylations, glycosylations).
- Peptide sets are well suited for method development (LC and MS).

Summary

- **Peptide libraries are well suited to cover sequence diversity.**
- **Peptide libraries can be used in numerous assay formats.**

Application: Histone Reader

- Several Histone-Reader domains were characterized at the Histone Code array
- Expected binding specificities (e.g. BRD4 [4]) or unexpected binding activities (C-terminal butyrylated Lys in two single isoforms of H2A as targets for Bromodomain, Data not shown) were observed.

Application: Histone Writer

- PRMT5-MEP50 is a protein arginine methyl-transferase
- Substrate specificity was investigated on peptide microarrays [5] (**Fig. 4**)

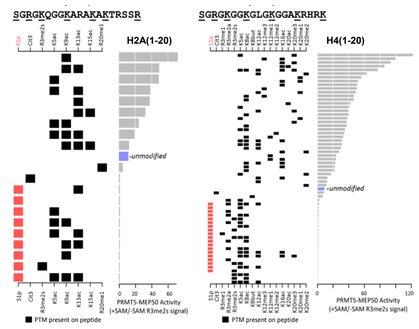


Fig. 4. PRMT5-MEP50 histone methyl-transferase activity is modulated by substrate PTMs.
High-density histone peptide arrays incubated with PRMT5-MEP50 in the presence or absence of SAM. The arrays were probed with anti-H2A/H4R3me2s antibody. Each row represents a discrete peptide.
R3 is the site of modification. Phosphorylation of S1 abrogates enzyme activity.
PRMT5-MEP50 acts consistently on the similar N-terminal tails of Histones H2A and H4.

Application: Histone Eraser

- Comprehensive characterization of substrate specificities for all human Sirtuins [6] (Fig. 5).
- Identification of substrates beyond the Histones.

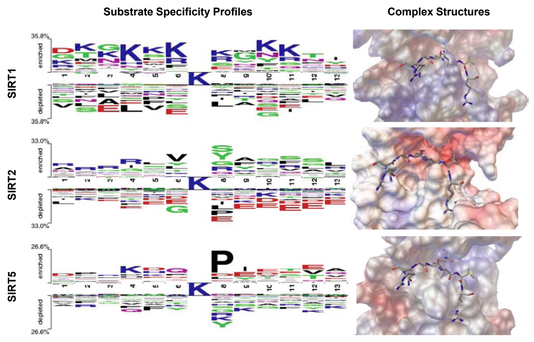


Fig. 5. Substrate specificty profiles (left) and complex models of the binding groove for three Sirtuins. Sequence logoos were calculated using the top 50 substrates from an Acetylome microarray experiment for each Sirtuin as positive sample and all remaining peptides as negative set. Structural models are based on an homology model (Sirt1) or experimental structures (Sirt2 and Sirt5).

References

1. Lek *et al.*, Analysis of protein-coding genetic variation in 60,706 humans. *Nature.* **2016**, 536, 285.
2. Wisniewski *et al.*, Mass spectrometric mapping of linker histone H1 variants reveals multiple acetylations, methylations, and phosphorylation as well as differences between cell culture and tissue. *Mol. Cell Proteomics.* **2007**, *6*, 72.
3. Feller *et al.*, Global and Specific Responses of the Histone Acetylome to Systematic Perturbation. *Mol Cell* **2015**, 57, 559.
4. Masch *et al.*, Peptide Microarrays for Profiling of Epigenetic Targets **2015** *in* Epigenetic Technological Applications, p169.
5. Ho *et al.,* Structure of the arginine methyltransferase PRMT5-MEP50 reveals a mechanism for substrate specificity. *PLoSOne* **2013**, 8(2):e57008; Burgos *et al.*, Histone H2A and H4 N-terminal tails are positioned by the MEP50 WD repeat protein for efficient methylation by the PRMT5 arginine methyltransferase. *J. Biol. Chem.* **2015**, 290, 9674
6. Rauh et al., An acetylome peptide microarray reveals specificities and deacetylation substrates for all human sirtuin isoforms. *Nat Commun.* **2013**, 4,2327.

* Correspondence should be addressed to Ulf Reimer: reimer@jpt.com

www.jpt.com

Innovative Peptide Solutions