CellPress

Molecular Cell

Perspective

A needed nomenclature for nucleosomes

Michael-Christopher Keogh,[1] Genevieve Almouzni,[2] Andrew J. Andrews,[3] Karim-Jean Armache,[4] Cheryl H. Arrowsmith,[5,6] Sung Hee Baek,[7] Mark T. Bedford,[8] Emily Bernstein,[9] Jacques Côté,[10] Yael David,[11] John M. Denu,[12] Beat Fierz,[13] Benjamin A. Garcia,[14] Karen C. Glass,[15] Or Gozani,[16] Kristian Helin,[17] Steven Henikoff,[18,19] Ole N. Jensen,[20] Steven Z. Josefowicz,[21] Neil L. Kelleher,[22] Tatiana G. Kutateladze,[23] Herbert H. Lindner,[24]

(Author list continued on next page)

1EpiCypher Inc, Durham, NC 27709, USA
2Institut Curie, PSL Research University, CNRS, Sorbonne Université, Nuclear Dynamics Unit, Equipe Labellisée Ligue Contre le Cancer, 75005 Paris, France
3University of North Carolina at Wilmington, Wilmington, NC 28403, USA
4Department of Biochemistry and Molecular Pharmacology, New York University Grossman School of Medicine, New York, NY 10016, USA
5Princess Margaret Cancer Centre, University Health Network, Toronto, ON M5G 1L7, Canada
6Structural Genomics Consortium and Department of Medical Biophysics, University of Toronto, Toronto, ON M5G 1L7, Canada
7Creative Research Initiatives Center for Epigenetic Code and Diseases, School of Biological Sciences, Seoul National University, Seoul 08826, South Korea
8Department of Epigenetics & Molecular Carcinogenesis, University of Texas MD Anderson Cancer Centre, Houston, TX 77230, USA
9Department of Dermatology, Department of Oncological Sciences and Tisch Cancer Institute, Icahn School of Medicine at Mount Sinai, New York, NY 10029, USA
10St-Patrick Research Group in Basic Oncology, Oncology Division of CHU de Québec-Université Laval Research Center, Laval University Cancer Research Center, Quebec City, Québec G1J 0J9, Canada
11Memorial Sloan Kettering Cancer Center, New York, NY 10065, USA
12Wisconsin Institute of Discovery, University of Wisconsin-Madison, Madison, WI 53706, USA
13SB ISIC LCBM, École Polytechnique Fédérale de Lausanne (EPFL), Station 6, CH-1015 Lausanne, Switzerland
14Department of Biochemistry and Molecular Biophysics, Washington University School of Medicine, St. Louis, MO 63110, USA
15Department of Pharmacology and University of Vermont Cancer Center, Larner College of Medicine, University of Vermont, Burlington, VT 05405, USA
16Department of Biology, Stanford University, Stanford, CA 94305, USA
17Institute of Cancer Research, London SW3 6JB, UK
18Basic Sciences Division, Fred Hutchinson Cancer Center, Seattle, WA 98109, USA
19Howard Hughes Medical Institute, Chevy Chase, MD 20815, USA
20Department of Biochemistry and Molecular Biology, University of Southern Denmark, 5230 Odense, Denmark
21Immunology and Microbial Pathogenesis Program and Department of Pathology and Laboratory Medicine, Weill Cornell Medical College, New York, NY 10065, USA

(Affiliations continued on next page)

SUMMARY

Histone post-translational modifications (PTMs) are crucial to eukaryotic genome regulation, with a range of reported functions and mechanisms of action. Though often studied individually, it has long been recognized that the modifications function by combinatorial synergy or antagonism. Interplay may involve PTMs on the same histone, within the same nucleosome (containing a histone octamer), or between nucleosomes in higher-order chromatin. Given this, the field must distinguish ever greater complexity, and the context in which it is studied, with brevity and precision. The proteoform was introduced to define individual forms of a protein by sequence and PTMs, followed by the nucleoform to describe the particular gathering of histones within an individual nucleosome. There is now a need to define specific forms of these entities in prose while providing space for experimental nuance. To this end, we introduce a nomenclature that can express discrete PTMs, proteoforms, nucleoforms, or situations where defined PTMs exist in an uncertain context. Though specifically designed for the chromatin field, adaptions of the framework could be used to describe—and thus dissect—how proteoforms are configured in functionally distinct complexes across biology.

INTRODUCTION

The core histones (H2A, H2B, H3, and H4) package the genome into nucleosomes and are a central component of chromatin biology.[1] Histone function is regulated by diverse post-translational modifications (PTMs) that operate in combination to transduce signals and recruit additional enzymatic activities, thus impacting DNA replication, transcription, and

Chao Lu,[25] Karolin Luger,[19,26] Parag Mallick,[27] Catherine A. Musselman,[28] Tom W. Muir,[29] Ljiljana Paša-Tolić,[30] Robert Schneider,[31] Xiaobing Shi,[32] Yang Shi,[33] Simone Sidoli,[34] Lloyd M. Smith,[35] Jessica K. Tyler,[36] Cynthia Wolberger,[37] Jerry L. Workman,[38] Brian D. Strahl,[39] and Nicolas L. Young[40,*]

[22]Departments of Chemistry and Molecular Biosciences, the Chemistry of Life Processes Institute, and the Proteomics Center of Excellence, Northwestern University, Evanston, IL 60208, USA
[23]Department of Pharmacology, University of Colorado School of Medicine, Aurora, CO 80045, USA
[24]Institute of Clinical Biochemistry, Biocenter, Medical University of Innsbruck, Innsbruck, Austria
[25]Department of Genetics and Development and Herbert Irving Comprehensive Cancer Center, Columbia University Irving Medical Center, New York, NY 10032, USA
[26]Department of Biochemistry, University of Colorado at Boulder, Boulder, CO 80309, USA
[27]Canary Center for Cancer Early Detection, School of Medicine, Stanford University, Stanford, CA 94305, USA
[28]Department of Biochemistry and Molecular Genetics, University of Colorado Anschutz Medical Campus, Aurora, CO 80045, USA
[29]Department of Chemistry, Princeton University, Princeton, Princeton, NJ 08544, USA
[30]Pacific Northwest National Laboratory, Richland, WA 99352, USA
[31]Institute of Functional Epigenetics, Helmholtz Zentrum München, Neuherberg, Germany
[32]Department of Epigenetics, Van Andel Institute, Grand Rapids, MI 49503, USA
[33]Ludwig Institute for Cancer Research, Nuffield Department of Medicine, University of Oxford, Oxford OX3 7DQ, UK
[34]Department of Biochemistry, Albert Einstein College of Medicine, Bronx, New York, NY 10461, USA
[35]Department of Chemistry, University of Wisconsin-Madison, Madison, WI 53706, USA
[36]Department of Pathology and Laboratory Medicine, Weill Cornell Medical College, New York, NY 10065, USA
[37]Department of Biophysics & Biophysical Chemistry, Johns Hopkins University School of Medicine, Baltimore, MD 21205, USA
[38]Stowers Institute for Medical Research, Kansas City, MO 64110, USA
[39]Department of Biochemistry and Biophysics and Lineberger Comprehensive Cancer Center, University of North Carolina at Chapel Hill, Chapel Hill, NC 27599, USA
[40]Verna & Marrs McLean Department of Biochemistry & Molecular Pharmacology and Department of Molecular and Cellular Biology, Baylor College of Medicine, Houston, TX 77030, USA
*Correspondence: nicolas.young@bcm.edu
https://doi.org/10.1016/j.molcel.2025.08.029

repair.[2] Discussions of this biology have been enabled by the 2005 Brno nomenclature[3] that communicates histone residue-position-modification type (e.g., H3K4me3 or H3K9ac): an intuitive clarity only slightly undermined by core histone proteins being canonically numbered from the second residue, as the initiator methionine is absent in the mature polypeptide (e.g., human H3.1 [UniProtKB P68431]: MA_1RTKQTARK …). We have an ever-growing capability to study function in the histone, nucleosome, or chromatin context, so the field now needs an expanded nomenclature to describe each element and any potential limits of related experimental datasets.

The nature of chromatin complexity (and how it is studied)

"Proteoform" was coined in 2012 to concisely describe individual protein molecules, and can encompass the differential products of transcriptional initiation or splicing, mutation, post-translational processing, modification, etc.[4,5] Thus, the physiological state of each "protein" is an ensemble of "proteoforms," with distinct forms potentially mediating distinct functions. Collated proteoforms in tabular formats are useful for mass spectrometry (MS)-focused efforts,[6] but do not work well to communicate biological function or mechanism and thus enable hypothesis-driven research.[7] The need for a concise and consistent proteoform notation is particularly urgent in chromatin biology. Histone proteoforms are often hypermodified (Figure 1). Thus, substrate engagement for reading, writing, erasing, or clipping most often occurs in the context of high heterogeneity.[1] The consideration of histone mutants (e.g., H3K27M associated with pediatric gliomas[8]), variants (e.g., centromere defining CenH3[9]), and variant-specific PTMs (e.g., γH2A.X associated with DNA double-strand breaks[10]) provides additional needs for clarity. Finally, this diversity is further encompassed within nucleosomes (each a core histone octamer wrapped by 147 bp DNA[11]), necessitating description of the "nucleoform."[12]

Top- and middle-down proteomics allow the quantitative study of intact proteoforms and can describe how PTMs co-occur on a single histone (in *cis*).[13–22] Advances in engineering fully defined histones and nucleosomes have enabled the dissection of PTM interplay in *cis* or *trans* (the latter within the same nucleosome but on different histones)[15,23–25] or between PTMs and other nucleosome surfaces (e.g., the "acidic patch" or wrapping DNA).[26–29] Emerging technologies, such as NanoPore protein sequencing[30] or Nuc-MS,[12] promise (but cannot yet routinely deliver) the complete characterization of native histones or even nucleosomes. The production of fully defined nucleosome arrays is already at a practical stage and stepping toward synthetic chromatin. Further inquiry along these many axes will accelerate as novel reagents and technologies become more widely available.

Elucidating chromatin biology with nucleoform-level information would be ideal. This is rarely possible, so any methodological limits on the gathered data should be considered. As an example, chromatin immunoprecipitation sequencing (ChIP-seq) and cleavage under targets & release using nuclease (CUT&RUN) are regularly used to map the genomic localization of individual PTMs in cell populations (Figure 2). The overlapping genomic localization of two PTMs in parallel experiments could indicate co-localization (whether co-incident or co-regulated)

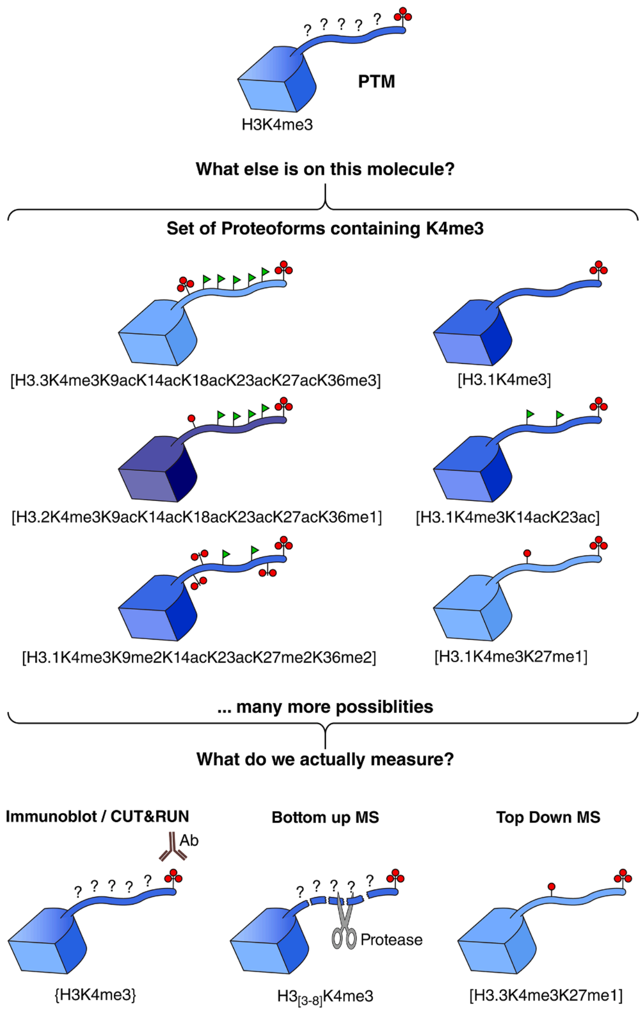


Figure 1. Most studies of histone PTMs and their role in genome regulation address the modifications as an attribute of a molecule, without knowledge of any other attributes

H3K4me3 is enriched at transcriptionally active gene promoters but found in many proteoforms, each with implications for PTM function. Despite this, most experimental approaches do not speak to the proteoform. Immunoblots or CUT&RUN uses antibodies that detect a single PTM and almost never its relationship with even adjacent modifications. Bottom-up proteomics can capture proximal PTMs on the same peptide but generally loses any connections to distal PTMs. Top-down MS methods can unambiguously determine the full proteoform, including variant identity and PTM co-occurrence. The new nomenclature addresses these concepts and any related experimental ambiguities.

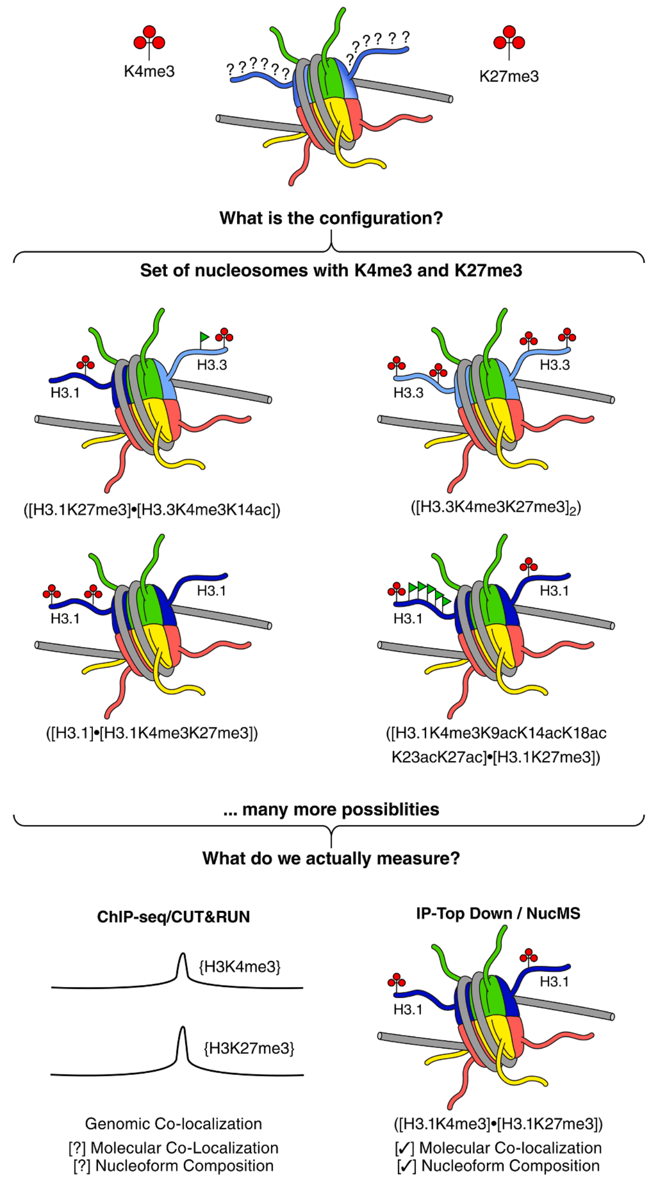
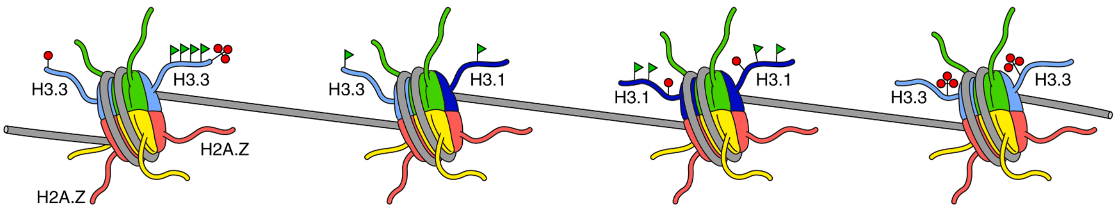


Figure 2. The configuration of histone PTMs and variants within the nucleoform can be essential to function but requires a means to effectively communicate molecular detail and data

In this example, a native nucleosome sample (blue, H3; green, H4; red, H2A; yellow, H2B) contains H3K4me3 and H3K27me3, but this could represent very different configurations (e.g., the two PTMs together in *trans* form the canonical "bivalent" signature[22]). Current methods to interrogate bulk samples are either unable to distinguish (e.g., immunoblot, ChIP-seq, and CUT&RUN), can discriminate some possibilities (e.g., sequential ChIP and middle-down MS), or have the potential to be definitive (e.g., combinatorial enrichment and NucMS[12,31]). The proposed nomenclature is designed to express the (un)certainty of the starting options and what the experimental data can(not) discriminate. Of note, it is now possible to engineer fully defined proteoforms and nucleoforms to support biochemical and genomic studies with unparalleled mechanistic insight.[25,26,32,33] Such reagents also need to be concisely described in prose.

at the histone or nucleosome level or mutually independent populations. The correct interpretation is essential to understand the related biology. Although bottom-up MS can inform on PTM co-localization on the same molecule, this is limited to the length of the proteolytic peptide and often does not resolve histone variant identity. Top-down MS can describe PTM co-localization in *cis* but not relationships in *trans*. Being able to specify the extent of knowledge about a peptide or protein, and to define knowns and ambiguities in an experimental dataset, is an essential aspect of the proposed notation.

Combinatorial mechanisms in chromatin biology are most often realized through the nucleosome.[11] This extends regulatory potential (and descriptive need) beyond the proteoform. In *cis*



([H2A.Z]$_2$•[H3.3K4me1]•[H3.3K4me3K9acK14acK18acK23ac]) ([H3.1K14ac]•[H3.3K14ac]) ([H3.1K14acK23acK36me1]$_2$) ([H3.3K36me3]$_2$)

Figure 3. An array of semi-synthetic nucleosomes with defined properties
Nucleosome #1 (blue, H3; green, H4; red, H2A; yellow, H2B) is homotypic for both H2A.Z and H3.3 (one copy as K4me1 only; the second K4me3, K9ac, K14ac, K18ac, and K23ac).
Nucleosome #2 is heterotypic for H3.1 and H3.3, with each modified at K14ac.
Nucleosome #3 is homotypic for H3.1, containing K14ac, K23ac, and K36me1.
Finally, nucleosome #4 is homotypic for H3.3K36me3. In each nucleosome, the other core histones are not noted and can thus be considered unmodified major histones. This complexity can be expressed in prose with precision as demonstrated beneath the figure. Importantly, a computer-readable derivative of the nomenclature could be parsed for downstream applications such as figure generation.

histone mechanisms indicate where two or more attributes co-occur and mediate function on the same molecule within the same nucleosome. In *trans* mechanisms are where two or more attributes on different molecules work in concert within a single nucleosome. Further increasing complexity, two different proteoforms of the same histone can co-occupy a "heterotypic" (aka "asymmetric") nucleosome—entities of great interest for the regulatory potential they provide.[34–36] Of note, homo-/heterotypic is our preferred designation over a-/symmetric because it follows the histone variant literature[37,38] and avoids confusion with hemi-/symmetric DNA methylation. Finally, trans-nucleosomal mechanisms are where attributes from neighboring or spatially co-localized (such as by "looping") nucleosomes function together. It is now possible to synthesize fully defined histones, nucleosomes, or even nucleosome arrays and study the regulatory potential at each level.

The proposed nomenclature

A nomenclature that can communicate all the above concepts defines proteoforms (Figure 1), their combination in a nucleoform (Figure 2), and the various nucleoforms in chromatin arrays (Figure 3). Further, it conveys effectively in prose, has intuitive biological meaning, can accommodate experimental imprecision, and enables a nuanced discussion of mechanisms and hypotheses. Finally, it does all this while encompassing the Brno nomenclature[3] that efficiently describes histone PTMs. Below, we provide some examples of the nomenclature in practice, but for more details, see Table 1.

Square brackets "[H3K4me3]" indicate a fully defined proteoform, as in a semi-synthetic histone,[15,25] where other sites of potential modification not denoted can be understood as definitively unmodified. Braces "{H3K4me3}" indicate a partially understood native proteoform, where one PTM is experimentally known to exist (as by immunoprecipitating H3K4me3 from a cell extract), but other sites of potential modification not denoted are understood to be of undefined status: the general situation with native material. However, if braces are not used, it should be assumed to be referring to discrete PTMs (e.g., H3K4me3 represents {H3K4me3}) because this is historical practice. If identifiable, the histone variant carrying the PTM is denoted immediately within the open bracket or braces: e.g., [H3.3K4me3]. If the species is a histone peptide, such as a synthetic reagent or measured by bottom-up proteomics, this is indicated by subscripted brackets identifying the start and end, separated by a hyphen: e.g., H3$_{[1-12]}$K4me3.

Parentheses "( )" are used to group proteoforms in a nucleosome, with "·" separating any distinguishing histones in alphabetical and numerical order. It is now possible to make semi-synthetic nucleosomes as fully defined homo- or heterotypics,[15,25,40] with ([H3K27me3]$_2$) a homotypic, and the ([H3]·[H3K4me3K14ac]) or ([H3K4me3]·[H3K14ac]) heterotypics respectively distributing the same PTM complement in *cis* or *trans*. The situation with native material (such as that enriched by a PTM-specific antibody) is more complex, and thus ({H3K27me3}) does not inform on the homo-/heterotypic state because these generally cannot be distinguished. For the advanced user, chromatin arrays can be expressed as a series of individual nucleosomes (optionally with linker DNA information: Table 1) and follow the 5′ to 3′ standard from left to right (Figure 3).

Discussing hypotheses and mechanisms

The capacity to concisely define histones and nucleosomes with rigor also allows the effective communication of nuanced hypotheses. The nomenclature readily supports the use of relational operators (>, <, or =) for comparing abundances or activities. For example, H3K4me3 is most commonly found with four or five *cis* acetyllysines and rarely in their absence,[15] thus the abundance of [H3K4me3K9acK14acK18acK23ac] >> [H3K4me3]. Figure 4 depicts the experimental formats used to resolve the multivalent nature of ZMYND11 (BS69): histone binding.[41] Although such graphical models are informative, there remains a need to concisely present distinct observations in prose. In this particular case, observed experimental binding of the ZMYND11 (Bromo-ZnF-PWWP) domain triplet is H3.1$_{[19-42]}$K36me3 < H3.3$_{[19-42]}$K36me3 <<< ([H3.3K36me3]$_2$).[41,42] This communicates that ZMYND11 binds K36me3 on histone variant H3.3 (containing S31 vs. the A31 of H3.1), but such binding is dramatically enhanced in the nucleosome vs. peptide context (because the PWWP domain synergistically engages DNA).

Table 1. Nomenclature examples

Type of information	Delimiters/operators	Example	Example description
A defined proteoform (semi-synthetic or top-down proteomics). Absence of a residue communicates not modified.	[ ]	[H3.2K27me3]	A fully defined histone H3.2 containing K27me3. All other sites of variable modification not noted are unmodified.
		[H3.2K14acK23ac]	A fully defined H3.2 containing only K14ac and K23ac.
		[H3.2]	A fully defined and completely unmodified H3.2.
Discrete PTM (usually native material, bottom-up proteomics, or marginalized top-down data). Absence of a residue communicates ambiguity.	None or { }	H3K27me3 or {H3K27me3}	A set of H3 molecules of unspecified isoform that contain K27me3.
		H3K27un or {H3K27un}	A set of H3 molecules that are unmodified at K27.
		H3.2 or {H3.2}	A set of H3.2 molecules in a completely undefined modification state.
		H3.2K14acK23ac or {H3,2K14acK23ac}	A set of H3.2 molecules that contain K14ac and K23ac.
		H2AXS139phY142ph or H2AXS139phY142ph	A set of H2AX molecules phosphorylated at both S139 and Y142.
A nucleosome. Semi-synthetic and fully defined. Absence of a histone family communicates major and unmodified.	( ) or (·)	([H3.2K14acK18acK23ac]$_2$)	A fully defined nucleosome homotypic for histone H3.2 acetylated at K14, K18, and K23. The other histones are major and unmodified.
		([H3.2]·[H3.2K14acK18acK23ac])	A defined nucleosome heterotypic for H3.2. One molecule is unmodified. The other is acetylated at K14, K18, and K23.
		([H3.2K14ac]·[H3.3K14ac])	A defined nucleosome heterotypic for H3. One molecule is major H3.2. The other is variant H3.3. Both are acetylated at K14.
'Native' material. Absence of a histone family communicates ambiguity		({H3K14ac})	A set of nucleosomes with at least one H3K14ac. All other sites of variable modification on H3 are undefined. The other histones are undefined.
		({H3K9me3K26me3})	A set of nucleosomes containing H3K9me3 and H3K36me3. The PTMs could be in *cis* or *trans*. Such materiaal could be enriched by sequential immunoprecipitation or with a combinatorial PTM reader.[39]
Nucleosome arrays	(·) (·)	([H3.3K4me3]$_2$) ([H3.2K14ac]$_2$) ([H3.2]·[H3.3])	A defined tri-nucleosome with unspecified DNA linker length. The most 5′ nucleosome is a [H3.3K4me3] homotypic. The second nucleosome is a [H3.2K14ac] homotypic. The most 3′ nucleosome is heterotypic for unmodified H3.2 and H3.3.
		([H3.3]$_2$)50([H2AZ.1]$_2$)	A defined di-nucleosome separated by 50 bp linker DNA. The 5′ nucleosome is homotypic for H3.3 (with all others in the octamer major histones as default). The 3′ nucleosome is homotpyic for H2AZ.1 (with all others in the octamer major histones as default).

(*Continued on next page*)

Table 1. *Continued*

Type of information	Delimiters/operators	Example	Example description
Not	!	{H3.2!K27me3}	A set of H3.2 molecules that do not contain K27me3.
		{H3.2!K27me3K36me3}	A set of H3.2 molecules that contain K36me3 but not K27me3.

DISCUSSION

The proposed nomenclature is consistent with current usage and requires little to no change for most common applications yet provides clarity and brevity for increasingly complex data and concepts. It can effectively describe a specific histone proteoform [H3K14acK18acK23acK27me3]. But it can also describe classes of proteoforms with shared attributes {H3K14acK27me3} that might represent a functional unit found in multiple proteoforms, including [H3K14acK18acK23acK27me3]. It can readily distinguish nucleosomes: homotypic ([H3K4me3 K27me3]$_2$) and *cis* or *trans* heterotypic ([H3]·[H3K4me3K27me3]) vs. ([H3K4me3]·[H3K27me3]). It can express functional relationships, such as [H4K20me2] is the preferred substrate of acetyltransferases writing H4K16ac, with the primary acetylated product [H4K16acK20me2][14] and subsequent observed states, including [H4K12acK16acK20me2] and [H4K8acK12acK16acK 20me2].

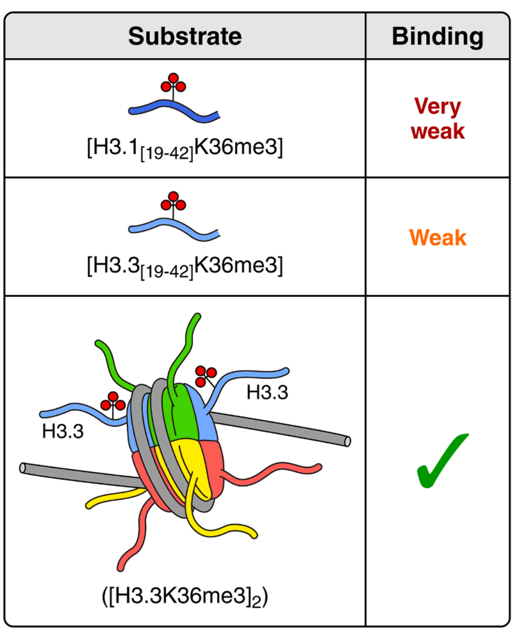


Figure 4. Molecular context impacts ZMYND11 (BD-ZnF-PWWP) binding

Blue: H3; green: H4; red: H2A; yellow: H2B. Weak engagement is observed with H3.3$_{[19-42]}$K36me3 over H3.1$_{[19-42]}$K36me3 peptides. However, binding is dramatically enhanced on ([H3.3K36me3]$_2$) nucleosomes. On a mechanistic level, ZMYND11 PWWP is co-operatively engaging H3.3S31un, the K36me3 PTM and nucleosomal DNA.[41] The further complexity—that this binding is inhibited by S31 phosphorylation, as in ({H3.3S31phK36me3})[43,44]—is not depicted.

The new nomenclature can communicate experimental nuance (and potential limitations): H3K4un vs. H3K4me3 vs. H3$_{[1-12]}$K4me3 vs. [H3.2K4me3] vs. ({H3K4me3}) vs. ([H3.2K4me3]$_2$) vs. ([H3.2]·[H3.2K4me3]), representing an unmodified amino acid, a PTM, a modified peptide, a fully defined proteoform, a native nucleosome containing a PTM (such as enriched via the same), a homotypic nucleoform, and a heterotypic nucleoform. The language of histone PTMs has always been considered to operate by combinatorial means,[2,26] and this nomenclature enables the user to posit hypotheses for such investigations, which can then be experimentally tested with defined proteoforms and nucleoforms. Ultimately, the nomenclature provides a clear path to even greater molecular specificity, such as additions to communicate species, linker length, linker histones, DNA modifications, or non-canonical nucleosomes (e.g., tetrasomes or hexasomes[45,46]) (Table 1). However, the initial version is purposely focused on the immediate need to accurately describe core histone proteoforms, variants, oncomutants, PTMs, and nucleoforms in chromatin and thus to enable the nuanced discussion of complex concepts and experimental data.

Finally, combinatorial modifications and differential associations are central to normal function and its dysregulation, from EGFR in cell growth and cancer to Tau in axonal transport and neurodegeneration.[6,47–50] The facile and expandable framework of our chromatin-focused nomenclature will provide a basis for colleagues in other fields to describe, and thus dissect, how proteoforms could be configured in functionally distinct complexes across biology.

ACKNOWLEDGMENTS

The figures were created by Patrick Lane, ScEYEnceStudios.

DECLARATION OF INTERESTS

M.-C.K., M.T.B., O.G., and B.D.S. are board members of *EpiCypher Inc.*, a commercial developer of reagents and platforms for chromatin studies.

REFERENCES

1. Taylor, B.C., and Young, N.L. (2021). Combinations of histone post-translational modifications. Biochem. J. *478*, 511–532. https://doi.org/10.1042/BCJ20200170.

2. Strahl, B.D., and Allis, C.D. (2000). The language of covalent histone modifications. Nature *403*, 41–45. https://doi.org/10.1038/47412.

3. Turner, B.M. (2005). Reading signals on the nucleosome with a new nomenclature for modified histones. Nat. Struct. Mol. Biol. *12*, 110–112. https://doi.org/10.1038/nsmb0205-110.

4. Aebersold, R., Agar, J.N., Amster, I.J., Baker, M.S., Bertozzi, C.R., Boja, E.S., Costello, C.E., Cravatt, B.F., Fenselau, C., Garcia, B.A., et al.

(2018). How many human proteoforms are there? Nat. Chem. Biol. *14*, 206–214. https://doi.org/10.1038/nchembio.2576.

5. Smith, L.M., and Kelleher, N.L.; Consortium for Top Down Proteomics (2013). Proteoform: a single term describing protein complexity. Nat. Methods *10*, 186–187. https://doi.org/10.1038/nmeth.2369.

6. Hollas, M.A.R., Robey, M.T., Fellers, R.T., LeDuc, R.D., Thomas, P.M., and Kelleher, N.L. (2022). The Human Proteoform Atlas: a FAIR community resource for experimentally derived proteoforms. Nucleic Acids Res. *50*, D526–D533. https://doi.org/10.1093/nar/gkab1086.

7. LeDuc, R.D., Schwämmle, V., Shortreed, M.R., Cesnik, A.J., Solntsev, S.K., Shaw, J.B., Martin, M.J., Vizcaino, J.A., Alpi, E., Danis, P., et al. (2018). ProForma: A Standard Proteoform Notation. J. Proteome Res. *17*, 1321–1325. https://doi.org/10.1021/acs.jproteome.7b00851.

8. Schwartzentruber, J., Korshunov, A., Liu, X.-Y., Jones, D.T.W., Pfaff, E., Jacob, K., Sturm, D., Fontebasso, A.M., Quang, D.-A.K., Tönjes, M., et al. (2012). Driver mutations in histone H3.3 and chromatin remodelling genes in paediatric glioblastoma. Nature *482*, 226–231. https://doi.org/10.1038/nature10833.

9. Palmer, D.K., O'Day, K., Wener, M.H., Andrews, B.S., and Margolis, R.L. (1987). A 17-kD centromere protein (CENP-A) copurifies with nucleosome core particles and with histones. J. Cell Biol. *104*, 805–815. https://doi.org/10.1083/jcb.104.4.805.

10. Fillingham, J., Keogh, M.-C., and Krogan, N.J. (2006). GammaH2AX and its role in DNA double-strand break repair. Biochem. Cell Biol. *84*, 568–577. https://doi.org/10.1139/o06-072.

11. Luger, K., Mäder, A.W., Richmond, R.K., Sargent, D.F., and Richmond, T.J. (1997). Crystal structure of the nucleosome core particle at 2.8 Å resolution. Nature *389*, 251–260. https://doi.org/10.1038/38444.

12. Schachner, L.F., Jooß, K., Morgan, M.A., Piunti, A., Meiners, M.J., Kafader, J.O., Lee, A.S., Iwanaszko, M., Cheek, M.A., Burg, J.M., et al. (2021). Decoding the protein composition of whole nucleosomes with Nuc-MS. Nat. Methods *18*, 303–308. https://doi.org/10.1038/s41592-020-01052-9.

13. Wang, T., Holt, M.V., and Young, N.L. (2018). The histone H4 proteoform dynamics in response to SUV4-20 inhibition reveals single molecule mechanisms of inhibitor resistance. Epigenetics Chromatin *11*, 29. https://doi.org/10.1186/s13072-018-0198-9.

14. Wang, T., Holt, M.V., and Young, N.L. (2018). Early Butyrate Induced Acetylation of Histone H4 is Proteoform Specific and Linked to Methylation State. Epigenetics *13*, 519–535. https://doi.org/10.1080/15592294.2018.1475979.

15. Jain, K., Marunde, M.R., Burg, J.M., Gloor, S.L., Joseph, F.M., Poncha, K.F., Gillespie, Z.B., Rodriguez, K.L., Popova, I.K., Hall, N.W., et al. (2023). An acetylation-mediated chromatin switch governs H3K4 methylation read-write capability. eLife *12*, e82596. https://doi.org/10.7554/eLife.82596.

16. Taylor, B.C., and Young, N.L. (2023). Histone H4 proteoforms and post-translational modifications in the Mus musculus brain with quantitative comparison of ages and brain regions. Anal. Bioanal. Chem. *415*, 1627–1639. https://doi.org/10.1007/s00216-023-04555-4.

17. Chen, Y., Hoover, M.E., Dang, X., Shomo, A.A., Guan, X., Marshall, A.G., Freitas, M.A., and Young, N.L. (2016). Quantitative Mass Spectrometry Reveals that Intact Histone H1 Phosphorylations are Variant Specific and Exhibit Single Molecule Hierarchical Dependence. Mol. Cell. Proteomics *15*, 818–833. https://doi.org/10.1074/mcp.M114.046441.

18. Young, N.L., DiMaggio, P.A., Plazas-Mayorca, M.D., Baliban, R.C., Floudas, C.A., and Garcia, B.A. (2009). High Throughput Characterization of Combinatorial Histone Codes. Mol. Cell. Proteomics *8*, 2266–2284. https://doi.org/10.1074/mcp.M900238-MCP200.

19. Dang, X., Scotcher, J., Wu, S., Chu, R.K., Tolić, N., Ntai, I., Thomas, P.M., Fellers, R.T., Early, B.P., Zheng, Y., et al. (2014). The first pilot project of the consortium for top-down proteomics: A status report. Proteomics *14*, 1130–1140. https://doi.org/10.1002/pmic.201300438.

20. Dang, X., Singh, A., Spetman, B.D., Nolan, K.D., Isaacs, J.S., Dennis, J.H., Dalton, S., Marshall, A.G., and Young, N.L. (2016). Label-Free Relative Quantitation of Isobaric and Isomeric Human Histone H2A and H2B Variants by Fourier Transform Ion Cyclotron Resonance Top-Down MS/MS. J. Proteome Res. *15*, 3196–3203. https://doi.org/10.1021/acs.jproteome.6b00414.

21. Holt, M.V., Wang, T., and Young, N.L. (2019). High-Throughput Quantitative Top-Down Proteomics: Histone H4. J. Am. Soc. Mass Spectrom. *30*, 2548–2560. https://doi.org/10.1007/s13361-019-02350-z.

22. Voigt, P., LeRoy, G., Drury, W.J., III, Zee, B.M., Son, J., Beck, D.B., Young, N.L., Garcia, B.A., and Reinberg, D. (2012). Asymmetrically Modified Nucleosomes. Cell *151*, 181–193. https://doi.org/10.1016/j.cell.2012.09.002.

23. Marunde, M.R., Popova, I.K., Weinzapfel, E.N., and Keogh, M.-C. (2022). The dCypher Approach to Interrogate Chromatin Reader Activity Against Posttranslational Modification-Defined Histone Peptides Histone peptides and Nucleosomes. In Chromatin: Methods and Protocols, J. Horsfield and J. Marsman, eds. (Springer), pp. 231–255. https://doi.org/10.1007/978-1-0716-2140-0_13.

24. Morgan, M.A.J., Popova, I.K., Vaidya, A., Burg, J.M., Marunde, M.R., Rendleman, E.J., Dumar, Z.J., Watson, R., Meiners, M.J., Howard, S.A., et al. (2021). A trivalent nucleosome interaction by PHIP/BRWD2 is disrupted in neurodevelopmental disorders and cancer. Genes Dev. *35*, 1642–1656. https://doi.org/10.1101/gad.348766.121.

25. Marunde, M.R., Fuchs, H.A., Burg, J.M., Popova, I.K., Vaidya, A., Hall, N.W., Weinzapfel, E.N., Meiners, M.J., Watson, R., Gillespie, Z.B., et al. (2024). Nucleosome conformation dictates the histone code. eLife *13*, e78866. https://doi.org/10.7554/eLife.78866.

26. Weinzapfel, E.N., Fedder-Semmes, K.N., Sun, Z.-W., and Keogh, M.-C. (2024). Beyond the tail: the consequence of context in histone post-translational modification and chromatin research. Biochem. J. *481*, 219–244. https://doi.org/10.1042/BCJ20230342.

27. Hicks, C.W., Rahman, S., Gloor, S.L., Fields, J.K., Husby, N.L., Vaidya, A., Maier, K.E., Morgan, M., Keogh, M.-C., and Wolberger, C. (2024). Ubiquitinated histone H2B as gatekeeper of the nucleosome acidic patch. Nucleic Acids Res. *52*, 9978–9995. https://doi.org/10.1093/nar/gkae698.

28. Thomas, J.F., Valencia-Sánchez, M.I., Tamburri, S., Gloor, S.L., Rustichelli, S., Godínez-López, V., De Ioannes, P., Lee, R., Abini-Agbomson, S., Gretarsson, K., et al. (2023). Structural basis of histone H2a lysine 119 deubiquitination by Polycomb Repressive Deubiquitinase BAP1/ASXL1. Preprint at bioRxiv. https://doi.org/10.1101/2023.02.23.529554.

29. Weinberg, D.N., Rosenbaum, P., Chen, X., Barrows, D., Horth, C., Marunde, M.R., Popova, I.K., Gillespie, Z.B., Keogh, M.-C., Lu, C., et al. (2021). Two competing mechanisms of DNMT3A recruitment regulate the dynamics of de novo DNA methylation at PRC1-targeted CpG islands. Nat. Genet. *53*, 794–800. https://doi.org/10.1038/s41588-021-00856-5.

30. Oukhaled, G., Mathé, J., Biance, A.L., Bacri, L., Betton, J.M., Lairez, D., Pelta, J., and Auvray, L. (2007). Unfolding of proteins and long transient conformations detected by single nanopore recording. Phys. Rev. Lett. *98*, 158101. https://doi.org/10.1103/PhysRevLett.98.158101.

31. Jooß, K., Schachner, L.F., Watson, R., Gillespie, Z.B., Howard, S.A., Cheek, M.A., Meiners, M.J., Sobh, A., Licht, J.D., Keogh, M.-C., et al. (2021). Separation and Characterization of Endogenous Nucleosomes by Native Capillary Zone Electrophoresis-Top-Down Mass Spectrometry. Anal. Chem. *93*, 5151–5160. https://doi.org/10.1021/acs.analchem.0c04975.

32. David, Y., and Muir, T.W. (2017). Emerging Chemistry Strategies for Engineering Native Chromatin. J. Am. Chem. Soc. *139*, 9090–9096. https://doi.org/10.1021/jacs.7b03430.

33. Shah, R.N., Grzybowski, A.T., Cornett, E.M., Johnstone, A.L., Dickson, B.M., Boone, B.A., Cheek, M.A., Cowles, M.W., Maryanski, D., Meiners, M.J., et al. (2018). Examining the Roles of H3K4 Methylation States with Systematically Characterized Antibodies. Mol. Cell *72*, 162–177.e7. https://doi.org/10.1016/j.molcel.2018.08.015.

34. Valsakumar, D., and Voigt, P. (2024). Nucleosomal asymmetry: a novel mechanism to regulate nucleosome function. Biochem. Soc. Trans. *52*, 1219–1232. https://doi.org/10.1042/BST20230877.

35. López, V.G., Valencia-Sánchez, M.I., Abini-Agbomson, S., Thomas, J.F., Lee, R., De Ioannes, P., Sosa, B.A., Armache, J.-P., and Armache, K.-J.

(2024). Read-write mechanisms of H2A ubiquitination by Polycomb repressive complex 1. Nature *636*, 755–761. https://doi.org/10.1038/s41586-024-08183-5.

36. Bryan, E., Valsakumar, D., Idigo, N.J., Warburton, M., Webb, K.M., McLaughlin, K.A., Spanos, C., Lenci, S., Major, V., Ambrosi, C., et al. (2025). Nucleosomal asymmetry shapes histone mark binding and promotes poising at bivalent domains. Mol. Cell *85*, 471–489.e12. https://doi.org/10.1016/j.molcel.2024.12.002.

37. Nekrasov, M., and Tremethick, D.J. (2021). Sequential Chromatin Immunoprecipitation to Identify Heterotypic Nucleosomes. Methods Mol. Biol. *2351*, 147–161. https://doi.org/10.1007/978-1-0716-1597-3_8.

38. Bernstein, E., and Hake, S.B. (2006). The nucleosome: a little variation goes a long way. Biochem. Cell Biol. *84*, 505–517. https://doi.org/10.1139/o06-085.

39. Mauser, R., Kungulovski, G., Keup, C., Reinhardt, R., and Jeltsch, A. (2017). Application of dual reading domains as novel reagents in chromatin biology reveals a new H3K9me3 and H3K36me2/3 bivalent chromatin state. Epigenetics Chromatin *10*, 45. https://doi.org/10.1186/s13072-017-0153-1.

40. Lechner, C.C., Agashe, N.D., and Fierz, B. (2016). Traceless Synthesis of Asymmetrically Modified Bivalent Nucleosomes. Angew. Chem. Int. Ed. Engl. *55*, 2903–2906. https://doi.org/10.1002/anie.201510996.

41. Wen, H., Li, Y., Xi, Y., Jiang, S., Stratton, S., Peng, D., Tanaka, K., Ren, Y., Xia, Z., Wu, J., et al. (2014). ZMYND11 links histone H3.3K36me3 to transcription elongation and tumour suppression. Nature *508*, 263–268. https://doi.org/10.1038/nature13045.

42. Wang, J., Qin, S., Li, F., Li, S., Zhang, W., Peng, J., Zhang, Z., Gong, Q., Wu, J., and Shi, Y. (2014). Crystal structure of human BS69 Bromo-ZnF-PWWP reveals its role in H3K36me3 nucleosome binding. Cell Res. *24*, 890–893. https://doi.org/10.1038/cr.2014.38.

43. Guo, R., Zheng, L., Park, J.W., Lv, R., Chen, H., Jiao, F., Xu, W., Mu, S., Wen, H., Qiu, J., et al. (2014). BS69/ZMYND11 reads and connects histone H3.3 lysine 36 trimethylation-decorated chromatin to regulated pre-mRNA processing. Mol. Cell *56*, 298–310. https://doi.org/10.1016/j.molcel.2014.08.022.

44. Armache, A., Yang, S., Martínez de Paz, A., Robbins, L.E., Durmaz, C., Cheong, J.Q., Ravishankar, A., Daman, A.W., Ahimovic, D.J., Klevorn, T., et al. (2020). Histone H3.3 phosphorylation amplifies stimulation-induced transcription. Nature *583*, 852–857. https://doi.org/10.1038/s41586-020-2533-0.

45. Tan, Z.Y., Cai, S., Noble, A.J., Chen, J.K., Shi, J., and Gan, L. (2023). Heterogeneous non-canonical nucleosomes predominate in yeast cells in situ. eLife *12*, RP87672. https://doi.org/10.7554/eLife.87672.

46. Zlatanova, J., Bishop, T.C., Victor, J.-M., Jackson, V., and van Holde, K. (2009). The nucleosome family: dynamic and growing. Structure *17*, 160–171. https://doi.org/10.1016/j.str.2008.12.016.

47. Malik, A., Afaq, S., Alwabli, A.S., and Al-Ghmady, K. (2019). Networking of predicted post-translational modification (PTM) sites in human EGFR. Bioinformation *15*, 448–456. https://doi.org/10.6026/97320630015448.

48. He, Q., Qu, M., Bao, H., Xu, Y., Shen, T., Tan, D., Barkat, M.Q., Xu, C., Zeng, L.-H., and Wu, X. (2023). Multiple post-translational modifications ensure EGFR functionality: Potential therapeutic targets to overcome its drug-resistance mutations. Cytokine Growth Factor Rev. *70*, 41–53. https://doi.org/10.1016/j.cytogfr.2023.03.003.

49. Alquezar, C., Arya, S., and Kao, A.W. (2020). Tau Post-translational Modifications: Dynamic Transformers of Tau Function, Degradation, and Aggregation. Front. Neurol. *11*, 595532. https://doi.org/10.3389/fneur.2020.595532.

50. Barber, K.W., and Rinehart, J. (2018). The ABCs of PTMs. Nat. Chem. Biol. *14*, 188–192. https://doi.org/10.1038/nchembio.2572.